



**HAL**  
open science

# Schémas aux volumes finis avec inégalités d'entropie discrètes pour des systèmes hyperboliques non linéaires

Ludovic Martaud

► **To cite this version:**

Ludovic Martaud. Schémas aux volumes finis avec inégalités d'entropie discrètes pour des systèmes hyperboliques non linéaires. Mathématiques [math]. Nantes Université, 2023. Français. NNT: . tel-04278166

**HAL Id: tel-04278166**

**<https://hal.science/tel-04278166>**

Submitted on 9 Nov 2023

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

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

# THÈSE DE DOCTORAT DE

NANTES UNIVERSITÉ

ÉCOLE DOCTORALE N° 641  
*Mathématiques et Sciences et Technologies  
de l'Information et de la Communication*  
Spécialité : *Mathématiques et leurs Interactions*

Par

**Ludovic MARTAUD**

**Schémas aux volumes finis avec inégalités d'entropie discrètes  
pour des systèmes hyperboliques non linéaires.**

**Thèse présentée et soutenue à Nantes, le 29 Septembre 2023**  
**Unité de recherche : Laboratoire de Mathématiques Jean Leray (LMJL)**

## **Rapporteurs avant soutenance :**

Rémi ABGRALL Full professor, Universität Zürich  
Marie-Hélène VIGNAL Maître de conférences, Université Paul Sabatier

## **Composition du Jury :**

Président :	Nicolas CROUSEILLES	Directeur de recherche, INRIA, Rennes
Examineurs :	Rémi ABGRALL	Full professor, Universität Zürich
	Marie-Hélène VIGNAL	Maître de conférences, Université Paul Sabatier
	Rodolphe TURPAULT	Professeur des universités, Bordeaux INP
	Marianne BESSEMOULIN-CHATARD	Chargée de recherche, CNRS, Nantes Université
Dir. de thèse :	Christophe BERTHON	Professeur des universités, Nantes Université
Co-dir. de thèse :	Mehdi BADSI	Maître de conférences, Nantes Université

# Remerciements

Je tiens tout d'abord à remercier Christophe Berthon et Mehdi Badsı pour leur encadrement ainsi que pour la confiance et la grande liberté qu'ils m'ont accordées. J'ai réellement apprécié ces trois années et j'espère que nous aurons d'autres occasions de travailler ensemble.

Je remercie également Rémi Abgrall et Marie-hélène Vignal d'avoir accepté de rapporter ma thèse ainsi que Marianne Bessemoulin-Chatard, Rodolphe Turpault et Nicolas Crouseilles pour leur participation au jury.

Ensuite, je remercie toutes les personnes que j'ai côtoyées au LMJL et plus généralement toutes celles et ceux que j'ai rencontrés durant mon parcours et qui ont rendus possible ce travail doctoral.

Enfin, je remercie ma famille et mes amis pour leur soutien constant et je leur dédie ce mémoire de thèse.

*À ma famille, à mes amis.*

# Table des matières

<b>Introduction</b>	<b>6</b>
<b>1 Approximations numériques de systèmes hyperboliques de lois de conservation</b>	<b>14</b>
1.1 Quelques propriétés des systèmes hyperboliques de lois de conservation . . . . .	14
1.2 Exemples de systèmes hyperboliques . . . . .	16
1.2.1 Le système de Saint Venant . . . . .	16
1.2.2 Le système de Ripa . . . . .	18
1.2.3 Les équations d'Euler . . . . .	19
1.3 Bibliographie de l'approximation numériques des systèmes hyperboliques de lois de conservation . . . . .	19
1.3.1 Quelques généralités sur les approximations numériques . . . . .	19
1.3.2 Les schémas numériques usuels du premier ordre . . . . .	21
1.3.3 La stabilité des schémas d'ordre élevé en espace . . . . .	22
1.3.4 L'obstacle de la discrétisation en temps . . . . .	25
1.3.5 Le terme source de topographie . . . . .	25
1.4 Approximations numériques avec la méthode des volumes finis . . . . .	27
1.4.1 Méthode des volumes finis . . . . .	27
1.4.2 Schéma de Godunov . . . . .	29
1.4.3 Schéma de <i>type</i> Godunov pour des systèmes d'équations homogènes . . . . .	31
1.4.4 Schéma de <i>type</i> Godunov pour des systèmes d'équations avec terme source . . . . .	36
<b>2 Solveurs de Riemann à deux états intermédiaires avec stabilité entropique locale pour des systèmes homogènes</b>	<b>39</b>
2.1 Condition suffisante de stabilité entropique locale . . . . .	39
2.2 Solveur de Riemann à deux états intermédiaires, consistant et entropique . . . . .	41
2.3 Application aux équations de Saint Venant homogènes . . . . .	43
2.4 Application aux équations d'Euler . . . . .	47
2.5 Résultats de simulations numériques . . . . .	53
2.5.1 Les équations de Saint Venant homogènes . . . . .	53
2.5.2 Les équations d'Euler . . . . .	56
2.6 Synthèse et conclusion . . . . .	57
<b>3 Solveurs de Riemann <i>well-balanced</i> et à stabilité entropique locale pour les systèmes de Saint Venant et de Ripa</b>	<b>59</b>
3.1 Schémas entropiques et <i>well-balanced</i> pour les équations de Saint Venant . . . . .	59
3.1.1 Condition suffisante de stabilité entropique locale . . . . .	59
3.1.2 Solveur de Riemann à deux états intermédiaires symétriques, consistant et entropique . . . . .	64
3.1.3 Schéma entropique et <i>well-balanced</i> pour le lac au repos . . . . .	70
3.1.4 Schéma entropique et <i>well-balanced</i> pour tous les états d'équilibres réguliers . . . . .	75
3.2 Schéma entropique et <i>well-balanced</i> pour les équations de Ripa . . . . .	84
3.2.1 Condition suffisante de stabilité entropique locale . . . . .	84

3.2.2	Solveur de Riemann à deux états intermédiaires symétriques, consistant et entropique . . . . .	90
3.2.3	Schéma entropique et <i>well-balanced</i> pour le lac au repos . . . . .	94
3.3	Résultats de simulations numériques . . . . .	104
3.3.1	Les équations de Saint Venant . . . . .	104
3.3.2	Les équations de Ripa . . . . .	111
3.4	Synthèse et conclusion . . . . .	112
<b>4</b>	<b>Schémas d'ordre élevé, non limités et à stabilité entropique discrète globale</b>	<b>114</b>
4.1	Schémas numériques dissipatifs . . . . .	114
4.2	Schémas non limités d'ordre deux . . . . .	117
4.3	Stabilité entropique discrète globale . . . . .	119
4.3.1	Notations et sommabilité . . . . .	119
4.3.2	Reformulation de l'inégalité de dissipation . . . . .	121
4.3.3	Existence de l'inégalité de dissipation . . . . .	124
4.3.4	Stabilité entropique discrète globale . . . . .	127
4.3.5	Étude de la condition CFL pour l'advection linéaire . . . . .	129
4.4	Montée en ordre supérieur . . . . .	132
4.4.1	Corrections dissipatives d'ordre $k$ . . . . .	132
4.4.2	Schémas non limités d'ordre trois et quatre . . . . .	135
4.4.3	Stabilité entropique discrète globale . . . . .	138
4.5	Résultats de simulations numériques . . . . .	141
4.5.1	Méthodes et paramétrages . . . . .	141
4.5.2	Schéma numérique du second ordre . . . . .	141
4.5.3	Schémas numériques du troisième et du quatrième ordre . . . . .	154
4.6	Synthèse et conclusion . . . . .	163
<b>5</b>	<b>Schémas d'ordre élevé, non limités et à stabilité entropique discrète globale sur maillages non structurés en dimension deux</b>	<b>165</b>
5.1	Approximations numériques des systèmes hyperboliques de lois de conservation en dimension deux . . . . .	165
5.1.1	Les inégalités d'entropie des systèmes hyperboliques de lois de conservation homogènes en dimension deux . . . . .	165
5.1.2	Quelques généralités sur les approximations numériques en dimension deux . . . . .	166
5.1.3	L'approximation des gradients en dimension deux . . . . .	167
5.1.4	La stabilité des schémas d'ordre élevé en dimension deux . . . . .	168
5.1.5	La méthode des volumes finis en deux dimensions d'espace . . . . .	170
5.2	Gradients et matrices hessiennes discrets non limités . . . . .	172
5.2.1	Définition de maillage admissible . . . . .	172
5.2.2	Gradients discrets du premier et du second ordre . . . . .	175
5.2.3	Matrices hessiennes discrètes du premier ordre . . . . .	178
5.2.4	Gradients et matrices hessiennes discrets des fonctions de flux . . . . .	182
5.3	Schémas non limités d'ordre deux et trois . . . . .	184
5.3.1	Définitions et notations . . . . .	184
5.3.2	Discrétisation des intégrales des fonctions de flux . . . . .	185
5.3.3	Définitions des corrections dissipatives . . . . .	188
5.4	Stabilité entropique discrète . . . . .	192
5.4.1	Inégalité de dissipation associée aux corrections du second ordre . . . . .	193
5.4.2	Inégalité de dissipation associée aux corrections du troisième ordre . . . . .	196
5.4.3	Stabilité entropique discrète globale . . . . .	202
5.5	Résultats de simulations numériques . . . . .	205
5.5.1	Méthodes et paramétrages . . . . .	205
5.5.2	Vortex isentropique . . . . .	207

5.5.3	Cas des problèmes de Riemann en dimension deux . . . . .	212
5.5.4	Écoulement dans un tunnel avec marche . . . . .	221
5.5.5	Double réflexion sur une rampe . . . . .	223
5.6	Synthèse et conclusion . . . . .	225
<b>Conclusion</b>		<b>227</b>
<b>A Solutions stationnaires régulières des équations de Saint Venant</b>		<b>229</b>
<b>Bibliographie</b>		<b>230</b>

# Introduction

## Contexte

Ce mémoire de thèse traite de l'approximation numérique des solutions de systèmes d'équations hyperboliques de lois de conservation. Pour  $(x, t)$  dans  $\mathbb{R} \times \mathbb{R}^+$ ,  $\Omega$  un ensemble convexe inclus dans  $\mathbb{R}^m$ , et une fonction  $f : \Omega \rightarrow \mathbb{R}^m$ , ces systèmes de  $m \geq 1$  équations sont de la forme

$$\partial_t w + \partial_x f(w) = 0. \quad (1)$$

Les variables  $x$  et  $t$  désignent respectivement l'espace et le temps et l'inconnue du système (1) est  $w : \mathbb{R} \times \mathbb{R}^+ \rightarrow \Omega$ . Ces systèmes d'équations aux dérivées partielles (EDP) sont utilisés dans la modélisation de phénomènes de transport tels que les écoulements en milieu peu profond [32], l'aérodynamique rapide [105] ou faiblement compressible [72]. Les exemples usuels sont l'équation de Lighthill, Whitham, Richards [153, 183] qui concerne la modélisation du trafic routier, l'équation de Buckley et Leverett qui décrit un écoulement diphasique non miscible en milieux poreux [39] ou encore les équations d'Euler [63, 88, 105] qui modélisent l'évolution d'un fluide compressible.

D'une manière générale, les solutions  $w$  des systèmes de la forme (1) sont continues en temps mais elles peuvent développer des discontinuités en espace. Dans ce cas, leur unicité est perdue [138, 139, 143, 190] et la sélection d'une solution s'opère via un critère supplémentaire qui s'exprime sous la forme d'inégalités d'entropie.

En notant  $\cdot^T$  l'opération de transposition usuelle de  $\mathbb{R}^m$ , une fonction convexe  $\eta : \Omega \rightarrow \mathbb{R}$  est appelée *entropie* s'il existe  $G : \Omega \rightarrow \mathbb{R}$  une fonction de classe  $C^1(\Omega, \mathbb{R})$  telle que  $\nabla G(w)^T = \nabla \eta(w)^T \nabla f(w)$ . Pour une entropie  $\eta$  donnée, il est admis qu'une solution d'un système hyperbolique de la forme (1) vérifie également une inégalité d'entropie qui s'écrit

$$\partial_t \eta(w) + \partial_x G(w) \leq 0. \quad (2)$$

Dans le cas d'équations scalaires (*i.e*  $m = 1$ ), et d'après [104], une solution de (1) est unique dès qu'elle vérifie l'inégalité (2) pour tous couples  $(\eta, G)$ .

Ainsi, l'objectif de ce mémoire de thèse est de proposer des *schémas numériques* qui approchent les solutions des systèmes de la forme (1). En notant  $w_i^n$  une approximation de  $w(x_i, t^n)$  dans  $\mathbb{R}^m$ , ces schémas numériques s'écrivent

$$\frac{w_i^{n+1} - w_i^n}{\Delta t} + \frac{\mathcal{F}_{i+\frac{1}{2}} - \mathcal{F}_{i-\frac{1}{2}}}{\Delta x} = 0. \quad (3)$$

Un schéma numérique (3) est souvent défini en vertu des propriétés mathématiques du système d'équations (1) mais d'après [119, 142], il doit nécessairement être consistant. Cette notion de consistance signifie qu'un schéma numérique (3) préserve les constantes de  $\Omega$  qui sont des solutions triviales du système (1). De plus, et en vertu des théorèmes de Lax [142] et de Lax-Wendroff [140, 149], la consistance doit être complétée par la notion de stabilité qui garantit que les erreurs d'approximations engendrées par le schéma ne vont pas s'amplifier au cours de la simulation numérique.

Il existe plusieurs définition de stabilité [53, 74, 149], mais d'après le Théorème de Lax-Wendroff [140, 149], il est admis que la stabilité privilégiée est *entropique*. Cette stabilité retranscrit au niveau discret une inégalité d'entropie (2) et elle s'écrit formellement

$$\frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} + \frac{\mathcal{G}_{i+\frac{1}{2}} - \mathcal{G}_{i-\frac{1}{2}}}{\Delta x} \leq 0. \quad (4)$$



Cette inégalité d'entropie discrète exprime une contrainte sur la quantité  $w_i^{n+1}$  et sa justification est double. Elle empêche d'abord la convergence du schéma vers des solutions non entropiques du système (1) [61, 132, 172] et elle évite ensuite l'apparition d'artefacts numériques tels que des oscillations parasites [92]. La Figure 1 illustre le défaut de convergence qui peut s'opérer lorsque la quantité  $w_i^{n+1}$  donnée par un schéma numérique (3) ne satisfait pas une inégalité d'entropie discrète (4).

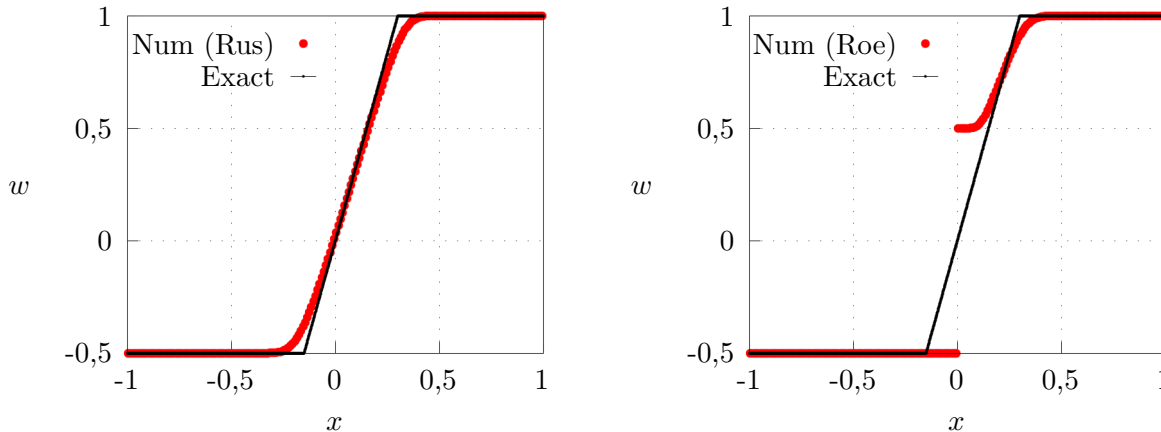


FIGURE 1 – À gauche (*resp.* à droite) résultat numérique donné par le schéma entropique de Rusanov (Rus) [185] (*resp.* non entropique de Roe [149, Section 4.14]) pour l'équation de Burgers définie par  $f(w) = w^2/2$ .

La simulation présentée dans la Figure 1 utilise un cas test proposé dans [172, Section 5] pour lequel la solution entropique de l'équation de Burgers se compose d'une onde de détente. Cette solution est convenablement restituée par le schéma de Rusanov [185] qui satisfait la stabilité entropique (4). En revanche, le schéma non entropique de Roe [149, Section 4.14] donne une solution constituée d'une onde de choc et d'une onde de détente. Il s'agit bien d'une solution faible de l'équation de Burgers car elle satisfait les relations de Rankine-Hugoniot mais cette solution n'est pas entropique.

Par ailleurs, ce mémoire de thèse s'intéresse également à l'approximation numérique des systèmes d'équations hyperboliques équipés d'un terme source. Pour une fonction  $S : \Omega \rightarrow \mathbb{R}^m$  et  $z : \mathbb{R} \rightarrow \mathbb{R}$  une fonction régulière donnée, les versions avec terme source du système (1) sont considérées sous la forme suivante :

$$\partial_t w + \partial_x f(w) = S(w) \partial_x z. \quad (5)$$

Ces versions modifiées du système (1) sont par exemple utilisées pour décrire un écoulement en milieu peu profond [32, 93] et qui s'opère sur un fond dont la topographie est modélisée par la fonction  $z$  donnée. L'approximation numérique de leurs solutions repose également sur des schémas consistants et qui vérifient une stabilité entropique discrète. Cependant, dans le cas de système d'équations avec terme source de la forme (5), ces deux aspects doivent être complétés par la propriété *well-balanced*.

La propriété *well-balanced* est un critère de précision supplémentaire qui assure la préservation exacte des solutions stationnaires non triviales (*i.e.* non constantes). Cette propriété prévient l'apparition d'erreurs numériques qui peuvent largement fausser les résultats de simulation [19, 111, 112]. Ce phénomène est illustré par la Figure 2 qui montre qu'un schéma numérique usuel commet une erreur de positionnement et d'amplitude pour la simulation de solution stationnaire non triviale.

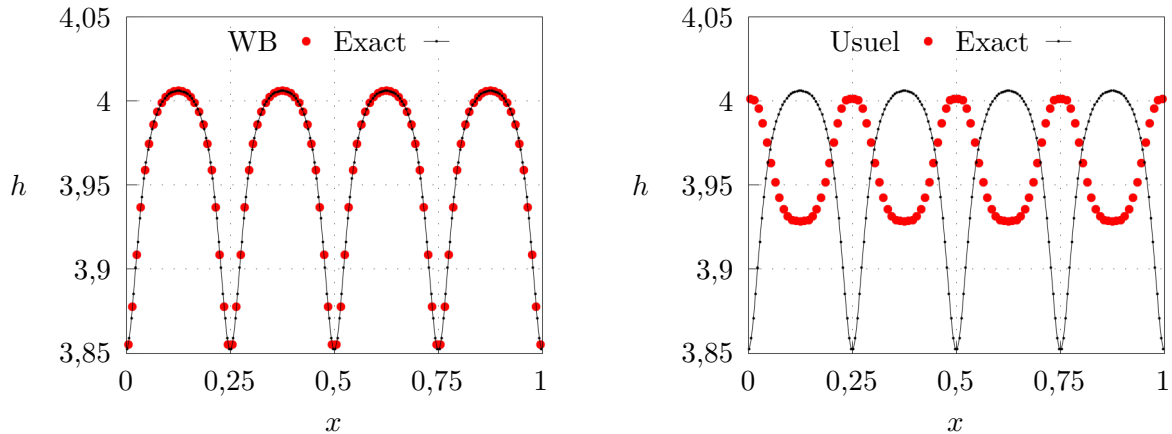


FIGURE 2 – À gauche (*resp.* à droite) résultats numériques d’un schéma *well-balanced* (WB) (*resp.* usuel) pour une solution stationnaire non constante des équations de Saint Venant avec terme source de topographie [32, Section 3.2].

Le développement de schémas numériques stables, précis et qui approchent les solutions des systèmes hyperboliques (1) ou (5) est un champ de recherche très actif illustré par la liste non exhaustive de publications [7, 22, 32, 34, 77, 96, 106, 118, 125, 127, 137, 230]. Ce mémoire de thèse s’inscrit dans ces recherches. La stabilité des schémas proposés est considérée au sens des inégalités d’entropie discrètes locales ou globales et la précision concerne des schémas du premier ordre dans le chapitre 2, *well-balanced* dans le chapitre 3 et d’ordre élevé dans les chapitres 4 et 5. Ces schémas stables et précis sont principalement appliqués aux équations de Saint Venant, de Ripa et d’Euler qui sont rappelées dans le chapitre 1. Les cinq chapitres de ce mémoire ainsi énoncés sont maintenant présentés à l’aide d’un plan détaillé.

## Plan détaillé et contributions

### Chapitre 1 : Approximations numériques de systèmes hyperboliques de lois de conservation

La section 1.1 du premier chapitre introduit les systèmes d’équations hyperboliques ainsi que les inégalités d’entropie qui leurs sont associées. Dans cette section, la notion de problème de Riemann ainsi que le Lemme de Harten [118] qui donne une expression de l’intégrale de  $x \mapsto w(x, t)$  sont également rappelés.

Ces généralités sont ensuite illustrées dans la section 1.2 avec les exemples usuels des systèmes d’équations de Saint Venant, de Ripa et d’Euler. Ces exemples montrent également l’influence du terme source qui modifie d’une part les inégalités d’entropie et qui induit, d’autre part, des solutions stationnaires régulières non triviales.

Ce chapitre aborde par la suite les difficultés relatives à l’approximation numérique des solutions des systèmes hyperboliques non linéaires. En s’appuyant sur une bibliographie, le paragraphe 1.3.1 introduit formellement les notions de précision et de stabilité des schémas puis motive la recherche de schémas précis vérifiant une stabilité entropique.

Ces éléments généraux établis, la section 1.3.2 présente des schémas numériques usuels du premier ordre. La plupart de ces schémas satisfont une stabilité entropique mais, en général, ils sont peu précis. Ensuite, trois difficultés particulières liées à l’approximation numérique sont détaillées dans des paragraphes dédiés.

D’abord, le paragraphe 1.3.3 concerne la stabilité des schémas formellement d’ordre élevé. La bibliographie exposée dans ce paragraphe montre d’abord une antinomie entre propriété de stabilité et ordre élevé [32, 116, 137]. La levée de cette antinomie est l’objet de nombreuses recherches et le paragraphe 1.3.3 en propose une présentation organisée par types de méthodes numériques.

Le second aspect évoqué dans le paragraphe 1.3.4 est la difficulté induite par la discrétisation temporelle. Les recherches bibliographiques synthétisées dans ce paragraphe montrent une nouvelle fois un paradoxe entre précision et stabilité. Les méthodes explicites en temps sont précises mais sujettes aux instabilités [159, 230], alors que leurs homologues implicites sont stables mais peuvent dégrader la qualité des résultats [230].

La dernière difficulté abordée dans le paragraphe 1.3.5 est la discrétisation des systèmes de Saint Venant et de Ripa avec un terme source de la forme  $S(w)\partial_x z$  où  $z : \mathbb{R} \rightarrow \mathbb{R}$  est une fonction régulière donnée. Ce paragraphe établit d’abord que la solution au problème de Riemann n’est pas unique [145, 146, 208, 209] puis il détaille les conséquences de cette perte d’unicité dans le cadre des approximations numériques. Le paragraphe 1.3.5 propose également un panorama des schémas numériques qui associent stabilité entropique et propriété *well-balanced*.

Enfin, ce premier chapitre se clôture avec la section 1.4 qui présente la méthode des volumes finis utilisée dans l’intégralité de ce mémoire. Les premiers paragraphes donnent l’expression d’un schéma numérique volume fini consistant et conservatif puis ils décrivent une notion formelle d’ordre spatial. Ces trois notions sont également complétées par la définition de stabilité entropique discrète. Ces définitions sont ensuite illustrées par le schéma de Godunov [106] puis par les schémas de type de Godunov. Pour ces derniers schémas, le formalisme général est rappelé puis un cas particulier est détaillé : le schéma de Harten Lax et Van Leer (HLL) [118]. Cette attention portée au schéma HLL est nécessaire car les recours à ses propriétés, et en particulier à sa stabilité entropique, sont nombreux dans ce mémoire de thèse. Des extensions des définitions usuelles de consistances sont ensuite proposées dans le paragraphe 1.4.4 pour le cas des systèmes avec un terme source de la forme  $S(w)\partial_x z$ .

## Chapitre 2 : Solveurs de Riemann à deux états intermédiaires avec stabilité entropique locale pour des systèmes homogènes

Le second chapitre propose des schémas numériques de type Godunov entropiques et définis par des solveurs de Riemann constants par morceaux à deux états intermédiaires. Pour un système hyperbolique homogène (1) donné, les sections 2.1 et 2.2 établissent des conditions sur le solveur de Riemann qui permettent d’obtenir, à l’aide de l’inégalité de Jensen, la stabilité entropique du schéma. Ces conditions de stabilité ainsi que la relation de consistance intégrale sont par exemple formulées dans [118] mais l’originalité des résultats proposés tient à leurs introductions dans la phase de construction du solveur.

À partir de ces conditions suffisantes de stabilité entropique et de consistance, des équations quadratiques sont proposées en Section 2.3 et 2.4 pour définir des solveurs de Riemann dédiés respectivement aux équations de Saint Venant homogènes et d’Euler. Pour les deux systèmes considérés, les propriétés de stabilité du schéma HLL [118] permettent d’établir que ces équations quadratiques admettent toujours des solutions. Chaque solution donne naissance à un schéma numérique qui, par construction, est consistant et entropique.

Ensuite, en Section 2.5, les schémas proposés sont illustrés et comparés aux schémas standards à l’aide de simulations numériques. Pour les cas tests usuels comme le tube à chocs de Sod [198], les résultats produits par les schémas développés sont satisfaisants. Néanmoins, ces schémas nécessitent parfois l’usage de limitations additionnelles qui assurent la robustesse dans les régions proches du bord de  $\Omega$ .

Enfin, ce second chapitre s’achève par la section 2.6 qui synthétise les résultats obtenus.

## Chapitre 3 : Solveurs de Riemann *well-balanced* et à stabilité entropique locale pour les systèmes de Saint Venant et de Ripa

Le troisième chapitre concerne l’approximation numérique des solutions des systèmes de Saint Venant et de Ripa tous deux avec un terme source de la forme  $S(w)\partial_x z$ . Ce chapitre se divise en quatre sections.

La première section 3.1 est consacrée aux équations de Saint Venant et elle débute par des rappels sur les propriétés du système d’équations ainsi que sur les schémas de type Godunov. Un premier résultat donné dans le paragraphe 3.1.1 établit des conditions suffisantes de consistance, de stabilité

entropique discrète locale et de propriété *well-balanced*. Les deux premières conditions sont ensuite utilisées dans le paragraphe 3.1.2 pour déduire un système d'équations sous-déterminé qui définit un solveur de Riemann à deux états intermédiaires constants.

Une première fermeture de ce système est proposée dans le paragraphe 3.1.3. Cette fermeture conduit à une équation quadratique dont le discriminant est paramétré par  $s_{LR}$  dans  $\mathbb{R}$  qui désigne formellement l'analogie discret de  $S(w)\partial_x z$ . Ce paramètre  $s_{LR}$  est choisi pour sa consistance mais également pour assurer l'existence de conditions telles que le discriminant de l'équation quadratique soit positif. Ainsi, après résolution explicite de cette équation quadratique, le paragraphe 3.1.3 présente un schéma de type Godunov consistant, entropique et *well-balanced* pour le lac au repos.

Une seconde fermeture est ensuite présentée dans le paragraphe 3.1.4. Cette seconde fermeture généralise la première et elle conduit à deux équations quadratiques. Le discriminant de la première équation est une nouvelle fois contrôlé par un choix consistant de  $s_{LR}$ . Le signe du discriminant de la seconde équation est assuré, si nécessaire, par une technique de limitation additionnelle. Par conséquent, ces deux équations quadratiques peuvent être résolues explicitement. Chaque solution donne formellement un schéma numérique consistant, entropique et le choix d'une solution particulière s'effectue selon le critère *well-balanced*. Ainsi, le paragraphe 3.1.4 montre l'existence d'un schéma de type Godunov consistant, entropique et *well-balanced* pour toutes les solutions stationnaires régulières.

La seconde section 3.2 est dédiée aux système de Ripa et elle se compose de trois sous-sections.

D'abord, le paragraphe 3.2.1 procède à des rappels sur les propriétés du système d'équations ainsi que sur l'approximation numérique des solutions. Ensuite, ce paragraphe donne des conditions suffisantes de consistance et de stabilité entropique pour un solveur de Riemann qui définit un schéma de type Godunov.

Ces conditions sont appliquées dans le paragraphe 3.2.2 à un solveur de Riemann constitué de deux états intermédiaires constants. Il en résulte trois équations et deux inéquations non linéaires dont la résolution directe reste difficile.

Par conséquent, le paragraphe 3.2.3 propose un système d'équations non linéaires qui permet de satisfaire les conditions énoncées précédemment. Sauf pour quelques cas particuliers, l'existence des solutions au système d'équations proposé n'a malheureusement pas pu être établie. Néanmoins, les très nombreuses résolutions conduites avec des méthodes usuelles de Newton ou de point fixe n'ont jamais montré de défaut de convergence. Ainsi, l'existence des solutions au système présenté à l'ouverture du paragraphe 3.2.3 est donnée sous la forme d'une conjecture. Ensuite, ce système est utilisé pour définir un schéma de type Godunov consistant, entropique et *well-balanced* pour le lac au repos.

La troisième section 3.3 est dédiée aux expériences numériques et sa première sous-section concerne les schémas numériques définis dans les paragraphes 3.1.3 et 3.1.4 pour les équations de Saint Venant.

Dans cette première sous-section 3.3.1, la propriété *well-balanced* pour le lac repos ainsi que les *transitions en zone sèche-mouillée* sont illustrées à l'aide du cas test usuel de l'île immergée [98]. Les résultats sont robustes et la propriété *well-balanced* est observée à la précision machine attendue.

La sous-section 3.3.1 propose également des tests et des comparaisons aux standards de la littérature [13, 35, 93] en considérant les trois problèmes de Goutal et Maurel [111]. Pour ces trois problèmes, les solutions numériques obtenues sont parfois localement mal positionnées et la propriété *well-balanced* n'est pas vérifiée dans le cas des équilibres mouvants atteints après régimes transitoires.

Enfin, les propriétés de stabilité des schémas proposés sont illustrées et comparées avec un cas test de rupture de barrage [166] également détaillé en Sous-section 3.3.1. Pour ce cas test, les résultats produits par les schémas des paragraphes 3.1.3 et 3.1.4 sont très satisfaisants. En particulier, les discontinuités des ondes de choc sont restituées avec très peu de cellules.

La section 3.3 se termine par le paragraphe 3.3.2 qui illustre le schéma dédié aux équations de Ripa. Le cas test considéré est une version modifiée de celui proposé par A. Chertock *et al* dans [58] et les résultats obtenus demeurent convenables.

Ce troisième chapitre se conclut par la section 3.4 qui résume les démarches et les résultats obtenus.

## Chapitre 4 : Schémas d'ordre élevé, non limités et à stabilité entropique discrète globale

Pour un système hyperbolique (1) donné, le chapitre 4 s'intéresse à la préservation des stabilités entropiques discrètes lorsque des montées en ordre des schémas numériques sont effectuées.

D'abord, la première section 4.1 propose une définition de schémas numériques dissipatifs. En complétant cette notion par une condition CFL, cette première section montre ensuite que si l'entropie  $\eta$  est suffisamment régulière alors un schéma numérique dissipatif satisfait une inégalité d'entropie discrète globale. Ainsi, les sections suivantes de ce chapitre proposent des schémas numériques d'ordre élevé et définis pour satisfaire cette propriété de dissipation.

La section 4.2 concerne des schémas numériques explicites en temps, formellement du second ordre et obtenus par une correction du schéma HLL standard [118]. Contrairement aux montées en ordre usuelles exposées dans la littérature [55, 57, 94, 116, 182, 194, 196, 219], les schémas proposés n'utilisent pas de techniques de limitations additionnelles. Ils s'expriment cependant à l'aide d'une séquence de matrices  $(\Theta_i^n)_{i \in \mathbb{Z}}$  dans  $\mathcal{M}_{m,m}(\mathbb{R})$  et dont le rôle est détaillé en Section 4.3.

Après introduction des sommabilités et des reformulations nécessaires, cette section 4.3 (paragraphe 4.3.3) donne des contraintes sur les matrices  $(\Theta_i^n)_{i \in \mathbb{Z}}$  en vertu de la propriété de dissipation attendue. Ce paragraphe 4.3.3 propose également des expressions de  $(\Theta_i^n)_{i \in \mathbb{Z}}$  qui satisfont les contraintes énoncées. Ainsi, et en utilisant la Section 4.1, la stabilité entropique discrète globale des schémas du second ordre proposés est établie dans le paragraphe 4.3.4.

La condition CFL utilisée dans ce dernier résultat n'est très certainement pas optimale et elle demeure très difficile à utiliser en pratique. Néanmoins, et d'après la sous-section 4.3.5, une relaxation de cette condition CFL est possible au moins dans le cas linéaire et à condition d'utiliser une discrétisation temporelle d'ordre élevé de type *Runge-Kutta Strong Stability Preserving* [107, 109, 110].

Ensuite, la section 4.4.1 généralise la structure d'ordre deux énoncée précédemment à un ordre  $k$  quelconque. En s'appuyant sur des discrétisations consistantes, une application de cette généralisation aux ordres trois et quatre est proposée en Section 4.4.2. Il en résulte des schémas numériques non limités et paramétrés par des séquences de matrices  $(\Theta_i^{O_k})_{(i,k) \in \mathbb{Z} \times \{3,4\}}$  qui permettent d'obtenir la propriété de dissipation.

Cette propriété est établie en Section 4.4.3 à l'aide de contraintes qui restreignent le choix des matrices  $(\Theta_i^{O_k})_{(i,k) \in \mathbb{Z} \times \{3,4\}}$ . Cette section montre également que ces contraintes peuvent être satisfaites et, par conséquent, les schémas non limités d'ordre trois et quatre proposés satisfont une stabilité entropique discrète globale.

Enfin, la section 4.5 évalue les performances des schémas d'ordre élevé à l'aide de simulations numériques et avec des paramétrages donnés en Section 4.5.1. Le paragraphe 4.5.2 concerne le second ordre pour lequel l'équation de Burgers et le système d'Euler sont considérés. Pour ces deux cas, plusieurs choix de matrices  $(\Theta_i^n)_{i \in \mathbb{Z}}$  sont proposés et des solutions régulières puis discontinues [141, 198] sont simulées. Dans le cas de solutions régulières, et puisque les schémas n'utilisent pas de limitation additionnelle, les ordres théoriques sont exactement retrouvés. Dans le cas de solutions discontinues, les résultats sont satisfaisants mais variables selon le choix des matrices  $(\Theta_i^n)_{i \in \mathbb{Z}}$  retenues. En particulier, une séquence de matrices génère très peu d'oscillations parasites et semble assurer le principe du maximum dans le cas scalaire.

Le paragraphe 4.5.3 est ensuite dédié aux schémas d'ordre trois et quatre pour lesquels l'évaluation est également conduite avec l'équation de Burgers et le système d'Euler. Des solutions régulières puis discontinues sont une nouvelle fois simulées et les ordres sont présentés sous la forme de tableaux.

Dans le cas de solutions régulières, les ordres théoriques attendus sont une nouvelle fois exactement mesurés. Dans le cas de solutions discontinues, les résultats sont conformes aux solutions de références mais ils présentent parfois des oscillations parasites d'amplitudes variables selon le choix des matrices  $(\Theta_i^{O_k})_{(i,k) \in \mathbb{Z} \times \{3,4\}}$ .

La section 4.6 conclut ce chapitre en récapitulant les résultats présentés.

## Chapitre 5 : Schémas d'ordre élevé, non limités et à stabilité entropique globale sur maillages non structurés en dimension deux

Le cinquième chapitre propose les extensions bi-dimensionnelles sur maillages non structurés des schémas numériques introduits dans le chapitre 4.

Tout d'abord, la section 5.1.1 présente les systèmes d'équations hyperboliques ainsi que les inégalités d'entropie dans un cadre bi-dimensionnel.

Ensuite, l'approximation numérique sur les maillages non structurés est détaillée à l'aide d'une bibliographie dédiée. Le paragraphe 5.1.2 donne les écritures formelles des schémas numériques et des inégalités d'entropie discrètes puis il souligne les difficultés inhérentes soulevées par les maillages non structurés. En particulier, l'approximation consistante des gradients est abordée dans le paragraphe 5.1.3 qui dresse un panorama des méthodes existantes. Par ailleurs, le paragraphe 5.1.4 détaille les pistes explorées pour articuler l'ordre élevé à une propriété de stabilité.

Enfin, la première section de ce chapitre s'achève par le paragraphe 5.1.5 qui présente la méthode des volumes finis bi-dimensionnelle. Les notions de conservation, de consistance, d'ordre spatial formel sont rappelées ainsi que la définition de stabilité entropique discrète. Un résultat technique d'intégration par parties discrète est également donné.

Ainsi, cette première section 5.1 donne le cadre nécessaire aux extensions des schémas numériques mono-dimensionnels proposés dans le chapitre 4. Puisqu'il s'agit de schémas d'ordre élevé, ces extensions débutent par des approximations consistantes des gradients et des matrices hessiennes de  $w$ .

À l'aide d'une notion de maillage admissible définie en Section 5.2.1, des approximations consistantes des gradients et des matrices hessiennes de  $w$  sont respectivement établies dans les paragraphes 5.2.2 et 5.2.3. Les approximations proposées sont linéaires, locales, consistantes et elles ne nécessitent pas de technique de limitation additionnelle. Ces quantités discrètes sont obtenues à l'aide du Théorème de Green-Gauss puis par des réécritures consistantes des développements de Taylor. De tels stratégies sont par exemple évoquées dans [28, 201] mais l'originalité des méthodes proposées tient à leurs expressions locales. Ces gradients et ces matrices hessiennes discrets sont ensuite utilisés en Section 5.2.4 pour définir des approximations des dérivées de la fonction composée  $f(w)$ .

Les dérivées discrètes de  $w$  et  $f(w)$  ainsi obtenues sont employées en Section 5.3 pour définir des schémas numériques explicites en temps, d'ordre deux et trois sur des maillages non structurés. Les *flux numériques* associés aux schémas reposent sur une discrétisation d'ordre élevé des intégrales des fonctions de flux puis sur des corrections dissipatives du même ordre respectivement données dans les paragraphes 5.3.2 et 5.3.3. Les corrections dissipatives dépendent d'une séquence de matrices  $(\Theta_{ij}^{Ok})_{(ij,k) \in \mathcal{E} \times \{2,3\}}$  définies sur les arêtes du maillage et qui sont destinées à assurer la stabilité des schémas proposés.

Des contraintes sont donc établies à cet effet en Section 5.4 laquelle propose également, dans les paragraphes 5.4.1 et 5.4.2, des expressions de matrices  $(\Theta_{ij}^{Ok})_{(ij,k) \in \mathcal{E} \times \{2,3\}}$  admissibles. Ces expressions permettent ensuite d'établir, sous condition CFL, l'existence de schémas numériques d'ordre élevé, sans limiteur de pente, définis sur des maillages non structurés, explicites en temps et qui vérifient une inégalité d'entropie discrète globale.

Ce résultat d'existence donné dans le paragraphe 5.4.3 repose sur un ensemble de paramètres pour lesquels des expressions explicites sont proposées. Il est cependant clair que ces expressions ne peuvent pas être utilisées en pratique. Par conséquent, la section 5.5 dédiée aux expériences numériques débute par un descriptif de mise en œuvre.

Après avoir détaillés la programmation des schémas, le paragraphe 5.5.1 présente d'abord les maillages utilisés pour les expériences numériques. Ces maillages se composent des grilles cartésiennes uniformes, des maillages définis par des triangles structurés puis par des triangles non structurés générés par le logiciel FreeFem++ [120]. Ensuite, le paragraphe 5.5.1 donne des choix heuristiques pour paramétrer les schémas d'ordre élevé ainsi qu'un critère de vérification de ces choix.

Les sections qui succèdent au paragraphe 5.5.1 illustrent la précision et la stabilité des schémas proposés avec des simulations numériques exclusivement dédiées aux équations d'Euler. Des résultats

de simulations sont d'abord donnés pour des solutions régulières [55, 199] puis pour des solutions discontinues [136, 224].

Puisque les schémas proposés n'utilisent pas de limiteurs de pente, les ordres théoriques sont exactement retrouvés. Une très légère dégradation est toutefois observée numériquement pour le schéma du troisième ordre sur des maillages non structurés.

La simulation de solutions discontinues donne des résultats acceptables. La stabilité entropique du schéma du second ordre est satisfaisante avec l'une des deux heuristiques proposées mais des instabilités sont parfois observées sur les maillages non structurés. Le schéma du troisième ordre facilite l'apparition d'instabilités de Kelvin-Helmholtz, mais ses performances de stabilité demeurent fragiles avec les heuristiques proposées.

L'ensemble des résultats établis sont enfin rappelés dans la section 5.6 qui clôture ce cinquième chapitre.

## Publications

Les travaux présentés dans ce mémoire de thèse ont donné naissance à trois publications. Elles comprennent tout d'abord, un *proceeding* relatif au schéma entropique et *well-balanced* développé dans le chapitre 3 (section 3.2) pour les équations de Ripa. Il s'intitule

- i) L. Martaud and C. Berthon. The existence of well-balanced entropy stable numerical scheme for the ripa model with the topography source term. In *Proceedings book. XXVII Congreso de Ecuaciones Diferenciales y XVII Congreso de Matemática Aplicada*, pages pp 89–96, July 2022

Ensuite, les schémas du second ordre présentés dans le chapitre 4 (sections 4.2 et 4.3) sont publiés dans l'article

- ii) M. Badsì, C. Berthon, and L. Martaud. A family of second-order dissipative finite volume schemes for hyperbolic systems of conservation laws. *The SMAI Journal of computational mathematics*, 9 :31–60, 01 2023

Enfin, les schémas du second ordre bi-dimensionnels définis dans le chapitre 5 sont l'objet de la publication

- iii) L. Martaud. Global entropy stability for a class of unlimited second-order schemes for 2D hyperbolic systems of conservation laws on unstructured meshes. *Journal of Computational Physics*, 487 :112176, 2023

Bien que ces publications aient été acceptées, une lecture attentive récente a montré la persistance de trois erreurs, chacune répartie dans un des trois articles. Ces manquements ne perturbent pas les résultats principaux mais ils requièrent néanmoins quelques corrections.

Des dispositions ont d'ores et déjà été prises en ce sens pour la publication *ii*). Un *erratum* à été envoyé à l'éditeur de la revue et une version corrigée est disponible sur HAL à l'adresse [hal.science/hal-03564325](https://hal.science/hal-03564325). Des dispositions seront également prises en ce qui concernent les publications *i*) et *iii*).

L'ensemble des corrections nécessaires figurent dans ce mémoire de thèse et par conséquent, quelques différences mineures subsistent entre les articles *i*), *ii*), *iii*) et les chapitres 3, 4, 5.

Par ailleurs, un article supplémentaire est en cours de rédaction. Cet article concerne les schémas entropiques et *well-balanced* développés dans le chapitre 3 (section 3.1) pour les équations de Saint Venant. Il s'intitule

- L. Martaud and C. Berthon. Fully well-balanced entropy stable Godunov numerical schemes for the Shallow Water equations with the topography source term. 2023.

# Chapitre 1

## Approximations numériques de systèmes hyperboliques de lois de conservation

### 1.1 Quelques propriétés des systèmes hyperboliques de lois de conservation

Pour un entier naturel non nul  $m$ , notons  $x$  la variable d'espace dans  $\mathbb{R}$  et  $t$  la variable de temps dans  $\mathbb{R}^+$ . Considérons un système hyperbolique de  $m$  équations de conservations en une dimension d'espace, d'inconnue  $w : \mathbb{R} \times \mathbb{R}^+ \rightarrow \mathbb{R}^m$  et s'écrivant

$$\begin{cases} \partial_t w + \partial_x f(w) = 0, & t > 0, \quad x \in \mathbb{R}, \\ w(x, t = 0) = w_0(x). \end{cases} \quad (1.1)$$

L'inconnue  $w$  prend ses valeurs dans un ensemble convexe non vide  $\Omega$  inclus dans  $\mathbb{R}^m$ ,  $f : \Omega \rightarrow \mathbb{R}^m$  désigne une fonction de flux donnée. Cette fonction de flux  $f$  est supposée être au moins de classe  $C^2(\Omega, \mathbb{R}^m)$ . On suppose de plus que le système (1.1) est strictement hyperbolique ce qui signifie que la matrice jacobienne de  $f$  notée  $\nabla f(w)$  est diagonalisable dans  $\mathbb{R}$  pour tous  $w$  de l'ensemble  $\Omega$  et que ses valeurs propres sont toutes de multiplicité un. À l'instant initial,  $w$  est égale à  $w_0$  une fonction mesurable donnée de  $L^1_{\text{Loc}}(\mathbb{R})$ .

Pour un système hyperbolique (1.1) donné, si la fonction de flux  $f$  est linéaire, alors la solution de ce système est connue et sa régularité est égale à celle de la donnée initiale  $w_0$ . En revanche dans le cas où  $f$  est non linéaire et même avec une donnée initiale  $w_0$  régulière, les solutions du système (1.1) peuvent admettre des discontinuités en temps fini. Dans ce cas, les solutions de (1.1) ne sont en général pas uniques [138, 139, 143, 190]. Il est admis dans ce mémoire de thèse que les solutions du système d'équations (1.1) sont caractérisées par les inégalités d'entropies.

En notant  $\cdot^T$  l'opération de transposition usuelle de  $\mathbb{R}^m$ , une fonction convexe  $\eta : \Omega \rightarrow \mathbb{R}$  est appelée *entropie* s'il existe  $G : \Omega \rightarrow \mathbb{R}$  une fonction de classe  $C^1(\Omega, \mathbb{R})$  appelé *flux d'entropie associé* à  $\eta$  et telle que

$$\nabla G(w)^T = \nabla \eta(w)^T \nabla f(w), \quad \forall w \in \Omega.$$

Pour un couple entropie, flux d'entropie  $(\eta, G)$  donné, on admet que les solutions du système (1.1) vérifient une *inégalité d'entropie* qui s'écrit

$$\partial_t \eta(w) + \partial_x G(w) \leq 0. \quad (1.2)$$

L'inégalité ci-dessus est locale car *a priori* définie en tout point  $x$  mais il est possible d'obtenir une caractérisation intégrale des solutions de (1.1). En effet, l'intégration de l'inégalité (1.2) selon la variable  $x$  donne l'*inégalité globale* suivante :

$$\int_{\mathbb{R}} \partial_t \eta(w(x, t)) \, dx \leq 0, \quad \forall t > 0. \quad (1.3)$$



Les inégalités d'entropie locales (1.2) ou globales (1.3) spécifient les solutions de (1.1) mais en général, il est impossible d'obtenir une expression analytique de  $w(x, t)$ . Toutefois, lorsque la donnée initiale  $w_0$  est constituée de deux constantes séparées par une discontinuité, le problème de Cauchy (1.1) devient un *problème de Riemann*. Dans ce cas et pour  $w_L$  et  $w_R$  deux constantes de  $\Omega$ ,  $w_0$  s'écrit

$$w_0(x) = \begin{cases} w_L & \text{si } x < 0, \\ w_R & \text{sinon.} \end{cases} \quad (1.4)$$

D'après le Lemme de Harten [118] rappelé ci-dessous, il est possible dans ce cas de donner une expression de l'intégrale de  $w$ .

**Lemme 1.1.1** (Lemme de Harten [118]). *Soit  $f : \Omega \rightarrow \mathbb{R}^m$  une fonction de flux et notons  $\{\mu_k(w)\}_{k \in \{1, \dots, m\}}$  l'ensemble des  $m$  valeurs propres de  $\nabla f(w)$ . Soient  $w_L$  et  $w_R$  deux constantes de  $\Omega$  et considérons le problème de Riemann formé de l'équation (1.1) avec la donnée initiale (1.4). Soient  $\Delta t > 0$ ,  $\Delta x > 0$  tels que*

$$\frac{\Delta t}{\Delta x} \max_{\substack{\alpha \in \{L, R\}, \\ k \in \{1, \dots, m\}}} |\mu_k(w_\alpha)| \leq \frac{1}{2}. \quad (1.5)$$

La solution du problème de Riemann (1.1)-(1.4) vérifie

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} w(x, \Delta t) dx = \frac{w_L + w_R}{2} - \frac{\Delta t}{\Delta x} (f(w_R) - f(w_L)). \quad (1.6)$$

*Démonstration.* En intégrant l'équation (1.1) sur le domaine  $[-\Delta x/2, \Delta x/2] \times [0, \Delta t]$ , il vient

$$\begin{aligned} 0 &= \int_0^{\Delta t} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} (\partial_t w + \partial_x f(w)) dx dt, \\ &= \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} w(x, \Delta t) dx - \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} w(x, t=0) dx + \int_0^{\Delta t} (f(w(\Delta x/2, t)) - f(w(-\Delta x/2, t))) dt. \end{aligned}$$

Sous la condition (1.5), les égalités  $w(\Delta x/2, t) = w_R$  et  $w(-\Delta x/2, t) = w_L$  sont satisfaites. Par conséquent, en utilisant l'expression de la condition initiale du problème de Riemann (1.4), on obtient

$$0 = \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} w(x, \Delta t) dx - \Delta x \frac{w_L + w_R}{2} + \Delta t (f(w_R) - f(w_L)).$$

En divisant par  $\Delta x > 0$  l'équation ci-dessus, on déduit l'égalité (1.6).  $\square$

Des extensions du système (1.1) seront également considérées dans ce mémoire. Ces versions se composent d'un terme additionnel appelé *terme source*. Pour  $z : \mathbb{R} \rightarrow \mathbb{R}$  une fonction régulière donnée, indépendante du temps et  $S : \Omega \rightarrow \mathbb{R}^m$ , les termes sources considérés se notent  $S(w)\partial_x z$ . Dans ce cas, l'extension du système (1.1) s'écrit

$$\begin{cases} \partial_t w + \partial_x f(w) = S(w)\partial_x z, & t > 0, \quad x \in \mathbb{R}, \\ w(x, t=0) = w_0(x). \end{cases} \quad (1.7)$$

Dans le cas d'un système de la forme (1.7), le Lemme de Harten n'est plus valide et seulement une approximation de l'intégrale de  $w$  peut être exhibée. Ce résultat est donné dans le lemme ci-dessous.

**Lemme 1.1.2.** *Soient  $f : \Omega \rightarrow \mathbb{R}^m$  une fonction de flux,  $z : \mathbb{R} \rightarrow \mathbb{R}$  une fonction régulière indépendante du temps et  $S : \Omega \rightarrow \mathbb{R}^m$ . Soient  $(w_L, z_L)$  et  $(w_R, z_R)$  deux constantes de  $\Omega \times \mathbb{R}$  et considérons le problème de Riemann*

$$\begin{cases} \partial_t w + \partial_x f(w) = S(w)\partial_x z, & t > 0, \quad x \in \mathbb{R}, \\ w(x, t=0) = \begin{cases} w_L & \text{si } x < 0, \\ w_R & \text{sinon.} \end{cases} \end{cases}$$

Soient  $\Delta t > 0$ ,  $\Delta x > 0$ . En écrivant  $\mathcal{O}$  la notation de Landau usuelle,  $w$  vérifie

$$\begin{aligned} \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} w(x, \Delta t) dx &= \\ \frac{w_L + w_R}{2} - \frac{\Delta t}{\Delta x} (f(w_R) - f(w_L)) + \frac{1}{\Delta x} \int_0^{\Delta t} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} S(w(x, t)) \partial_x z dx dt &+ \mathcal{O}(\Delta x) + \mathcal{O}(\Delta t). \end{aligned} \quad (1.8)$$

*Démonstration.* En intégrant l'équation sur le domaine  $[-\Delta x/2, \Delta x/2] \times [0, \Delta t]$ , il suit

$$\begin{aligned} \int_0^{\Delta t} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} S(w(x, t)) \partial_x z dx dt &= \int_0^{\Delta t} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} (\partial_t w + \partial_x f(w)) dx dt, \\ &= \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} w(x, \Delta t) dx - \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} w(x, t=0) dx + \int_0^{\Delta t} (f(w(\Delta x/2, t)) - f(w(-\Delta x/2, t))) dt, \\ &= \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} w(x, \Delta t) dx - \Delta x \frac{w_L + w_R}{2} + \int_0^{\Delta t} (f(w(\Delta x/2, t)) - f(w(-\Delta x/2, t))) dt. \end{aligned}$$

Si  $\Delta t > 0$  est suffisamment petit alors, pour tout  $t$  dans  $[0, \Delta t]$ , il existe des régions de l'espace localisées autour de  $x = \Delta x/2$  et de  $x = -\Delta x/2$  au sein desquelles la discontinuité située en  $x = 0$  n'a pas d'effet. Par conséquent, dans de telles régions, la donnée initiale est constante et  $w$  est régulière. Comme la fonction de flux  $f$  est également régulière, un développement de Taylor du premier ordre en  $\Delta t$  et en  $\Delta x$  donne dans ce cas

$$\begin{aligned} f(w(\Delta x/2, t)) &= f(w_R) + \mathcal{O}(\Delta x) + \mathcal{O}(\Delta t), \\ f(w(-\Delta x/2, t)) &= f(w_L) + \mathcal{O}(\Delta x) + \mathcal{O}(\Delta t). \end{aligned}$$

En insérant les développements ci-dessus dans l'égalité qui les précède puis en divisant par  $\Delta x > 0$ , on en déduit l'estimation (1.8).  $\square$

Le terme  $S(w) \partial_x z$  donne également naissance à des solutions stationnaires non triviales. En effet, si les solutions stationnaires régulières d'un système homogène (1.1) se résument à des solutions constantes, les solutions stationnaires régulières du système (1.7) vérifient

$$\partial_x f(w) = S(w) \partial_x z. \quad (1.9)$$

Les solutions stationnaires définies ci-dessus feront l'objet d'une attention particulière dans la suite de ce mémoire de thèse.

## 1.2 Exemples de systèmes hyperboliques

Cette section décrit des exemples usuels de systèmes hyperboliques homogènes (1.1) ou de systèmes avec terme source qui s'écrivent sous la forme (1.7). Dans ce second cas, le terme  $S(w) \partial_x z$  induit des solutions stationnaires non triviales définies (1.9). Ce terme source modifie également les inégalités d'entropie (1.2) et les exemples ci-après illustrent comment ces inégalités peuvent s'écrire dans le cas de systèmes avec terme source.

### 1.2.1 Le système de Saint Venant

Détaillées dans [32, 93], les équations de Saint Venant modélisent un écoulement d'eau en milieu peu profond caractérisé par une épaisseur verticale faible par rapport à l'étendue horizontale. En notant  $g > 0$  la constante de gravitation,  $h(x, t) \geq 0$  la hauteur d'eau et  $u(x, t)$  la vitesse dans  $\mathbb{R}$ , les équations de Saint-Venant sont données par

$$\partial_t \begin{pmatrix} h \\ hu \end{pmatrix} + \partial_x \begin{pmatrix} hu \\ hu^2 + \frac{gh^2}{2} \end{pmatrix} = \begin{pmatrix} 0 \\ -gh \end{pmatrix} \partial_x z. \quad (1.10)$$

Dans le système ci-dessus,  $z : \mathbb{R} \rightarrow \mathbb{R}$  décrit la topographie du fond sur lequel l'écoulement s'opère. Avec les notations introduites dans la section 1.1, l'inconnue  $w$ , la fonction de flux  $f$  et le terme source  $S$  s'écrivent dans ce cas

$$w = (h, hu)^T, \quad f(w) = \left( (hu), \frac{(hu)^2}{h} + \frac{gh^2}{2} \right)^T, \quad S(w) = (0, -gh)^T. \quad (1.11)$$

L'inconnue  $w$  prend ses valeurs dans l'espace convexe

$$\Omega = \left\{ (h, hu) \in \mathbb{R}^2 \mid h \geq 0, hu \in \mathbb{R} \right\}.$$

Dans l'espace  $\Omega$  ci-dessus,  $h$  peut être nul et dans ce cas la fonction  $f$  donnée par (1.11) n'est plus définie. D'après [15, 166], la convention admise dans ce cas est de considérer que si  $h$  est nul alors  $hu/h$  est également nul. Par conséquent, la variable  $u$  se définit par

$$u = \begin{cases} \frac{(hu)}{h}, & \text{si } h > 0, \\ 0, & \text{sinon.} \end{cases}$$

Lorsque la fonction  $z$  est constante, le terme source s'annule et le système (1.10) devient homogène. Dans ce cas le système de Saint Venant est strictement hyperbolique ce qui veut dire que  $\nabla f(w)$  est diagonalisable dans  $\mathbb{R}$  et que ses valeurs propres sont de plus distinctes [165, 166]. Le système de Saint Venant (1.10) est associé au couple entropie, flux d'entropie  $(\eta, G)$  suivant :

$$\eta(w) = \frac{(hu)^2}{2h} + \frac{gh^2}{2}, \quad G(w) = \frac{(hu)^3}{2h^2} + gh(hu), \quad \forall w \in \Omega.$$

L'inégalité d'entropie associée au système s'écrit

$$\partial_t \eta(w) + \partial_x G(w) \leq -ghu \partial_x z. \quad (1.12)$$

D'après l'inégalité ci-dessus, l'introduction du terme  $S(w) \partial_x z$  dans les équations de Saint Venant (1.10) modifie l'inégalité d'entropie (1.2) en ajoutant une quantité dépendante de  $z$  à droite de l'inégalité. Selon les auteurs de [32, 93] et puisque  $\partial_t z = 0$ , l'inégalité (1.12) peut se reformuler de manière équivalente sous une forme conservative qui s'écrit

$$\partial_t (\eta(w) + ghz) + \partial_x (G(w) + g(hu)z) \leq 0.$$

La modification de l'inégalité (1.2) n'est pas la seule conséquence du terme  $S(w) \partial_x z$  dans les équations de Saint Venant (1.10). En effet, ce terme implique également l'existence de solutions stationnaires qui vérifient

$$\partial_x \begin{pmatrix} hu \\ hu^2 + \frac{gh^2}{2} \end{pmatrix} = \begin{pmatrix} 0 \\ -gh \end{pmatrix} \partial_x z.$$

En admettant une régularité suffisante, la résolution directe du système d'équations ci-dessus donne les solutions stationnaires suivantes :

$$hu = \text{cste}, \quad \frac{u^2}{2} + g(h + z) = \text{cste}. \quad (1.13)$$

Ces solutions stationnaires sont aussi appelées *états d'équilibres* et parmi eux, le *lac au repos* fera l'objet d'une attention particulière dans la suite de ce mémoire de thèse. Cet état du lac au repos est caractérisé par une vitesse  $u$  nulle. Dans ce cas, les solutions (1.13) deviennent

$$u = 0, \quad h + z = \text{cste}. \quad (1.14)$$

## 1.2.2 Le système de Ripa

Le système de Ripa est une version enrichie du système de Saint Venant qui décrit également la température de l'eau. La hauteur d'eau et la vitesse sont respectivement notées  $h(x, t) > 0$  et  $u(x, t)$  dans  $\mathbb{R}$ . À partir de  $\varphi : \mathbb{R}_*^+ \rightarrow \mathbb{R}_*^+$  une fonction positive, régulière, bijective et de  $\theta : \mathbb{R} \times \mathbb{R}^+ \rightarrow \mathbb{R}_*^+$ , le champs de température est  $\varphi(\theta(x, t)) > 0$ . Avec ces notations, le système de Ripa s'écrit

$$\partial_t \begin{pmatrix} h \\ hu \\ h\theta \end{pmatrix} + \partial_x \begin{pmatrix} hu \\ hu^2 + \frac{gh^2\varphi(\theta)}{2} \\ h\theta u \end{pmatrix} = \begin{pmatrix} 0 \\ -gh\varphi(\theta) \\ 0 \end{pmatrix} \partial_x z. \quad (1.15)$$

Les quantités  $w$ ,  $f(w)$ ,  $S(w)$  pour le système de Ripa s'écrivent

$$\begin{aligned} w &= (h, hu, h\theta)^T, \\ f(w) &= \left( (hu), \frac{(hu)^2}{h} + \frac{gh^2\varphi((h\theta)/h)}{2}, \frac{(hu)(h\theta)}{h} \right)^T, \\ S(w) &= \left( 0, -gh\varphi((h\theta)/h), 0 \right)^T. \end{aligned}$$

L'espace convexe dans lequel  $w$  prend ses valeurs est

$$\Omega = \left\{ (h, hu, h\theta) \in \mathbb{R}^3 \mid h > 0, hu \in \mathbb{R}, h\theta > 0 \right\}.$$

D'après [1, 78], le système de Ripa (1.15) est strictement hyperbolique lorsque  $z$  est constante. Dans [78], Desveaux *et al* proposent de contraindre la fonction  $\varphi$  avec les conditions suivantes :

$$\varphi''(\theta)\varphi(\theta) - \frac{\varphi'(\theta)^2}{2} > 0, \quad \varphi(\theta) - \theta\varphi'(\theta) + \frac{\theta^2}{2}\varphi''(\theta) > 0, \quad \forall \theta \in \mathbb{R}_*^+.$$

Si  $\varphi$  vérifie les conditions ci-dessus alors les auteurs de [78] montrent que le système d'équations de Ripa (1.15) peut être équipé d'une paire entropie convexe, flux d'entropie  $(\eta, G)$  donnée par

$$\eta(w) = \frac{(hu)^2}{2h} + \frac{g\varphi((h\theta)/h)h^2}{2}, \quad G(w) = \frac{(hu)^3}{2h^2} + gh\varphi((h\theta)/h)(hu), \quad \forall w \in \Omega.$$

Les auteurs de [78] montrent également que les solutions de (1.15) vérifient l'inégalité d'entropie suivante :

$$\partial_t \left( \eta(w) + gh\varphi((h\theta)/h)z \right) + \partial_x \left( G(w) + g\varphi((h\theta)/h)(hu)z \right) \leq 0.$$

Ainsi, l'introduction de  $S(w)\partial_x z$  dans les équations de Ripa (1.15) transforme l'inégalité (1.2) en une autre inégalité conservative mais dépendante de la donnée  $z$ . Le terme  $S(w)\partial_x z$  entraîne aussi l'existence de solutions stationnaires qui vérifient

$$\partial_x \begin{pmatrix} hu \\ hu^2 + \frac{gh^2\varphi(\theta)}{2} \\ h\theta u \end{pmatrix} = \begin{pmatrix} 0 \\ -gh\varphi(\theta) \\ 0 \end{pmatrix} \partial_x z.$$

Les solutions stationnaires régulières du système d'équation ci-avant se répartissent en deux familles. La première regroupe les solutions stationnaires à vitesse non nulle vérifiant

$$hu = \text{cste}, \quad \theta = \text{cste}, \quad \frac{u^2}{2} + g\varphi(\theta)(h + z) = \text{cste}. \quad (1.16)$$

La seconde famille est une version du lac au repos pour les équations de Ripa (1.15). Elle contient les équilibres à vitesse nulle obtenus en passant à la limite  $u \rightarrow 0$  dans les expressions ci-dessus. Cette seconde famille d'équilibre s'écrit

$$u = 0, \quad \theta = \text{cste}, \quad h + z = \text{cste}. \quad (1.17)$$

### 1.2.3 Les équations d'Euler

Les équations d'Euler décrivent l'évolution d'un fluide compressible et sont par exemple détaillées dans [63, 88, 105]. Les grandeurs de ce fluide sont la masse volumique  $\rho(x, t) > 0$ , la vitesse  $u(x, t)$  dans  $\mathbb{R}$ , l'énergie totale  $\rho E(x, t) > 0$  et  $p(x, t) > 0$  la pression. Les équations d'Euler sont données par

$$\partial_t \begin{pmatrix} \rho \\ \rho u \\ \rho E \end{pmatrix} + \partial_x \begin{pmatrix} \rho u \\ \rho u^2 + p \\ (\rho E + p)u \end{pmatrix} = 0. \quad (1.18)$$

Ces équations doivent être accompagnées d'une relation de fermeture sur  $p$  appelée équation d'état. Dans la suite, on adoptera le modèle du gaz parfait. En notant  $\gamma$  la constante adiabatique dans ]1, 3],  $p$  vérifie dans ce cas l'équation

$$p(w) = (\gamma - 1) \left( \rho E - \frac{(\rho u)^2}{2\rho} \right).$$

Le système (1.18) est homogène et les quantités  $w$  et  $f$  sont définies par

$$w = (\rho, \rho u, \rho E)^T, \quad f(w) = \left( (\rho u), \frac{(\rho u)^2}{\rho} + p(w), ((\rho E) + p(w)) \frac{(\rho u)}{\rho} \right)^T.$$

L'espace des valeurs admissibles de  $w$  est

$$\Omega = \left\{ (\rho, \rho u, \rho E) \in \mathbb{R}^3 \mid \rho > 0, \rho u \in \mathbb{R}, E - u^2/2 > 0 \right\}.$$

Le système d'Euler (1.18) est strictement hyperbolique [105]. Les entropies associées au système d'Euler (1.18) sont par exemple détaillées dans [22, 88, 138]. Elles reposent sur une fonction  $\mathcal{K} : \mathbb{R} \rightarrow \mathbb{R}$  telle que

$$\mathcal{K}'(y) > 0, \quad \mathcal{K}'(y) - \gamma \mathcal{K}''(y) > 0, \quad \forall y \in \mathbb{R}.$$

Pour une fonction  $\mathcal{K}$  satisfaisant les inégalités strictes ci-dessus, les couples d'entropies  $(\eta, G)$  associés au système d'Euler (1.18) s'écrivent

$$\eta(w) = -\rho \mathcal{K}(\ln(p/\rho^\gamma)), \quad G(w) = -(\rho u) \mathcal{K}(\ln(p/\rho^\gamma)), \quad \forall w \in \Omega.$$

Dans la suite, le choix  $\mathcal{K}(y) = y$  est proposé et par conséquent le couple  $(\eta, G)$  s'écrit

$$\eta(w) = -\rho \ln(p/\rho^\gamma), \quad G(w) = -(\rho u) \ln(p/\rho^\gamma), \quad \forall w \in \Omega.$$

## 1.3 Bibliographie de l'approximation numériques des systèmes hyperboliques de lois de conservation

Ce mémoire de thèse est consacré à la construction de méthodes numériques approchant les solutions du système (1.1) ou du système (1.7) et vérifiant des équivalents discrets de (1.2) ou bien de (1.3). La construction de telles méthodes constitue un champs de recherche très actif depuis de nombreuses années. Les critères généraux, les méthodes explorées ainsi que les résultats obtenus sont maintenant détaillés à l'aide d'une bibliographie.

### 1.3.1 Quelques généralités sur les approximations numériques

Les solutions des systèmes homogènes (1.1) ou leurs versions avec terme source (1.7) sont approchées sur des maillages du temps et de l'espace respectivement donnés par une séquence  $(t^n)_{n \in \mathbb{N}}$  et un ensemble d'interfaces  $(x_{i+\frac{1}{2}})_{i \in \mathbb{Z}}$ . Ces deux maillages sont supposés uniformes et les pas constants sont respectivement notés  $\Delta t > 0$  et  $\Delta x > 0$ . À la date  $t^n$  et sur chaque intervalle  $]x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}[$ , la quantité  $w(\cdot, t^n)$  est approchée par un objet noté formellement  $w_i^n$  et pour lequel de très nombreuses définitions sont possibles. Il s'agit par exemple d'une quantité ponctuelle dans [92, 116, 184], d'une constante de

$\mathbb{R}^m$  dans [32, 149, 211] ou bien d'un polynôme à valeurs dans  $\mathbb{R}^m$  dans [7, 125, 197]. Si  $t^0$  désigne le premier instant discret alors  $w_i^0$  s'établit à l'aide de la donnée initiale  $w_0$ . Par conséquent, il faut définir un *schéma numérique* qui permet d'obtenir  $(w_i^{n+1})_{i \in \mathbb{Z}}$  à partir de  $(w_i^n)_{i \in \mathbb{Z}}$  pour tout  $n$  dans  $\mathbb{N}$ . En notant  $\mathcal{F}_{i+\frac{1}{2}}$  un *flux numérique* approchant formellement la quantité  $f(w(x_{i+\frac{1}{2}}, t^n))$ , un schéma dédié aux systèmes d'équations homogènes (1.1) peut s'écrire

$$\frac{w_i^{n+1} - w_i^n}{\Delta t} + \frac{\mathcal{F}_{i+\frac{1}{2}} - \mathcal{F}_{i-\frac{1}{2}}}{\Delta x} = 0, \quad \forall i \in \mathbb{Z}.$$

Comme  $w_i^n$  et  $\mathcal{F}_{i+\frac{1}{2}}$  sont des approximations respectives de  $w(\cdot, t^n)$  et de  $f(w(x_{i+\frac{1}{2}}, t^n))$ , il convient de caractériser l'erreur commise à l'aide d'un critères de précision.

Le critère de précision le plus usuel est la notion d'ordre [18, 32, 55, 144, 219] mais l'étude bibliographique menée ne montre pas l'existence d'une définition générale et unificatrice de cette notion. Néanmoins, la notion d'ordre peut s'introduire formellement de la manière suivante. Étant donnée une quantification de l'erreur d'approximation, un schéma numérique est dit *d'ordre  $k$  en espace* si cette erreur s'écrit formellement  $\mathcal{O}_{\Delta x \rightarrow 0}(\Delta x^k)$ . La notion d'ordre en temps se caractérise de la même manière en substituant  $\Delta x$  par  $\Delta t$ . Un schéma d'ordre spatial  $k$  est en général couplé à une discrétisation en temps du même ordre [16, 125, 194, 230]. Les discrétisations temporelles usuelles sont détaillées dans [41, 102, 108]. La construction de schémas numériques d'ordre élevé en espace se justifie selon au moins deux axes.

D'abord et d'après [148], les schémas du premier ordre introduisent artificiellement des artefacts qui régularisent les solutions numériques calculées. Cette altération de *diffusion numérique* conduit souvent à sous estimer les extrema de  $w$ . Il est admis que les montées en ordre corrigent au moins partiellement ce défaut. La seconde motivation à la recherche de schémas d'ordre élevé est que, pour une précision de simulation numérique donnée, travailler sur des maillages grossiers avec un schéma d'ordre élevé est plus économique en ressource et en temps de calcul que de travailler sur des maillages fins avec un schéma du premier ordre [168, 193, 210].

Malheureusement, sans précaution particulière les schémas d'ordre élevé sont nécessairement oscillants voire instables dans les zones de discontinuités [149, 193]. Un résultat célèbre à ce propos est le théorème de Godunov [106] qui stipule qu'un schéma linéaire ne créant pas d'extrema est au plus du premier ordre. Ainsi, la précision doit être complétée de critères de stabilités.

L'un des premiers critères de stabilité d'un schéma numérique est la préservation du domaine  $\Omega$  qui consiste à garantir que  $(w_i^n)_{i \in \mathbb{Z}}$  demeure dans  $\Omega$  pour tout  $n$  dans  $\mathbb{N}$ . Ce critère permet par exemple d'assurer le sens physique des grandeurs calculées (pression positive, hauteur d'eau positive, *etc*). Dans le cas d'une fonction de flux  $f$  linéaire ou bien de systèmes linéarisés, la stabilité des schémas est parfois conduite dans le cadre des fonctions de carré sommables  $L^2$  munies d'une norme notée  $\|\cdot\|_2$ . Dans ce cas, la séquence mise à jour  $(w_i^{n+1})_{i \in \mathbb{Z}}$  doit vérifier

$$\sum_{i \in \mathbb{Z}} \|w_i^{n+1}\|_2^2 \Delta x \leq \sum_{i \in \mathbb{Z}} \|w_i^n\|_2^2 \Delta x.$$

Cette stabilité peut s'établir grâce à l'analyse de Fourier-Von Neumann [53, 74] ou bien directement algébriquement [75]. Dans le cas d'équations scalaires, la stabilité  $L^2$  est aussi parfois complétée d'un critère *Total Variation Diminishing* (TVD) [149]. Ce critère TVD donne un contrôle sur les oscillations numériques parasites mais est insuffisant sur le plan théorique [63].

La stabilité entropique généralise ces notions et consiste à montrer que la séquence mise à jour  $(w_i^{n+1})_{i \in \mathbb{Z}}$  vérifie un analogue discret de (1.2). En notant  $\mathcal{G}_{i+\frac{1}{2}}$  une approximation formelle de  $G(w(x_{i+\frac{1}{2}}, t^n))$ , cet analogue discret s'écrit sous la forme d'une inégalité d'entropie discrète locale donnée par

$$\frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} + \frac{\mathcal{G}_{i+\frac{1}{2}} - \mathcal{G}_{i-\frac{1}{2}}}{\Delta x} \leq 0, \quad \forall i \in \mathbb{Z}. \quad (1.19)$$

L'intérêt des stabilités entropiques est double. D'un point de vue théorique d'abord, si on dispose d'un schéma conservatif et consistant alors d'après [119], ce schéma peut converger vers une solution

faible de (1.1). Toutefois, la solution limite obtenue à convergence n'est pas nécessairement entropique [61, 132]. D'après le théorème de Lax Wendroff [140, 149], ce caractère entropique peut se retrouver à condition de retranscrire une stabilité entropique au niveau discret. D'un point de vue numérique ensuite, une inégalité d'entropie discrète est un critère de stabilité fort et on peut citer [92] où une expérience montre qu'une stabilité entropique élimine plus d'artefacts numériques qu'une stabilité  $L^2$  usuelle.

Les stabilités entropiques ainsi que la notion d'ordre sont valables sans distinction particulière pour les systèmes homogènes (1.1) et leurs versions avec terme source (1.7). En revanche l'introduction du terme  $S(w)\partial_x z$  et les solutions stationnaires qu'il implique nécessitent la définition d'une caractéristique supplémentaire appelée *propriété well-balanced*.

Lorsque le terme source  $S(w)\partial_x z$  est présent, la simulation numérique des solutions stationnaires de (1.9) doit être très précise [19, 111, 112]. Dans le cas contraire, les erreurs numériques peuvent s'amplifier au cours du temps et rendre les résultats inexploitable. La notion de précision d'un schéma numérique repose dans ce cas sur la propriété *well-balanced*. En notant  $\mathcal{S}$  un ensemble de solutions de (1.9) à valeurs dans  $\Omega$ , un schéma numérique est dit *well-balanced pour  $\mathcal{S}$* , s'il préserve exactement les états de  $\mathcal{S}$ . C'est à dire,

$$\text{si } w_i^n \in \mathcal{S}, \quad \forall i \in \mathbb{Z}, \quad \text{alors, } w_i^{n+1} = w_i^n, \quad \forall i \in \mathbb{Z}.$$

Un même schéma peut par exemple combiner la propriété *well-balanced* et être d'ordre élevé [48, 166, 170]. D'une manière générale, un schéma numérique idéal serait d'ordre élevé en espace et en temps, vérifierait l'inégalité d'entropie discrète (1.19) et serait de plus *well-balanced* dans le cas de  $S(w)\partial_x z$  non nul. L'élaboration d'un tel schéma constitue un champs un recherche très actif et les sections à venir dressent l'éventail des résultats.

### 1.3.2 Les schémas numériques usuels du premier ordre

De très nombreux schémas numériques localement entropiques et du premier ordre ont été proposés pour approcher les solutions des systèmes homogènes (1.1). S'il existe les  $E$ -schémas [36, 87, 149], les schémas cinétiques [31, 134] et d'autres schémas décrits dans [99, 101, 123], la plupart des schémas usuels peuvent se décrire avec la méthode des volumes finis. Dans cette méthode,  $w_i^n$  est une constante de  $\mathbb{R}^m$  qui vérifie formellement

$$w_i^n \approx \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} w(x, t^n) dx, \quad \forall i \in \mathbb{Z}.$$

Par construction et comme l'illustre la Figure 1.1, la séquence  $(w_i^n)_{i \in \mathbb{Z}}$  peut s'interpréter comme une juxtaposition de problèmes de Riemann locaux, chacun défini à une interface du maillage. C'est à partir de ce constat que S. Godunov propose dans [106] un schéma numérique construit sur la donnée de  $w(x, t^n + \Delta t)$  qui désigne la solution exacte de la juxtaposition des problèmes de Riemann. L'obtention de cette solution exacte est en générale difficile voire impossible en raison des non linéarités de  $f$ . Toutefois, après résolution des problèmes de Riemann, la séquence mise à jour  $(w_i^{n+1})_{i \in \mathbb{Z}}$  se déduit en intégrant  $w(x, t^n + \Delta t)$  sur l'intervalle  $]x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}[$ .

La principale difficulté du schéma de Godunov [106] réside dans le calcul de  $w(x, t^n + \Delta t)$ . Il est possible de contourner cet obstacle avec des schémas de *type* Godunov qui s'appuient sur une approximation de  $w(x, t^n + \Delta t)$ . Le schéma de type Godunov le plus simple est celui de Harten, Lax et van Leer (HLL). Introduit dans [118, 119], ce schéma vérifie une inégalité d'entropie discrète de la forme (1.19). On peut aussi citer le schéma de Harten Lax et van Leer *Contact preserving* (HLLC) [213]. Ce schéma dédié aux équations d'Euler préserve exactement les ondes de contact stationnaires et satisfait une stabilité entropique discrète locale. Une inégalité d'entropie discrète locale (1.19) est également vérifiée par le schéma détaillé dans [118, Section b.ii]. Ce schéma de type Godunov est construit pour une capture précise des ondes de chocs.

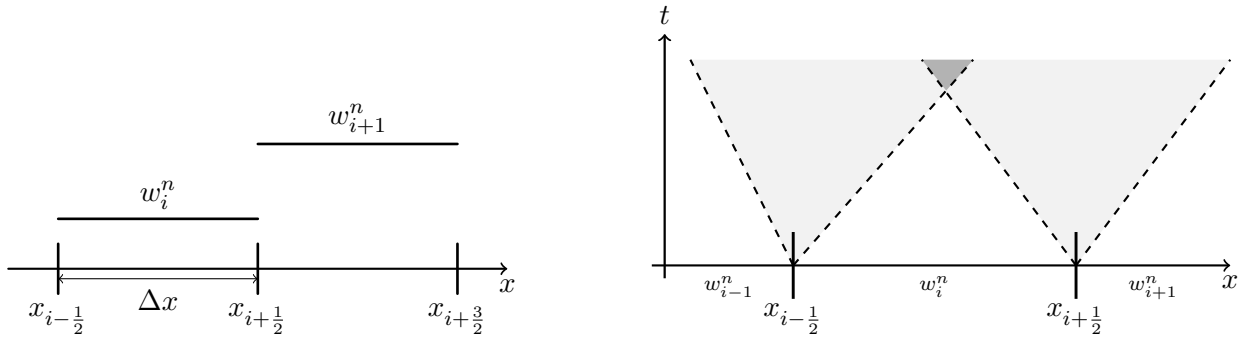


FIGURE 1.1 – À gauche, approximation constante par morceaux de  $w(\cdot, t^n)$  sur le maillage spatial défini par interfaces  $(x_{i+\frac{1}{2}})_{i \in \mathbb{Z}}$  de pas constant  $\Delta x > 0$ . À droite, juxtaposition dans le plan  $(x, t)$  de problèmes de Riemann autour des interfaces  $x_{i-\frac{1}{2}}$ ,  $x_{i+\frac{1}{2}}$  où les zones d'influences des solutions sont représentées en gris.

Enfin, il existe les schémas de relaxation [21, 32, 49, 64, 86] construits sur un découplage des équations (1.1) via l'introduction de variables supplémentaires. Ces schémas de relaxation revêtent des caractères implicites ou demandent des calculs très techniques. Par exemple, la preuve de l'inégalité d'entropie discrète (1.19) dans [21] nécessite d'établir une équation supplémentaire, l'existence de fonctions implicites et enfin l'existence d'un minimum local.

Les schémas numériques usuels rappelés dans cette section sont tous du premier ordre. La section suivante présente les pistes explorées pour conjuguer ordre élevé de résolution et propriétés de stabilité.

### 1.3.3 La stabilité des schémas d'ordre élevé en espace

La construction d'une méthode numérique qui articule la stabilité entropique et l'ordre élevé en espace est une tâche difficile qui s'illustre par exemple avec les citations ci-dessous.

*"both high resolution and strong entropy enforcement, a task that seems self-contradictory"*

A. Harten, 1982, [116].

*"It is extremely difficult to obtain second-order schemes that verify an entropy inequality."*

F. Bouchut, 2004, [32].

*"Methods that ensure the validity of local fully discrete entropy inequalities are still rare."*

D. Kuzmin, 2021, [137].

Néanmoins, la littérature foisonne de méthodes formellement d'ordre élevé en espace et satisfaisants certains critères de stabilité. Ces méthodes sont décrites dans les paragraphes ci-après qui débiteront par les deux paradigmes ADER et MOOD.

Proposés dans [212], les schémas de type *Arbitrary high order DERivative* (ADER) [46, 62, 84, 214, 215] se construisent par la résolution de problèmes de Riemann généralisés. À la date  $t^n$ , ces problèmes sont obtenus en considérant des polynômes d'ordre quelconque sur chaque cellule du maillage puis leurs résolutions consistent à donner une expression de  $w(x, t^n + \Delta t)$ . En incorporant la solution d'un problème de Riemann généralisé dans les flux numériques usuels [83], on obtient un schéma d'ordre croissant avec le degré des polynômes employés. Néanmoins, les solutions analytiques de ces problèmes étant très souvent inaccessibles, la résolution pratique est conduite avec un solveur implicite local [84, 210].

Le paradigme *Multi-dimensional Optimal Order Detection* (MOOD) [60, 79] contrôle *a posteriori* la stabilité des schémas d'ordre élevé. À la date  $t^n$  et étant donné des critères de stabilité  $\mathcal{C}$ , on procède à une itération en temps du schéma d'ordre élevé. Si la solution candidate ainsi obtenue vérifie  $\mathcal{C}$  alors elle est retenue comme approximation à la date  $t^{n+1}$ . Dans le cas contraire et tant que  $\mathcal{C}$  n'est pas vérifié, on réduit successivement et localement l'ordre du schéma pour générer une nouvelle solution candidate. Comme les schémas du premier ordre vérifient *a priori*  $\mathcal{C}$  la procédure itérative MOOD



est nécessairement finie. Cette approche est par exemple appliquée dans [77] pour forcer une inégalité d'entropie locale (1.19) et dans [207] pour préserver des domaine  $\Omega$ . Le paradigme MOOD a aussi été couplé aux méthodes ADER dans [158] afin d'accroître qualitativement la stabilité générale.

Dans son article fondateur [106], S. Godounov utilise la projection standard  $L^2$  pour obtenir un schéma du premier ordre mais satisfaisant une stabilité entropique locale discrète de la forme (1.19). L'idée suivie dans [33, 63, 65] est alors de construire des opérateurs préservant à la fois l'ordre élevé et cette stabilité entropique. Une telle méthode est explicitement appliquée dans [63] pour les équations d'Euler (1.18). Il en résulte un schéma du second ordre vérifiant une version discrète de toutes les inégalités d'entropie associées aux équations d'Euler. Malheureusement, les principes de ces méthodes sont implicites, très non linéaires et intimement liés au système étudié.

Une autre idée est de conduire les montées en ordre en corrigeant les schémas du premier ordre qui eux possèdent une stabilité entropique locale (1.19). D'abord sans analyse de stabilité dans [233], c'est dans cet esprit que les schémas *Monotonic Upstream-centered Scheme for Conservation Laws* (MUSCL) [219] et les schémas de A. Harten [116] ont été introduits. Ces schémas sont formellement du second ordre mais seules des preuves TVD dans le cas scalaire sont données dans [116]. Par ailleurs, les auteurs de [22, 77, 174] montrent que les schémas MUSCL satisfont des formes incomplètes de stabilité entropique. En notant  $\bar{\eta}_i$  une approximation de  $\eta(w_i^n)$ , ces premières versions incomplètes de (1.19) s'écrivent

$$\frac{\eta(w_i^{n+1}) - \bar{\eta}_i}{\Delta t} + \frac{\mathcal{G}_{i+\frac{1}{2}} - \mathcal{G}_{i-\frac{1}{2}}}{\Delta x} \leq 0, \quad \forall i \in \mathbb{Z}. \quad (1.20)$$

Le second ordre des techniques MUSCL est obtenu par l'introduction d'une reconstruction affine de la solution numérique. Ses extensions naturelles sont les méthodes *Essentially Non-Oscillatory* (ENO) et *Weighted Essentially Non-Oscillatory* (WENO) [117] qui produisent des schémas d'ordres croissants avec le degré du polynôme de reconstruction. L'originalité de ces schémas provient d'une reconstruction polynomiale choisie pour limiter les oscillations parasites [94, 179, 193, 232]. Avec ce type de méthodes des schémas d'ordre trois à cinq vérifiant un critère TVD [229] ou un critère de stabilité  $L^2$  [29] ont pu être proposés. Pour autant, les reconstructions polynomiales sont fortement non linéaires ce qui rend souvent les méthodes (W)ENO hermétiques aux preuves de stabilité formelles. Par exemple le résultat de stabilité  $L^2$  de [29] repose sur des expériences numériques. Ce verrou est maintenant levé et plusieurs schémas d'ordre élevé vérifiant certaines stabilités entropiques sont établis. Dans [230], Zakerzadeh *et al* montrent qu'une famille de schémas d'ordre  $k$  dérivés des méthodes (W)ENO vérifient une inégalité d'entropie discrète globale de la forme

$$\sum_{i \in \mathbb{Z}} \frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} \Delta x \leq 0. \quad (1.21)$$

De leurs côtés, les auteurs de [51, 57, 94] proposent des inégalités d'entropie semi-discrètes. Ces secondes versions incomplètes de (1.19) sont discrètes en espace mais continues en temps. Elles s'écrivent

$$\frac{d}{dt} \eta(w_i(t)) + \frac{\mathcal{G}_{i+\frac{1}{2}}(t) - \mathcal{G}_{i-\frac{1}{2}}(t)}{\Delta x} \leq 0, \quad t > 0, \quad \forall i \in \mathbb{Z}. \quad (1.22)$$

Enfin, un schéma numérique décrit dans [92] satisfait une version globale de l'inégalité ci-dessus donnée par

$$\frac{d}{dt} \sum_{i \in \mathbb{Z}} \eta(w_i(t)) \Delta x \leq 0, \quad t > 0. \quad (1.23)$$

Les inégalités d'entropie de la forme (1.21)-(1.22)-(1.23) sont obtenues par un couplage des reconstructions (W)ENO avec la stratégie des variables entropiques.

La stratégie des variables entropiques [47, 144, 205] repose tout d'abord sur la connaissance d'une fonction  $\psi : \Omega \rightarrow \mathbb{R}$  appelée *potentiel entropique* et vérifiant  $\psi = \nabla \eta^T f - G$ . À partir d'un potentiel entropique donné, la méthode des variables entropiques est ensuite fondée sur la détermination d'un flux numérique  $\mathcal{F}_{i+\frac{1}{2}}$  satisfaisant l'équation

$$(\nabla \eta(w_{i+1}^n) - \nabla \eta(w_i^n))^T \mathcal{F}_{i+\frac{1}{2}} = \psi(w_{i+1}^n) - \psi(w_i^n), \quad \forall i \in \mathbb{Z}. \quad (1.24)$$

Un flux numérique vérifiant l'équation ci-dessus donne automatiquement un schéma numérique conservant l'entropie au niveau semi-discret. Cette conservation s'écrit sous la forme de l'égalité suivante :

$$\frac{d}{dt} \sum_{i \in \mathbb{Z}} \eta(w_i(t)) \Delta x = 0, \quad t > 0. \quad (1.25)$$

Néanmoins, un calcul direct et explicite de  $\mathcal{F}_{i+\frac{1}{2}}$  à partir de (1.24) n'est pas toujours possible [206] et demande nécessairement la connaissance du potentiel  $\psi$ . Ce potentiel est explicitement établi pour les systèmes usuels comme Saint Venant ou Euler, mais son obtention pour un système quelconque est numériquement difficile [94]. Enfin, pour passer de la conservation de l'entropie (1.25) à une stabilité entropique semi-discrète (1.23), un flux numérique satisfaisant l'équation (1.24) doit être corrigé. Les approches typiques en ce sens sont l'ajout d'un opérateur de diffusion [45, 51, 94, 205] ou bien une utilisation judicieuse des conditions de bords [7, 44].

Dans le cadre des éléments finis Galerkin discontinus (*Discontinuous Galerkin*, DG) [56, 195, 196], les variables entropiques sont unifiées au formalisme plus large des opérateurs *Summation By Part* (SBP) [7, 55]. Ce cadre théorique autorise des calculs semi-discrets analogues à ceux du continu ce qui simplifie considérablement les preuves de stabilité. L'ordre du schéma se sélectionne ensuite simplement avec le choix des fonctions de base des éléments finis considérés. C'est dans ce formalisme que des schémas DG d'ordre deux à quatre satisfaisant des inégalités d'entropie semi discrètes locales (1.22) ou globales (1.23) sont respectivement présentés dans [137, 182] et dans [7, 45].

La méthode des éléments finis peut également servir de cadre pour les schémas numériques aux *résidus distribués* (RD) [2, 69]. L'originalité de ces schémas repose d'abord sur la substitution de  $w$  par une approximation au sein de la formulation faible des équations (1.1). Cette substitution permet ensuite de construire un résidu distribué qui définit le schéma numérique et qui permet le calcul de l'approximation considérée.

Plusieurs schémas RD monotones, d'ordre deux à quatre sont par exemple proposés dans [4, 6] mais les résidus de stabilité présentés dans [3, 8, 9] permettent désormais d'obtenir une stabilité entropique. En s'appuyant sur ces résidus et en notant  $t^0$  la date initiale de simulation et  $t^N$  la date finale, les auteurs de [5] proposent par exemple un schéma RD d'ordre quelconque pour les équations d'Euler et qui satisfait

$$\sum_{i \in \mathbb{Z}} \eta(w_i^N) \Delta x \leq \sum_{i \in \mathbb{Z}} \eta(w_i^0) \Delta x. \quad (1.26)$$

Le point commun de la reconstruction polynomiale et de la méthode des variables entropiques est qu'elles utilisent toutes deux des limiteurs [4, 55, 57, 94, 116, 182, 194, 196, 219]. La recherche bibliographique n'a pas montré l'existence d'une définition formelle de technique de limitation mais on peut néanmoins citer l'exemple canonique des schémas MUSCL [219]. Les oscillations parasites générées artificiellement par ces schémas sont contrôlées par un limiteur de pente  $\phi$  vérifiant  $\phi(r)|_{r \leq 0} = 0$  et  $\phi(r)|_{r > 0} \leq 2 \min(1, r)$ . Une fois choisi, le limiteur est introduit dans l'approximation de la dérivée de la manière suivante :

$$\partial_x w(x_{i+\frac{1}{2}}, t^n) \approx \phi\left(\frac{w_i^n - w_{i-1}^n}{\Delta x} \frac{\Delta x}{w_{i+1}^n - w_i^n}\right) \frac{w_{i+1}^n - w_i^n}{\Delta x}, \quad \forall i \in \mathbb{Z}.$$

Par conséquent, une technique de limitation peut être interprétée comme une correction *ad-hoc* forçant une propriété de stabilité souhaitée. Le choix d'un limiteur n'est pas simple. Il dépend fortement du problème étudié et peut engendrer une dégradation de l'ordre théorique [16, 196, 227].

Ainsi, plusieurs champs sont explorés pour conjuguer discrétisation spatiale d'ordre élevé et stabilité entropique. Selon la bibliographie présentée dans ce mémoire, des schémas numériques d'ordre élevé vérifiant (1.19) ne semblent pas exister et très peu en vérifient une version globale donnée par (1.21). Bien que plus faible, cette stabilité entropique globale a déjà produit numériquement des solutions ayant des comportements proches de celles présentant une stabilité entropique locale [49, 125, 127].

La très large majorité des méthodes donnent naissance à des stabilités entropiques semi-discrètes (1.22)-(1.23) ou incomplètes (1.20). Cependant, des expériences numériques exposées dans [77] et dans

[34] montrent respectivement qu'un schéma vérifiant une inégalité d'entropie incomplète ou semi-discrète ne jauge pas totalement certains artefacts numériques. Ce constat contraint donc à travailler avec des discrétisations temps et espace complètes et avec des inégalités d'entropie de la forme (1.19), (1.21). Le transfert d'une stabilité entropique en temps discret constitue une difficulté majeure. Les aspects et les causes de cette difficulté sont présentés dans la prochaine section.

### 1.3.4 L'obstacle de la discrétisation en temps

*"In contrast, the entropy stability of explicit time discretization is by large an open problem"*

T. Chen, 2020, [56].

Les discrétisations temporelles explicites sont intrinsèquement contraires aux inégalités d'entropie et sont donc sujettes aux instabilités [159, 230]. Le contrôle de ces instabilités est parfois possible [16, 230] mais peut conduire à des conditions CFL très restrictives voire rédhibitoires. D'une manière générale, la stabilité entropique discrète des schémas d'ordre élevé et explicites en temps est véritablement problématique. Dans le cadre des schémas MUSCL et (W)ENO, cette thématique fait encore aujourd'hui l'objet de conjectures [10, 230] et dans le cadre DG, elle n'est quasiment pas explorée [56]. Remarquons toutefois que si on dispose d'un schéma explicite d'ordre un en temps et vérifiant une stabilité entropique discrète, alors on peut effectuer une montée en ordre temporel à l'aide des méthodes de *Runge-Kutta Strong Stability Preserving* (RKSSP) [107, 109, 110]. Ces méthodes explicites sont construites à partir de combinaisons convexes de discrétisations du premier ordre. En s'appuyant sur la convexité des entropies  $\eta$ , on peut alors montrer que ces méthodes RKSSP préservent les stabilités entropiques [76, 230].

La discrétisation du temps peut en outre se réaliser avec des méthodes implicites [40, 122, 137, 144, 197, 228] qui sont par nature plus stables que leurs homologues explicites [230]. Ces méthodes demandent néanmoins à chaque itération en temps, la résolution d'un système d'équations non linéaires couplant généralement toutes les inconnues du maillage. Elles entraînent de plus une augmentation significative d'artefact comme la diffusion numérique ce qui peut dégrader la qualité générale des simulations. Ces discrétisations temporelles implicites sont formalisées dans le cadre DG sous le nom de *Discontinuous Galerkin Space and Time* (DGST). Dans ce formalisme, le temps est considéré comme une variable d'espace supplémentaire [96, 125, 127] ce qui conduit à des estimations globales temps et espace de la forme (1.26).

Enfin, des méthodes de discrétisations temporelles de type Runge Kutta ont aussi été proposées pour transférer une inégalité d'entropie semi-discrète (1.22)-(1.23) vers une inégalité discrète globale de la forme (1.21). Explicites dans le cas  $L^2$  [133], ces méthodes requièrent néanmoins, dans le cas d'une entropie  $\eta$  quelconque, la résolution d'une équation non linéaire à chaque itération en temps [178].

### 1.3.5 Le terme source de topographie

L'introduction d'une fonction  $z$  quelconque dans les équations de Saint Venant (1.10) constitue une difficulté réelle à l'élaboration de schémas numériques approchant la quantité  $w = (h, hu)^T$ .

La principale difficulté provient de ce que la solution du problème de Riemann pour les équations de Saint Venant (1.10) n'est pas unique [145, 146]. Par conséquent, les simulations numériques peuvent donc converger vers différentes solutions [34, 42]. L'étude du problème de Riemann avec une fonction  $z$  discontinue est l'objet de plusieurs travaux dans lesquels des traitements différents des solutions multiples sont avancés. Dans [12] l'unicité est assurée par la continuité du débit  $hu$  alors que dans [20] la conservation de la masse et des relations de Rankine–Hugoniot spécifiques sont employées. De leurs côtés, les auteurs de [231] proposent de maximiser le taux de croissance de la quantité  $u^2/2 + gh$ . Enfin, d'autres manières de sélectionner une solution sont proposées dans [11, 59, 114, 146]. Toutes ces études produisent des méthodes numériques qui demandent malheureusement la résolution d'équations fortement non linéaires à chaque itération en temps et pour chaque cellule du maillage. À partir de ces constats et en notant  $\hat{w} = (h, hu, z)^T$ , il demeure difficile de construire des schémas numériques

*well-balanced* et à stabilité entropique discrète pour le couple

$$\hat{\eta}(\hat{w}) = \frac{(hu)^2}{2h} + \frac{gh^2}{2} + ghz, \quad \hat{G}(\hat{w}) = \frac{(hu)^3}{2h^2} + g(hu)(h+z). \quad (1.27)$$

La littérature consacrée à ces schémas reste toutefois vaste. Les paragraphes ci-dessous en proposent une présentation organisée en quatre catégories distribuées selon le critère de stabilité entropique.

Tout d'abord, la première catégorie se compose de schémas *well-balanced* pour lesquels aucun résultat de stabilité entropique n'est établi. La plupart de ces schémas [98, 103, 152, 165, 166, 175] s'écrivent sous la forme de schéma de type Godunov tels qu'évoqués dans la section 1.3.2. La propriété *well-balanced* est alors assurée grâce à des approximations constantes par morceaux et par une discrétisation judicieuse du terme source. D'autres schémas sont construits à partir de méthodes spectrales [135] où d'éléments finis [30, 151, 226]. D'une manière générale, des montées en ordre de ces schémas sont proposées [48, 103, 129, 135, 152, 165, 166].

Ensuite, la seconde catégorie concerne les schémas vérifiant une inégalité d'entropie semi-discrète pour le couple  $(\hat{\eta}, \hat{G})$  donné par (1.27). En notant  $\hat{\mathcal{G}}_{i+\frac{1}{2}}$  une approximation formelle de  $\hat{G}(\hat{w}(x_{i+\frac{1}{2}}, t^n))$ , ces inégalités s'écrivent

$$\frac{d}{dt} \hat{\eta}(\hat{w}_i(t)) + \frac{\hat{\mathcal{G}}_{i+\frac{1}{2}}(t) - \hat{\mathcal{G}}_{i-\frac{1}{2}}(t)}{\Delta x} \leq 0, \quad t > 0.$$

On peut alors inclure à cette catégorie, des schémas DG [100, 177, 223, 225] d'ordre élevé, *well-balanced* pour le lac au repos et conçus à partir d'une version des opérateurs SBP décrits dans la section 1.3.3. C'est également dans cette seconde catégorie que se trouve la méthode de reconstruction hydrostatique [13]. Cette méthode est très utilisée et repose sur une discrétisation du terme source déduite d'une analyse formelle des cas  $u \ll \sqrt{gh}$ . Pour un flux numérique quelconque donné, elle permet de restituer la propriété *well-balanced* pour le lac au repos (1.14) [13, 14, 54]. Elle est de plus facile à programmer [13]. Malheureusement, un résultat de [14] montre qu'une méthode numérique construite avec les principes de la reconstruction hydrostatique ne peut donner naissance qu'à des stabilités entropiques semi-discrètes. On peut enfin joindre à cette catégorie, des schémas *well-balanced* pour des solutions stationnaires à vitesse non nulle [35, 93].

La troisième catégorie comprend des schémas vérifiant une inégalité d'entropie locale incomplète de la forme

$$\frac{\hat{\eta}(\hat{w}_i^{n+1}) - \hat{\eta}(\hat{w}_i^n)}{\Delta t} + \frac{\hat{\mathcal{G}}_{i+\frac{1}{2}} - \hat{\mathcal{G}}_{i-\frac{1}{2}}}{\Delta x} \leq \mathcal{O}(\Delta x^2). \quad (1.28)$$

De tels résultats ont d'abord pu être montrés sur un schéma cinétique [14] *well balanced* pour le lac au repos et ensuite sur des schémas de type Godunov. Dans ce dernier cas, on compte un schéma [15] *well-balanced* pour le lac au repos et des schémas [23, 25, 26] *well balanced* pour un état stationnaire quelconque (1.13).

Enfin, la dernière catégorie concerne des schémas *well-balanced* pour le lac au repos et présentant une stabilité entropique discrète complète. Dans [32, 34], Bouchut *et al* proposent des schémas vérifiant une inégalité d'entropie locale qui s'écrit

$$\frac{\hat{\eta}(\hat{w}_i^{n+1}) - \hat{\eta}(\hat{w}_i^n)}{\Delta t} + \frac{\hat{\mathcal{G}}_{i+\frac{1}{2}} - \hat{\mathcal{G}}_{i-\frac{1}{2}}}{\Delta x} \leq 0. \quad (1.29)$$

Ces schémas s'appuient notamment sur des techniques de relaxation et sont d'après l'auteur lui-même complexes à mettre en œuvre : *"it is possible to obtain the solution explicitly. However we shall not describe this in detail here, because this would be too lengthy"*, F. Bouchut, 2004, [32]. Enfin les auteurs de [86, 126] proposent des schémas *well-balanced* pour le lac au repos et satisfaisant respectivement des estimations globales de la forme (1.21), (1.26) écrites pour le couple  $(\hat{\eta}, \hat{G})$  donné par (1.27).

D'après le panorama précédent, il est clair que la littérature regorge de méthodes pour approcher les solutions des équations de Saint Venant (1.10). Si la propriété *well-balanced* a largement été exhibée, la propriété de stabilité entropique est quant à elle plus rare et souvent incomplète. Ceci s'explique certainement par la démarche de conception des schémas numériques qui s'articule d'abord autour du

critère *well-balanced*, la stabilité entropique étant analysée dans un second temps. Cette seconde tâche est ardue compte-tenu du caractère non linéaire de l'entropie donnée par (1.27). On peut en outre remarquer que l'étude bibliographique n'a pas montré l'existence d'un schéma numérique vérifiant une inégalité d'entropie locale complète (1.19) et *well-balanced* pour tous les équilibres (1.13).

Si on s'intéresse maintenant au système de Ripa (1.15) alors comme précédemment, la solution au problème de Riemann n'est pas unique [208, 209]. En revanche la littérature consacrée à l'approximation numérique des solutions de (1.15) est moins riche que celle dédiée aux équations de Saint Venant. Les schémas approchant les solutions du système de Ripa (1.15) peuvent globalement se répartir selon deux groupes. Le premier contient des schémas *well-balanced* n'ayant aucun résultat de stabilité. Ils se répartissent autour des équilibres à vitesses nulles dans [1, 95, 115, 124, 216] et autour des équilibres mobiles dans [38]. Le second groupe concerne des schémas *well-balanced* pour le lac au repos, d'ordre élevé et pour lesquels une propriété de préservation du domaine  $\Omega$  est établie. Ces schémas se basent sur des montées en ordre de type MUSCL [58, 186, 187], DG [131, 150, 176] ou bien (W)ENO [181]. Plus récemment, la propriété *well-balanced* a été obtenue pour des équilibres mouvant (1.16) avec des schémas de type MUSCL [82].

D'après la description ci-dessus, il est clair que des schémas approchant les solutions de (1.15) et vérifiant une stabilité entropique font défaut. Il ne semble exister que deux schémas numériques dédiés aux équations de Ripa (1.15) et relatant des stabilités entropiques. Le premier, décrit dans [203], est *well-balanced* pour le lac au repos (1.17) et vérifie une inégalité d'entropie locale (1.19) démontrée seulement dans le cas où  $z$  est constante. Le second, introduit dans [78] est *well-balanced* pour des équilibres à vitesses nulles et satisfait une inégalité d'entropie incomplète de la forme (1.28).

Ainsi, énormément de pistes ont été explorées pour allier précision et stabilité des schémas numériques. Cependant, la précision caractérisée par une discrétisation en temps explicite, des propriétés *well-balanced* et des discrétisations d'ordre élevé en espace est contraire à la stabilité entropique souhaitée. Les chapitres suivants vont présenter des schémas du premier ordre explicites en temps, stables et précis. Ces schémas vont être élaborés à partir de la méthode des volumes finis et en particulier avec des schémas de type Godunov [118] détaillés dans la prochaine section.

## 1.4 Approximations numériques avec la méthode des volumes finis

### 1.4.1 Méthode des volumes finis

La méthode des volumes finis utilise d'abord un maillage du temps dans  $\mathbb{R}^+$ . Un tel maillage est donné par une séquence  $(t^n)_{n \in \mathbb{N}}$  de pas uniforme  $\Delta t > 0$  et qui vérifie

$$t^0 = 0, \quad t^{n+1} = t^n + \Delta t, \quad \forall n \in \mathbb{N}.$$

Ensuite, cette méthode requière un maillage de l'espace dans  $\mathbb{R}$  défini par un ensemble d'*interfaces*  $(x_{i+\frac{1}{2}})_{i \in \mathbb{Z}}$  de pas uniforme  $\Delta x > 0$  tel que

$$x_{i+\frac{1}{2}} = x_{i-\frac{1}{2}} + \Delta x, \quad x_i = x_{i-\frac{1}{2}} + \Delta x/2, \quad \forall i \in \mathbb{Z}.$$

En s'appuyant sur ces interfaces, la notation  $C_i$  désigne les *cellules* du maillage, chacune de longueur  $\Delta x > 0$  et définie par l'intervalle ouvert  $]x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}[$ . De la même manière,  $C_{i+\frac{1}{2}}$  désigne l'intervalle ouvert  $]x_i, x_{i+1}[$ .

Avec la donnée d'un maillage du temps et de l'espace, la méthode des volumes finis permet d'approcher les solutions d'un système hyperbolique homogène de  $m$  lois de conservation en une dimension d'espace de la forme

$$\begin{cases} \partial_t w + \partial_x f(w) = 0, & t > 0, \quad x \in \mathbb{R}, \\ w(x, t = 0) = w_0(x). \end{cases} \quad (1.30)$$

Tout d'abord, la donnée initiale  $w_0$  est discrétisée à l'aide d'une séquence  $(w_i^0)_{i \in \mathbb{Z}}$  dans  $\mathbb{R}^m$  définie par

$$w_i^0 = \frac{1}{\Delta x} \int_{C_i} w_0(x) dx, \quad \forall i \in \mathbb{Z}.$$

Ensuite, en considérant  $(w_i^n)_{i \in \mathbb{Z}}$  une suite d'éléments de  $\mathbb{R}^m$  ainsi que  $\chi_{\mathcal{A}}$  la fonction indicatrice d'un ensemble  $\mathcal{A}$  quelconque, il est possible de définir  $w_{\Delta} : \mathbb{R} \times \mathbb{N} \rightarrow \mathbb{R}^m$  une fonction constante par morceaux telle que

$$w_{\Delta}(x, t^n) = \sum_{i \in \mathbb{Z}} w_i^n \chi_{C_i}(x). \quad (1.31)$$

Par construction,  $w_{\Delta}(x, t^0)$  est une approximation constante par morceaux de la fonction  $w_0$  qui s'identifie à  $w(\cdot, t^0)$ . Ainsi, pour obtenir une approximation de  $w(\cdot, t^n)$  pour tout  $n$  dans  $\mathbb{N}$ , il est suffisant de définir un schéma numérique permettant de calculer  $(w_i^{n+1})_{i \in \mathbb{Z}}$  à partir de  $(w_i^n)_{i \in \mathbb{Z}}$ . Une définition d'un schéma numérique volumes finis est donnée ci-après.

**Définition 1.4.1.1** (Schéma numérique volumes finis conservatif). *Soit  $l$  un entier et  $\mathcal{F} : (\mathbb{R}^m)^{(2l+2)} \rightarrow \mathbb{R}^m$  une fonction de flux numérique continue. Soit  $u : \mathbb{R} \rightarrow \mathbb{R}^m$  une fonction régulière et notons*

$$u_i = \frac{1}{\Delta x} \int_{C_i} u(x) dx, \quad \forall i \in \mathbb{Z}.$$

Un schéma numérique volumes finis approchant les solutions de (1.30) s'écrit

$$w_i^{n+1} = w_i^n - \frac{\Delta t}{\Delta x} (\mathcal{F}(w_{i-l}^n, \dots, w_{i+l+1}^n) - \mathcal{F}(w_{i-l-1}^n, \dots, w_{i+l}^n)), \quad \forall i \in \mathbb{Z}. \quad (1.32)$$

Un tel schéma est dit

i) *conservatif* ce qui signifie qu'aux contributions des termes de bord près, la séquence  $(w_i^{n+1})_{i \in \mathbb{Z}}$  vérifie

$$\sum_{i \in \mathbb{Z}} w_i^{n+1} \Delta x = \sum_{i \in \mathbb{Z}} w_i^n \Delta x.$$

ii) *consistant* si le flux numérique  $\mathcal{F}$  vérifie  $\mathcal{F}(w, \dots, w) = f(w)$ , pour tout  $w$  dans  $\Omega$ ,

iii) *d'ordre  $k$*  si, pour une interface  $x_{i+\frac{1}{2}}$  fixée, le flux numérique  $\mathcal{F}$  vérifie également

$$\mathcal{F}(u_{i-l}, \dots, u_{i+l+1}) = f(u(x_{i+\frac{1}{2}})) + \mathcal{O}(\Delta x^k).$$

L'écriture conservative (1.32) est entièrement définie par la donnée de  $\mathcal{F}$ . Le caractère conservatif est usuel et ne sera pas davantage détaillé. Le lecteur intéressé pourra se référer à [104, Chapitre 3]. La condition de consistance décrite dans la Définition 1.4.1.1-ii) assure que le schéma volume finis (1.32) préserve les constantes de  $\Omega$  qui sont des solutions du système homogène (1.30). La notion d'ordre spatiale de la Définition 1.4.1.1-iii) est une formulation forte et formelle de la notion introduite dans [32, Proposition 2.6]. D'autres définitions d'ordre spatial sont possibles et le lecteur intéressé pourra par exemple consulter [18]. Comme il n'y a pas de confusion possible, la notation  $\mathcal{F}_{i+\frac{1}{2}} = \mathcal{F}(w_{i-l}^n, \dots, w_{i+l+1}^n)$  sera désormais employée et par conséquent (1.32) se lit

$$w_i^{n+1} = w_i^n - \frac{\Delta t}{\Delta x} (\mathcal{F}_{i+\frac{1}{2}} - \mathcal{F}_{i-\frac{1}{2}}), \quad \forall i \in \mathbb{Z}.$$

Avec les éléments ci-dessus, il est possible de définir la stabilité entropique discrète d'un schéma numérique volumes finis (1.32) de la manière suivante :

**Définition 1.4.1.2** (Stabilité entropique discrète). *Soit un système hyperbolique de la forme (1.30) équipé d'un couple entropie, flux d'entropie  $(\eta, G)$  et dont les solutions vérifient une inégalité d'entropie de la forme*

$$\partial_t \eta(w) + \partial_x G(w) \leq 0.$$

i) *Soient  $l$  un entier et  $\mathcal{G} : (\mathbb{R}^m)^{(2l+2)} \rightarrow \mathbb{R}$  une fonction continue. La fonction  $\mathcal{G}$  est dite consistante avec  $G$  si*

$$\mathcal{G}(w, \dots, w) = G(w), \quad \forall w \in \Omega.$$

ii) Un schéma numérique volumes finis (1.32) approchant les solutions de (1.30) est dit à stabilité entropique locale pour le couple  $(\eta, G)$  s'il existe une fonction  $\mathcal{G}$  consistante avec  $G$  et telle que

$$\frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} + \frac{\mathcal{G}(w_{i-l}^n, \dots, w_{i+l+1}^n) - \mathcal{G}(w_{i-l-1}^n, \dots, w_{i+l}^n)}{\Delta x} \leq 0, \quad \forall i \in \mathbb{Z}. \quad (1.33)$$

Dans ce cas, la fonction  $\mathcal{G}$  est appelée flux numérique d'entropie.

iii) Un schéma numérique volumes finis (1.32) approchant les solutions de (1.30) est dit à stabilité entropique globale pour le couple  $(\eta, G)$  si

$$\sum_{i \in \mathbb{Z}} \frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} \Delta x \leq 0.$$

Comme cela ne porte pas à confusion, la notation  $\mathcal{G}_{i+\frac{1}{2}} = \mathcal{G}(w_{i-l}^n, \dots, w_{i+l+1}^n)$  est adoptée et par conséquent l'inégalité d'entropie discrète (1.33) se lit désormais

$$\frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} + \frac{\mathcal{G}_{i+\frac{1}{2}} - \mathcal{G}_{i-\frac{1}{2}}}{\Delta x} \leq 0, \quad \forall i \in \mathbb{Z}.$$

L'un des premiers schémas numériques volumes finis vérifiant une inégalité d'entropie discrète (1.33) est le schéma de Godunov introduit dans [106] et dédié aux systèmes d'équations homogènes.

### 1.4.2 Schéma de Godunov

Pour un système hyperbolique homogène donné, considérons la séquence de problèmes de Riemann posée sur les intervalles  $(C_{i+\frac{1}{2}})_{i \in \mathbb{Z}}$  et définie par

$$\begin{cases} \partial_t w + \partial_x f(w) = 0, & t \in ]t^n, t^n + \Delta t], \quad x \in C_{i+\frac{1}{2}}, \\ w(x, t = t^n) = \begin{cases} w_i^n & \text{si } x < x_{i+\frac{1}{2}}, \\ w_{i+1}^n & \text{sinon,} \end{cases} \end{cases} \quad \forall i \in \mathbb{Z}. \quad (1.34)$$

La construction du schéma de Godunov s'articule autour de l'étude des problèmes de Riemann ci-dessus et elle est constituée de deux étapes distinctes détaillées ci-dessous.

La première étape *d'évolution* consiste à résoudre les problèmes de Riemann (1.34) en supposant qu'ils n'interagissent pas entre eux. En notant  $\{\mu_k(w)\}_{k \in \{1, \dots, m\}}$  l'ensemble des  $m$  valeurs propres de la matrice  $\nabla f(w)$ , cette indépendance des problèmes de Riemann se traduit par une contrainte sur  $\Delta t > 0$  qui s'écrit

$$\frac{\Delta t}{\Delta x} \max_{\substack{\alpha \in \{i, i+1\}, \\ k \in \{1, \dots, m\}}} |\mu_k(w_\alpha^n)| \leq \frac{1}{2}. \quad (1.35)$$

Lorsqu'elle est satisfaite, l'inégalité ci-dessus permet d'étudier (1.34) comme une juxtaposition de problèmes de Riemann indépendants chacun posé à une interface. Par conséquent et pour  $i$  fixé, la solution d'un problème de Riemann (1.34) localisé autour d'une interface  $x_{i+\frac{1}{2}}$  ne dépend que des quantités  $w_i^n$  et  $w_{i+1}^n$  situées de part de d'autre de l'interface. Ainsi, pour  $i$  dans  $\mathbb{Z}$  donné et en définissant la fonction

$$\xi_{i+\frac{1}{2}}^n : (x, t) \mapsto \frac{x - x_{i+\frac{1}{2}}}{t^n + t}, \quad \forall (x, t) \in C_{i+\frac{1}{2}} \times [0, \Delta t],$$

il est possible d'écrire la solution de (1.34) de la manière suivante :

$$(i, x, t) \mapsto w_{\mathcal{R}}(\xi_{i+\frac{1}{2}}^n(x, t), w_i^n, w_{i+1}^n), \quad \forall (i, x, t) \in \mathbb{Z} \times C_{i+\frac{1}{2}} \times [0, \Delta t].$$

Par définition et à  $i$  fixé, la fonction  $(x, t) \mapsto w_{\mathcal{R}}(\cdot, w_i^n, w_{i+1}^n)$  vérifie

$$w_{\mathcal{R}}(\xi_{i+\frac{1}{2}}^n(x, 0), w_i^n, w_{i+1}^n) = \begin{cases} w_i^n & \text{si } x < x_{i+\frac{1}{2}}, \\ w_{i+1}^n & \text{sinon.} \end{cases}$$

Cette fonction vérifie également le Lemme de Harten rappelé en Section 1.1 et qui donne l'égalité suivante :

$$\frac{1}{\Delta x} \int_{C_{i+\frac{1}{2}}} w_{\mathcal{R}}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n) dx = \frac{w_i^n + w_{i+1}^n}{2} - \frac{\Delta t}{\Delta x} (f(w_{i+1}^n) - f(w_i^n)). \quad (1.36)$$

De plus, puisque les constantes sont trivialement solutions des systèmes d'équations homogènes, il résulte

$$w_{\mathcal{R}}(\xi_{i+\frac{1}{2}}^n(x, t), w, w) = w, \quad \forall w \in \mathbb{R}^d. \quad (1.37)$$

Avec la collection des solutions  $w_{\mathcal{R}}(\cdot, w_i^n, w_{i+1}^n)$ , on dispose par construction d'une approximation de  $w(\cdot, t)$  pour  $t$  dans  $[t^n, t^n + \Delta t]$ . Cette approximation est notée  $w_{\Delta, \mathcal{R}} : \mathbb{R} \times [t^n, t^n + \Delta t] \rightarrow \mathbb{R}^m$  et s'écrit

$$w_{\Delta, \mathcal{R}}(x, t^n + t) = \sum_{i \in \mathbb{Z}} w_{\mathcal{R}}(\xi_{i+\frac{1}{2}}^n(x, t), w_i^n, w_{i+1}^n) \chi_{C_{i+\frac{1}{2}}}(x).$$

À partir de l'expression ci-dessus, la seconde étape de *projection* du schéma de Godunov consiste en une projection  $L^2$  de  $w_{\Delta, \mathcal{R}}(\cdot, t^n + \Delta t)$  sur l'espace des fonctions constantes par morceaux. Cette projection conduit à l'expression de  $(w_i^{n+1})_{i \in \mathbb{Z}}$  qui suit

$$w_i^{n+1} = \frac{1}{\Delta x} \int_{C_i} w_{\Delta, \mathcal{R}}(x, t^n + \Delta t) dx, \quad \forall i \in \mathbb{Z}. \quad (1.38)$$

Plus précisément, pour une cellule  $C_i$  donnée et en vertu de (1.35), la quantité  $w_i^{n+1}$  est seulement influencée par les deux problèmes de Riemann situés aux frontières de la cellule. Par conséquent pour tout  $i$  dans  $\mathbb{Z}$ , il suit

$$w_i^{n+1} = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_i} w_{\mathcal{R}}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), w_{i-1}^n, w_i^n) dx + \frac{1}{\Delta x} \int_{x_i}^{x_{i+\frac{1}{2}}} w_{\mathcal{R}}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n) dx. \quad (1.39)$$

Il est possible de récrire le schéma de Godunov ci-dessus sous une forme conservative telle que proposée dans la Définition 1.4.1.1 Cette reformulation équivalente est par exemple conduite dans [118] et donne l'expression du flux numérique suivante :

$$\mathcal{F}_{i+\frac{1}{2}} = f(w_{\mathcal{R}}(0, w_i^n, w_{i+1}^n)).$$

En s'appuyant sur l'égalité (1.37), la consistance du flux numérique ci-dessus est immédiate. Le schéma de Godunov (1.39) satisfait également une inégalité d'entropie rappelée dans le lemme ci-après.

**Lemme 1.4.2.1** (Stabilité entropique locale du schéma de Godunov). *Soit un système hyperbolique homogène de  $d$  lois de conservation muni d'un couple entropie convexe, flux d'entropie  $(\eta, G)$  et dont les solutions sont approchées par le schéma de Godunov (1.39). Un tel schéma satisfait une inégalité d'entropie discrète locale qui s'écrit*

$$\frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} + \frac{G(w_{\mathcal{R}}(0^-, w_i^n, w_{i+1}^n)) - G(w_{\mathcal{R}}(0^+, w_{i-1}^n, w_i^n))}{\Delta x} \leq 0, \quad \forall i \in \mathbb{Z}. \quad (1.40)$$

*Démonstration.* Par construction,  $(x, t) \mapsto w_{\Delta, \mathcal{R}}(x, t^n + t)$  vérifie l'inégalité

$$\int_0^{\Delta t} \int_{C_i} (\partial_t \eta(w_{\Delta, \mathcal{R}}) + \partial_x G(w_{\Delta, \mathcal{R}})) dx dt \leq 0.$$

En considérant la fonction  $w_{\Delta}$  définie dans (1.31) et en exploitant l'égalité  $w_{\Delta, \mathcal{R}}(\cdot, t^n) = w_{\Delta}(\cdot, t^n)$ , il suit

$$\int_0^{\Delta t} \int_{C_i} (\partial_t \eta(w_{\Delta, \mathcal{R}}) + \partial_x G(w_{\Delta, \mathcal{R}})) dx dt$$



$$\begin{aligned}
&= \int_{C_i} \eta(w_{\Delta, \mathcal{R}}(x, t^{n+1})) dx - \int_{C_i} \eta(w_{\Delta}(x, t^n)) dx + \int_0^{\Delta t} G(w_{\Delta, \mathcal{R}}(x_{i+\frac{1}{2}}^-, t)) dt - \int_0^{\Delta t} G(w_{\Delta, \mathcal{R}}(x_{i-\frac{1}{2}}^+, t)) dt, \\
&= \int_{C_i} \eta(w_{\Delta, \mathcal{R}}(x, t^{n+1})) dx - \Delta x \eta(w_i^n) + \int_0^{\Delta t} G(w_{\Delta, \mathcal{R}}(x_{i+\frac{1}{2}}^-, t)) dt - \int_0^{\Delta t} G(w_{\Delta, \mathcal{R}}(x_{i-\frac{1}{2}}^+, t)) dt.
\end{aligned} \tag{1.41}$$

Les quantités  $G(w_{\Delta, \mathcal{R}}(x_{i+\frac{1}{2}}^-, t))$  et  $G(w_{\Delta, \mathcal{R}}(x_{i-\frac{1}{2}}^+, t))$  du calcul ci-dessus désignent respectivement les limites

$$G(w_{\Delta, \mathcal{R}}(x, t)) \xrightarrow[x < x_{i+\frac{1}{2}}]{x \rightarrow x_{i+\frac{1}{2}}} G(w_{\Delta, \mathcal{R}}(x_{i+\frac{1}{2}}^-, t)), \quad G(w_{\Delta, \mathcal{R}}(x, t)) \xrightarrow[x > x_{i-\frac{1}{2}}]{x \rightarrow x_{i-\frac{1}{2}}} G(w_{\Delta, \mathcal{R}}(x_{i-\frac{1}{2}}^+, t)).$$

D'après [118, 204], il est admis que ces limites existent. En utilisant l'expression des solutions des problèmes de Riemann  $w_{\mathcal{R}}$ , ces limites sont alors constantes et on en déduit

$$\int_0^{\Delta t} G(w_{\Delta, \mathcal{R}}(x_{i+\frac{1}{2}}^-, t)) dt = \int_0^{\Delta t} G(w_{\mathcal{R}}(0^-, w_i^n, w_{i+1}^n)) dt = \Delta t G(w_{\mathcal{R}}(0^-, w_i^n, w_{i+1}^n)).$$

Comme un calcul similaire peut être conduit pour les quantités situées aux voisinages de  $x_{i-\frac{1}{2}}$ , on obtient à l'aide de (1.41) l'égalité suivante :

$$\begin{aligned}
\int_0^{\Delta t} \int_{C_i} (\partial_t \eta(w_{\Delta, \mathcal{R}}) + \partial_x G(w_{\Delta, \mathcal{R}})) dx dt &= \int_{C_i} \eta(w_{\Delta, \mathcal{R}}(x, t^{n+1})) dx - \Delta x \eta(w_i^n) \\
&+ \Delta t G(w_{\mathcal{R}}(0^-, w_i^n, w_{i+1}^n)) - \Delta t G(w_{\mathcal{R}}(0^+, w_{i-1}^n, w_i^n)).
\end{aligned} \tag{1.42}$$

Selon [173], l'inégalité de Jensen demeure vraie pour les fonctions convexes définies sur un espace convexe  $\Omega$  inclus dans  $\mathbb{R}^m$ . Comme  $\eta$  satisfait les conditions requises, l'application de ce résultat à la définition (1.38) donne

$$\eta(w_i^{n+1}) \leq \frac{1}{\Delta x} \int_{C_i} \eta(w_{\Delta, \mathcal{R}}(x, t^{n+1})) dx.$$

En combinant l'inégalité ci-dessus à (1.42) puis en divisant par  $\Delta t > 0$ , on déduit que la séquence  $(w_i^{n+1})_{i \in \mathbb{Z}}$  mise à jour par le schéma de Godunov satisfait

$$\frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} + \frac{G(w_{\mathcal{R}}(0^-, w_i^n, w_{i+1}^n)) - G(w_{\mathcal{R}}(0^+, w_{i-1}^n, w_i^n))}{\Delta x} \leq 0.$$

L'inégalité ci-dessus correspond à la forme (1.40) annoncée. D'après (1.37), il suit immédiatement la consistance  $G(w_{\mathcal{R}}(0^\pm, w, w)) = G(w)$  et par conséquent (1.40) est une inégalité d'entropie discrète locale ce qui conclut la preuve.  $\square$

Le schéma de Godunov présenté dans cette section satisfait une inégalité d'entropie donnée par (1.40) mais est conditionné à la détermination des solutions  $w_{\mathcal{R}}$  des problèmes de Riemann (1.34). Malheureusement, l'obtention de  $w_{\mathcal{R}}$  est en général très difficile voire impossible. C'est à partir de ce constat que se bâtissent les schémas de *type* Godunov qui reposent sur une approximation de  $w_{\mathcal{R}}$ . L'élaboration et quelques exemples de ces schémas de type Godunov sont décrits dans les sections suivantes.

### 1.4.3 Schéma de *type* Godunov pour des systèmes d'équations homogènes

Pour un système hyperbolique homogène donné, considérons la juxtaposition de problèmes de Riemann (1.34). Pour un indice  $i$  de  $\mathbb{Z}$  fixé, un schéma numérique de type Godunov consiste à substituer la solution exacte  $w_{\mathcal{R}}$  décrite dans la section 1.4.2 par une approximation. Cette approximation est une fonction appelée *solveur de Riemann* approché et elle est notée  $\tilde{w} : \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ . Par construction, on dispose d'une séquence de solveur  $(\tilde{w}(\cdot, w_i^n, w_{i+1}^n))_{i \in \mathbb{Z}}$  et on impose que ces solveurs n'interagissent pas entre eux. Cette contrainte de non interaction illustrée par la Figure 1.2 se traduit par une condition restrictive sur le pas de temps appelée *condition CFL de non interaction*.

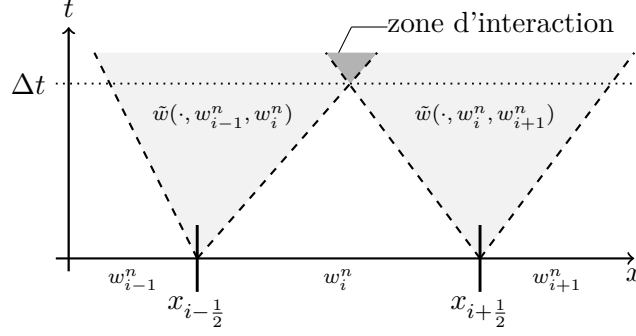


FIGURE 1.2 – Juxtaposition dans le plan  $(x, t)$  de solveurs de Riemann autour des interfaces  $x_{i-\frac{1}{2}}$ ,  $x_{i+\frac{1}{2}}$  où les zones d'influences de chaque solveur sont représentées en gris.

Comme un solveur de Riemann est une approximation de  $w_{\mathcal{R}}$ , il doit vérifier certaines conditions qui sont détaillées dans la définition qui suit.

**Définition 1.4.3.1** (Solveur de Riemann pour les systèmes hyperboliques homogènes). *Pour  $(i, n)$  dans  $\mathbb{Z} \times \mathbb{N}$ , notons*

$$\xi_{i+\frac{1}{2}}^n : (x, t) \mapsto \frac{x - x_{i+\frac{1}{2}}}{t^n + t}, \quad \forall (x, t) \in C_{i+\frac{1}{2}} \times [0, \Delta t].$$

Soit  $\tilde{w} : \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ . La fonction  $\tilde{w}$  est appelée solveur de Riemann si les conditions suivantes sont satisfaites.

- i) Il existe  $\Delta t > 0$  tel qu'une condition CFL de non interaction est satisfaite.
- ii) La fonction  $\tilde{w}$  est localement intégrable selon sa première variable :

$$\xi \mapsto \tilde{w}(\xi, w, v) \in L_{\text{Loc}}^1(\mathbb{R}), \quad \forall (w, v) \in (\mathbb{R}^m)^2.$$

- iii) La fonction  $\tilde{w}$  est consistante ce qui s'écrit

$$\tilde{w}(\xi, w, w) = w, \quad \forall (\xi, w) \in \mathbb{R} \times \mathbb{R}^m. \quad (1.43)$$

- iv) Lorsque qu'une condition CFL de non interaction est satisfaite, la fonction  $\tilde{w}$  vérifie une relation de consistance intégrale qui s'écrit

$$\frac{1}{\Delta x} \int_{C_{i+\frac{1}{2}}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n) dx = \frac{w_i^n + w_{i+1}^n}{2} - \frac{\Delta t}{\Delta x} (f(w_{i+1}^n) - f(w_i^n)). \quad (1.44)$$

Avec la donnée d'un solveur de Riemann satisfaisant la définition ci-dessus, on dispose d'une approximation de  $w(\cdot, t^{n+1})$ . En effet, d'après (1.36), la relation de consistance intégrale (1.44) impose l'égalité entre la valeur moyenne du solveur de Riemann  $\tilde{w}$  et la valeur moyenne de la solution exacte  $w_{\mathcal{R}}$ . Un schéma de type Godunov se construit ensuite en projetant chacun des solveurs sur l'espace des fonctions constantes par morceaux. Comme une condition CFL de non interaction est satisfaite, la quantité  $w_i^{n+1}$  est uniquement influencée par les deux solveurs de Riemann situés aux frontières de la cellule  $C_i$ . Par conséquent pour tout  $i$  dans  $\mathbb{Z}$ , un schéma de type Godunov se lit

$$w_i^{n+1} = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_i} \tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), w_{i-1}^n, w_i^n) dx + \frac{1}{\Delta x} \int_{x_i}^{x_{i+\frac{1}{2}}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n) dx. \quad (1.45)$$

Il est possible de récrire un schéma de type Godunov sous la forme conservative de la Définition 1.4.1.1. Cette réécriture est proposée dans le lemme ci-dessous.

**Lemme 1.4.3.1** (Forme conservative des schémas de type Godunov). *Considérons un schéma de type Godunov (1.45) construit à partir d'un solveur de Riemann  $\tilde{w}$  satisfaisant la Définition 1.4.3.1. Soit  $\mathcal{F} : (\mathbb{R}^m)^2 \rightarrow \mathbb{R}^m$  telle que*

$$\begin{aligned} \mathcal{F}(w_i^n, w_{i+1}^n) &= \frac{f(w_{i+1}^n) + f(w_i^n)}{2} - \frac{\Delta x}{4\Delta t} (w_{i+1}^n - w_i^n) + \frac{1}{2\Delta t} \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n) dx \\ &\quad - \frac{1}{2\Delta t} \int_{x_i}^{x_{i+\frac{1}{2}}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n) dx, \end{aligned} \quad (1.46)$$

Les assertions suivantes sont vérifiées.

i)  $\mathcal{F}$  est un flux numérique consistant.

ii) Le schéma de type Godunov (1.45) se récrit de manière équivalente sous la forme conservative suivante :

$$w_i^{n+1} = w_i^n - \frac{\Delta t}{\Delta x} (\mathcal{F}(w_i^n, w_{i+1}^n) - \mathcal{F}(w_{i-1}^n, w_i^n)), \quad \forall i \in \mathbb{Z}. \quad (1.47)$$

*Démonstration.* La fonction  $\mathcal{F}$  est visiblement continue et sa consistance se déduit immédiatement de l'égalité (1.43). Par conséquent  $\mathcal{F}$  est un flux numérique. Puisqu'une condition CFL de non interaction est satisfaite, on peut scinder les intégrales figurant dans (1.45) afin de récrire un schéma de type Godunov sous la forme

$$\begin{aligned} w_i^{n+1} &= \frac{1}{2\Delta x} \int_{x_i}^{x_{i+\frac{1}{2}}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n) dx + \frac{1}{2\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_i} \tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), w_{i-1}^n, w_i^n) dx \\ &\quad + \frac{1}{2\Delta x} \int_{C_{i-\frac{1}{2}}} \tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), w_{i-1}^n, w_i^n) dx + \frac{1}{2\Delta x} \int_{C_{i+\frac{1}{2}}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n) dx \\ &\quad - \frac{1}{2\Delta x} \int_{x_{i-1}}^{x_{i-\frac{1}{2}}} \tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), w_{i-1}^n, w_i^n) dx - \frac{1}{2\Delta x} \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n) dx. \end{aligned}$$

En insérant la relation de consistance intégrale (1.44) dans l'égalité ci-dessus, il suit

$$\begin{aligned} w_i^{n+1} &= \frac{1}{2\Delta x} \int_{x_i}^{x_{i+\frac{1}{2}}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n) dx + \frac{1}{2\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_i} \tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), w_{i-1}^n, w_i^n) dx \\ &\quad + \frac{1}{2} \left( \frac{w_{i-1}^n + w_i^n}{2} - \frac{\Delta t}{\Delta x} (f(w_i^n) - f(w_{i-1}^n)) \right) + \frac{1}{2} \left( \frac{w_i^n + w_{i+1}^n}{2} - \frac{\Delta t}{\Delta x} (f(w_{i+1}^n) - f(w_i^n)) \right) \\ &\quad - \frac{1}{2\Delta x} \int_{x_{i-1}}^{x_{i-\frac{1}{2}}} \tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), w_{i-1}^n, w_i^n) dx - \frac{1}{2\Delta x} \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n) dx. \end{aligned}$$

En utilisant l'expression de  $\mathcal{F}$  donnée dans (1.46) dans l'expression ci-dessus, on en déduit la forme conservative (1.47). Comme l'implication réciproque se montre en remontant les calculs précédents, ceci termine également la preuve.  $\square$

D'après la définition (1.45), un schéma de type Godunov est entièrement défini par la donnée du solveur de Riemann qui lui est associé. On peut aussi remarquer que le schéma de Godunov exposé dans la section 1.4.2 est un cas particulier des schémas de type Godunov obtenu en posant  $\tilde{w} = w_{\mathcal{R}}$ . Le solveur de Harten, Lax et van Leer [118] est un autre exemple usuel de solveur de Riemann approché.

Introduit dans [118], le solveur de Harten Lax et van Leer (HLL) s'inscrit dans la Définition 1.4.3.1. Comme le montre cette définition, un solveur de Riemann répond à des critères qui sont tous localisés autour d'une interface  $x_{i+\frac{1}{2}}$  ayant à sa gauche un état  $w_i^n$  et à sa droite un état  $w_{i+1}^n$ . Ainsi, pour alléger les écritures et sans perdre de généralité au propos, on peut donc poser les notations

$$t^n = 0, \quad x_{i+\frac{1}{2}} = 0, \quad w_i^n = w_L, \quad w_{i+1}^n = w_R. \quad (1.48)$$

Avec ces notations, détaillons maintenant le solveur de Riemann HLL. Considérons  $\{\mu_k(w)\}_{k \in \{1, \dots, m\}}$  l'ensemble des  $m$  valeurs propres de la matrice  $\nabla f(w)$  puis deux réels différents  $\lambda_L, \lambda_R$  de part et d'autre de l'interface représentées sur la Figure 1.3 et tels que

$$\lambda_L < \min_{\substack{\alpha \in \{L, R\}, \\ k \in \{1, \dots, m\}}} \mu_k(w_\alpha) \leq \max_{\substack{\alpha \in \{L, R\}, \\ k \in \{1, \dots, m\}}} \mu_k(w_\alpha) < \lambda_R. \quad (1.49)$$

À gauche (*resp.* à droite) du cône illustré par la Figure 1.3 et défini par  $(\lambda_L, \lambda_R)$ , le solveur de Riemann HLL vaut  $w_L$  (*resp.*  $w_R$ ). Dans le cône  $(\lambda_L, \lambda_R)$ , le solveur est constitué d'un état constant noté  $w^{\text{HLL}}$ .

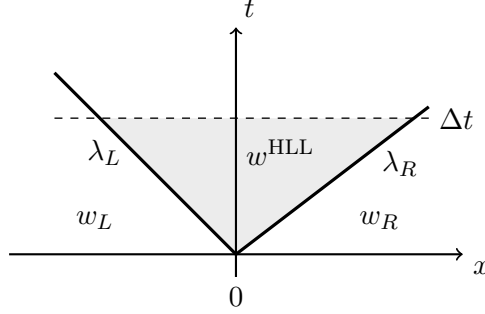


FIGURE 1.3 – Solveur HLL standard [118] autour d'une interface localisée  $x = 0$  et représenté dans le plan  $(x, t)$ .

Ainsi, le solveur de Riemann HLL, noté  $\tilde{w}^{\text{HLL}}(\cdot, w_L, w_R)$ , est défini par

$$\tilde{w}^{\text{HLL}}(x/\Delta t, w_L, w_R) = \begin{cases} w_L & \text{si } \frac{x}{\Delta t} \leq \lambda_L, \\ w^{\text{HLL}} & \text{si } \lambda_L < \frac{x}{\Delta t} \leq \lambda_R, \\ w_R & \text{si } \lambda_R < \frac{x}{\Delta t}. \end{cases} \quad (1.50)$$

En utilisant la définition précédente et si on impose la condition CFL de non interaction suivante :

$$0 < \max_{\alpha \in \{L, R\}} |\lambda_\alpha| \frac{\Delta t}{\Delta x} \leq \frac{1}{2}, \quad (1.51)$$

alors, on peut scinder l'intégrale du solveur (1.50) sur le domaine  $[-\Delta x/2, \Delta x/2]$  comme suit :

$$\begin{aligned} \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \tilde{w}^{\text{HLL}}(x/\Delta t, w_L, w_R) dx &= \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\lambda_L \Delta t} w_L dx + \frac{1}{\Delta x} \int_{\lambda_L \Delta t}^{\lambda_R \Delta t} w^{\text{HLL}} dx + \frac{1}{\Delta x} \int_{\lambda_R \Delta t}^{\frac{\Delta x}{2}} w_R dx, \\ &= \frac{w_L + w_R}{2} + \frac{\Delta t}{\Delta x} (\lambda_R - \lambda_L) w^{\text{HLL}} + \frac{\Delta t}{\Delta x} (\lambda_L w_L - \lambda_R w_R). \end{aligned} \quad (1.52)$$

La quantité  $w^{\text{HLL}}$  de l'équation ci-dessus est inconnue mais se détermine avec la relation de consistance intégrale (1.44). Cette relation appliquée au solveur (1.50)-(1.52) produit l'unique expression de  $w^{\text{HLL}}$  qui est donnée par

$$w^{\text{HLL}} = \frac{\lambda_R w_R - \lambda_L w_L}{\lambda_R - \lambda_L} - \frac{f(w_R) - f(w_L)}{\lambda_R - \lambda_L}. \quad (1.53)$$

Avec la définition ci-dessus, il est facile de vérifier que le solveur HLL (1.50) est consistant au sens de (1.43) et qu'il est intégrable selon sa première variable.

Un schéma de type Godunov (1.45) construit à partir du solveur HLL est nommé par abus de langage *schéma HLL* [118]. Ce schéma satisfait une inégalité d'entropie discrète locale démontrée dans le lemme suivant.

**Lemme 1.4.3.2** (Stabilité entropique locale du schéma HLL). *Soit un système hyperbolique homogène de  $m$  lois de conservation en une dimension d'espace de la forme*

$$\begin{cases} \partial_t w + \partial_x f(w) = 0, & t > 0, \quad x \in \mathbb{R}, \\ w(x, t = 0) = w_0(x). \end{cases} \quad (1.54)$$

*Le système d'équations ci-dessus est muni d'un couple entropie convexe, flux d'entropie  $(\eta, G)$  et ses solutions vérifient une inégalité d'entropie de la forme*

$$\partial_t \eta(w) + \partial_x G(w) \leq 0. \quad (1.55)$$

*Si les solutions de (1.54) sont approchées par un schéma de type Godunov (1.45) associé au solveur HLL (1.50) alors le schéma numérique satisfait une inégalité d'entropie discrète locale de la forme*

$$\frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} + \frac{\mathcal{G}_{i+\frac{1}{2}} - \mathcal{G}_{i-\frac{1}{2}}}{\Delta x} \leq 0, \quad \forall i \in \mathbb{Z}.$$

*Démonstration.* Pour simplifier les écritures, les notations (1.48) sont adoptées. D'après [118], si un solveur de Riemann satisfait l'inégalité suivante :

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \eta(\tilde{w}^{\text{HLL}}(x/\Delta t, w_L, w_R)) \, dx \leq \frac{\eta(w_L) + \eta(w_R)}{2} - \frac{\Delta t}{\Delta x} (G(w_R) - G(w_L)),$$

alors il existe un flux numérique d'entropie  $\mathcal{G}$  tel que le schéma de type Godunov associé au solveur vérifie une inégalité d'entropie discrète. Par conséquent, l'objectif est de montrer que le solveur HLL satisfait l'inégalité ci-dessus. Un calcul analogue à (1.52) donne

$$\begin{aligned} \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \eta(\tilde{w}^{\text{HLL}}(x/\Delta t, w_L, w_R)) \, dx &= \frac{\eta(w_L) + \eta(w_R)}{2} + \frac{\Delta t}{\Delta x} (\lambda_R - \lambda_L) \eta(w^{\text{HLL}}) \\ &+ \frac{\Delta t}{\Delta x} (\lambda_L \eta(w_L) - \lambda_R \eta(w_R)). \end{aligned}$$

En considérant les deux résultats précédents, on doit donc montrer l'inégalité large  $\eta(w^{\text{HLL}}) \leq \eta^{\text{HLL}}$  avec

$$\eta^{\text{HLL}} = \frac{\lambda_R \eta(w_R) - \lambda_L \eta(w_L)}{\lambda_R - \lambda_L} - \frac{G(w_R) - G(w_L)}{\lambda_R - \lambda_L}. \quad (1.56)$$

L'intégration de l'équation (1.54) sur le domaine  $[0, \Delta t] \times [-\Delta x/2, \Delta x/2]$ , jointe à la donnée initiale du problème de Riemann donne

$$\begin{aligned} 0 &= \int_0^{\Delta t} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} (\partial_t w + \partial_x f(w)) \, dx dt, \\ &= \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} (w(x, \Delta t) - w_0(x)) \, dx + \int_0^{\Delta t} (f(w(\Delta x/2, t)) - f(w(-\Delta x/2, t))) \, dt, \\ &= \int_{-\frac{\Delta x}{2}}^{\lambda_L \Delta t} w_L \, dx + \int_{\lambda_L \Delta t}^{\lambda_R \Delta t} w(x, \Delta t) \, dx + \int_{\lambda_R \Delta t}^{\frac{\Delta x}{2}} w_R \, dx - \left( \int_{-\frac{\Delta x}{2}}^0 w_L \, dx + \int_0^{\frac{\Delta x}{2}} w_R \, dx \right) \\ &\quad + \Delta t (f(w_R) - f(w_L)), \\ &= \int_{\lambda_L \Delta t}^{\lambda_R \Delta t} w(x, \Delta t) \, dx + \lambda_L \Delta t w_L - \lambda_R \Delta t w_R + \Delta t (f(w_R) - f(w_L)). \end{aligned} \quad (1.57)$$

En introduisant le résultat précédent dans la définition de  $w^{\text{HLL}}$  donnée par (1.53) on obtient

$$w^{\text{HLL}} = \frac{1}{(\lambda_R - \lambda_L) \Delta t} \int_{\lambda_L \Delta t}^{\lambda_R \Delta t} w(x, \Delta t) \, dx, \quad (1.58)$$

où pour rappel  $\lambda_L < \lambda_R$ . Comme  $\eta$  est convexe, l'inégalité de Jensen décrite dans [173] permet alors de déduire

$$\eta(w^{\text{HLL}}) \leq \frac{1}{(\lambda_R - \lambda_L)\Delta t} \int_{\lambda_L \Delta t}^{\lambda_R \Delta t} \eta(w(x, \Delta t)) dx. \quad (1.59)$$

Or,  $w(x, \Delta t)$  est la solution exacte de (1.54) et vérifie l'inégalité d'entropie (1.55). En intégrant cette inégalité sur  $[0, \Delta t] \times [-\Delta x/2, \Delta x/2]$ , un calcul analogue à (1.57) produit l'inégalité suivante :

$$0 \geq \int_{\lambda_L \Delta t}^{\lambda_R \Delta t} \eta(w(x, \Delta t)) dx + \lambda_L \Delta t \eta(w_L) - \lambda_R \Delta t \eta(w_R) + \Delta t (G(w_R) - G(w_L)).$$

Finalement, avec cette dernière inégalité associée à (1.58)-(1.59), on peut déduire

$$\begin{aligned} \eta(w^{\text{HLL}}) &= \eta\left(\frac{1}{(\lambda_R - \lambda_L)\Delta t} \int_{\lambda_L \Delta t}^{\lambda_R \Delta t} w(x, \Delta t) dx\right) \leq \frac{1}{(\lambda_R - \lambda_L)\Delta t} \int_{\lambda_L \Delta t}^{\lambda_R \Delta t} \eta(w(x, \Delta t)) dx, \\ &\leq \frac{\lambda_R \eta(w_R) - \lambda_L \eta(w_L)}{\lambda_R - \lambda_L} - \frac{G(w_R) - G(w_L)}{\lambda_R - \lambda_L}, \end{aligned}$$

ce qui termine la preuve.  $\square$

Les schémas de type Godunov donnés dans cette section sont seulement valables pour les systèmes d'équations homogènes de la forme (1.54). La section suivante expose la construction des schémas de type Godunov pour le cas des systèmes hyperboliques avec terme source.

#### 1.4.4 Schéma de *type* Godunov pour des systèmes d'équations avec terme source

Soit  $z : \mathbb{R} \rightarrow \mathbb{R}$  une fonction régulière et indépendante du temps. Considérons tout d'abord la suite de réels  $(z_i)_{i \in \mathbb{Z}}$  telle que

$$z_i = \frac{1}{\Delta x} \int_{C_i} z(x) dx, \quad \forall i \in \mathbb{Z}.$$

Pour un système hyperbolique donné, considérons ensuite la séquence de problèmes de Riemann posée sur les intervalles  $(C_{i+\frac{1}{2}})_{i \in \mathbb{Z}}$  et définie par

$$\begin{cases} \partial_t w + \partial_x f(w) = S(w) \partial_x z, & t \in ]t^n, t^n + \Delta t], \quad x \in C_{i+\frac{1}{2}}, \\ w(x, t = t^n) = \begin{cases} w_i^n & \text{si } x < x_{i+\frac{1}{2}}, \\ w_{i+1}^n & \text{sinon,} \end{cases} \end{cases} \quad \forall i \in \mathbb{Z}. \quad (1.60)$$

Considérons enfin la séquence  $(\mathfrak{S}_{i+\frac{1}{2}}^n)_{i \in \mathbb{Z}}$  d'éléments de  $\mathbb{R}^m$  telle que

$$\mathfrak{S}_{i+\frac{1}{2}}^n = \frac{1}{\Delta t \Delta x} \int_0^{\Delta t} \int_{C_{i+\frac{1}{2}}} S(w(x, t^n + t)) \partial_x z dx dt, \quad \forall i \in \mathbb{Z}.$$

D'après les sections 1.4.2 et 1.4.3, si une condition CFL de non interaction est satisfaite, alors la séquence de problèmes de Riemann (1.60) peut s'étudier comme une juxtaposition de problèmes de Riemann locaux et distincts. Dans ce cas, pour  $i$  dans  $\mathbb{Z}$  fixé, le Lemme 1.1.2 de la section 1.1, montre qu'une solution de (1.60) vérifie

$$\begin{aligned} \frac{1}{\Delta x} \int_{C_{i+\frac{1}{2}}} w(x, t^n + \Delta t) dx &= \frac{w_i^n + w_{i+1}^n}{2} - \frac{\Delta t}{\Delta x} (f(w_{i+1}^n) - f(w_i^n)) \\ &+ \Delta t \mathfrak{S}_{i+\frac{1}{2}}^n + \mathcal{O}(\Delta x) + \mathcal{O}(\Delta t). \end{aligned} \quad (1.61)$$

Si la quantité  $\mathfrak{S}_{i+\frac{1}{2}}^n$  est connue, alors aux termes d'erreur  $\mathcal{O}(\Delta x)$  et  $\mathcal{O}(\Delta t)$  près, l'équation ci-dessus s'identifie à une relation de consistance intégrale. Par conséquent, il est d'abord nécessaire de définir une approximation de  $\mathfrak{S}_{i+\frac{1}{2}}^n$ .

**Définition 1.4.4.1** (Approximation consistante de  $(\mathfrak{S}_{i+\frac{1}{2}}^n)_{i \in \mathbb{Z}}$ ). Notons  $\hat{w} = (w, z)^\top$  dans  $\mathbb{R}^{m+1}$ . Soient  $l$  un entier,  $\hat{\Sigma} : (\mathbb{R}^{m+1})^{2l+2} \rightarrow \mathbb{R}^m$  et  $(\hat{S}_{i+\frac{1}{2}}^n)_{i \in \mathbb{Z}}$  une suite d'élément de  $\mathbb{R}^m$  telle que

$$\hat{S}_{i+\frac{1}{2}}^n = \hat{\Sigma}(\hat{w}_{i-l}^n, \dots, \hat{w}_{i+l+1}^n) \frac{z_{i+1} - z_i}{\Delta x}, \quad \forall i \in \mathbb{Z}.$$

Pour  $i$  dans  $\mathbb{Z}$  donné, la quantité  $\hat{S}_{i+\frac{1}{2}}^n$  définie ci-dessus est dite consistante avec  $\mathfrak{S}_{i+\frac{1}{2}}^n$  si

$$\hat{\Sigma}(\hat{w}, \dots, \hat{w}) = S(w). \quad (1.62)$$

Avec la définition précédente, il est désormais possible de construire une relation de consistance intégrale approchée en négligeant les termes d'erreur  $\mathcal{O}(\Delta x)$  et  $\mathcal{O}(\Delta t)$  dans (1.61). Cette relation associée à des critères supplémentaires permet de définir un solveur de Riemann pour les problèmes de Riemann de la forme (1.60).

**Définition 1.4.4.2** (Solveur de Riemann pour les systèmes hyperboliques avec terme source). Notons  $\hat{w} = (w, z)^\top$  dans  $\mathbb{R}^{m+1}$ . Pour  $(i, n)$  dans  $\mathbb{Z} \times \mathbb{N}$ , notons également

$$\xi_{i+\frac{1}{2}}^n : (x, t) \mapsto \frac{x - x_{i+\frac{1}{2}}}{t^n + t}, \quad \forall (x, t) \in C_{i+\frac{1}{2}} \times [0, \Delta t].$$

Soit  $(\hat{S}_{i+\frac{1}{2}}^n)_{i \in \mathbb{Z}}$  une suite d'élément de  $\mathbb{R}^m$  satisfaisant la définition de consistance 1.4.4.1. Une fonction  $\tilde{w} : \mathbb{R} \times \mathbb{R}^{m+1} \times \mathbb{R}^{m+1} \rightarrow \mathbb{R}^m$  est appelée solveur de Riemann pour les systèmes hyperboliques avec terme source si les conditions suivantes sont satisfaites.

- i) Il existe  $\Delta t > 0$  tel qu'une condition CFL de non interaction est satisfaite.
- ii) La fonction  $\tilde{w}$  est localement intégrable selon sa première variable :

$$\xi \mapsto \tilde{w}(\xi, \hat{w}, \hat{v}) \in L_{\text{Loc}}^1(\mathbb{R}), \quad \forall (\hat{w}, \hat{v}) \in (\mathbb{R}^{m+1})^2.$$

- iii) La fonction  $\tilde{w}$  est consistante ce qui s'écrit

$$\tilde{w}(\xi, \hat{w}, \hat{w}) = w, \quad \forall (\xi, \hat{w}) \in \mathbb{R} \times \mathbb{R}^{m+1}.$$

- iv) Lorsque qu'une condition CFL de non interaction est satisfaite, la fonction  $\tilde{w}$  vérifie la relation de consistance intégrale suivante :

$$\frac{1}{\Delta x} \int_{C_{i+\frac{1}{2}}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n) dx = \frac{w_i^n + w_{i+1}^n}{2} - \frac{\Delta t}{\Delta x} (f(w_{i+1}^n) - f(w_i^n)) + \Delta t \hat{S}_{i+\frac{1}{2}}^n. \quad (1.63)$$

La littérature [25, 32, 38, 95, 151, 203] consacrée à l'approximation numérique des solutions des systèmes de la forme (1.60) fait apparaître plusieurs définitions de consistance. La définition proposée ci-dessus forme une condition suffisante à la définition de consistance suggérée dans [32, Eq 4.17] et elle dégénère vers la Définition 1.4.3.1 lorsque la séquence  $(z_i)_{i \in \mathbb{Z}}$  est localement constante.

Avec la Définition 1.4.4.2, un schéma de type Godunov pour les systèmes d'équations de la forme (1.60) s'obtient en projetant chacun des solveurs sur l'espace des fonctions constantes par morceaux. Comme une condition CFL de non interaction est satisfaite, la quantité  $w_i^{n+1}$  est seulement influencée par les deux solveurs de Riemann situés aux frontières de la cellule  $C_i$ . Ainsi, pour tout  $i$  dans  $\mathbb{Z}$ , un schéma de type Godunov pour les systèmes hyperboliques avec terme source s'écrit

$$w_i^{n+1} = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_i} \tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), \hat{w}_{i-1}^n, \hat{w}_i^n) dx + \frac{1}{\Delta x} \int_{x_i}^{x_{i+\frac{1}{2}}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n) dx. \quad (1.64)$$

Il est possible de réécrire un schéma de type Godunov défini ci-dessus sous une forme qui fait apparaître un flux numérique. Cette réécriture est détaillée dans le prochain lemme.

**Lemme 1.4.4.1** (Schémas numériques volumes finis pour les systèmes hyperboliques avec terme source). *Considérons un schéma de type Godunov (1.64) construit à partir d'un solveur de Riemann  $\tilde{w}$  satisfaisant la Définition 1.4.4.2. Soit  $\hat{\mathcal{F}} : (\mathbb{R}^{m+1})^2 \rightarrow \mathbb{R}^m$  telle que*

$$\begin{aligned} \hat{\mathcal{F}}(\hat{w}_i^n, \hat{w}_{i+1}^n) &= \frac{f(w_{i+1}^n) + f(w_i^n)}{2} - \frac{\Delta x}{4\Delta t} (w_{i+1}^n - w_i^n) + \frac{1}{2\Delta t} \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n) dx \\ &- \frac{1}{2\Delta t} \int_{x_i}^{x_{i+\frac{1}{2}}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n) dx, \end{aligned}$$

*En considérant la séquence  $(\hat{S}_{i+\frac{1}{2}}^n)_{i \in \mathbb{Z}}$  satisfaisant la Définition 1.4.4.1, le schéma de type Godunov (1.64) se réécrit de manière équivalente sous la forme*

$$w_i^{n+1} = w_i^n - \frac{\Delta t}{\Delta x} (\hat{\mathcal{F}}(\hat{w}_i^n, \hat{w}_{i+1}^n) - \hat{\mathcal{F}}(\hat{w}_{i-1}^n, \hat{w}_i^n)) + \frac{\Delta t}{2} (\hat{S}_{i+\frac{1}{2}}^n + \hat{S}_{i-\frac{1}{2}}^n), \quad \forall i \in \mathbb{Z}. \quad (1.65)$$

*Démonstration.* En considérant la relation de consistance intégrale qui inclue la séquence  $(\hat{S}_{i+\frac{1}{2}}^n)_{i \in \mathbb{Z}}$  et qui est donnée dans (1.63), la preuve se mène avec les mêmes calculs que ceux utilisés dans la preuve du Lemme 1.4.3.1.  $\square$



## Chapitre 2

# Solveurs de Riemann à deux états intermédiaires avec stabilité entropique locale pour des systèmes homogènes

Ce chapitre propose des schémas de type Godunov pour approcher les solutions d'un système hyperbolique homogène de  $m$  lois de conservation qui s'écrit

$$\begin{cases} \partial_t w + \partial_x f(w) &= 0, & t > 0, & x \in \mathbb{R}, \\ w(x, t = 0) &= w_0(x). \end{cases} \quad (2.1)$$

Le système ci-dessus est muni d'une entropie convexe  $\eta$  associée à un flux d'entropie  $G$  qui vérifie  $\nabla G(w)^T = \nabla \eta(w)^T \nabla f(w)$ . En conséquence, les solutions du système (2.1) vérifient une inégalité d'entropie de la forme

$$\partial_t \eta(w) + \partial_x G(w) \leq 0. \quad (2.2)$$

Les schémas de type Godunov proposés sont définis par des solveurs de Riemann constants par morceaux qui possèdent deux états intermédiaires. Par conséquent, ces schémas sont comparables au schéma HLLC [213], aux schémas de relaxation [21, 32] ou encore au schéma de [118, Section b.ii]. En effet, les schémas décrits dans [21, 32, 213] et dans [118, Section b.ii] sont également définis par des solveurs de Riemann à deux états intermédiaires et ils vérifient tous une stabilité entropique discrète locale. En revanche, le schéma [118, Section b.ii] nécessite la détermination de deux constantes pour lesquelles les auteurs ne donnent pas de formule de calcul. Comme le choix de ces constantes n'est pas simple, l'utilisation pratique de ce schéma peut parfois être problématique. La mise en œuvre des schémas [21, 32, 213] est accessible mais la démonstration de l'inégalité d'entropie discrète demeure délicate.

L'idée développée dans ce chapitre est l'introduction d'une condition suffisante de stabilité entropique dès la détermination du solveur de Riemann. Ainsi, par construction, le schéma de type Godunov résultant contient la stabilité entropique recherchée ce qui simplifie considérablement les preuves.

### 2.1 Condition suffisante de stabilité entropique locale

D'après la Section 1.4.3, un schéma de type Godunov approchant les solutions de (2.1) est défini par la donnée d'un solveur de Riemann  $\tilde{w} : \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$  qui satisfait la Définition 1.4.3.1. Avec cette donnée et en considérant la notation

$$\xi_{i+\frac{1}{2}}^n : (x, t) \mapsto \frac{x - x_{i+\frac{1}{2}}}{t^n + t}, \quad \forall (i, n, x, t) \in \mathbb{Z} \times \mathbb{N} \times ]x_i, x_{i+1}[ \times ]0, \Delta t],$$

un schéma numérique de type Godunov s'écrit

$$w_i^{n+1} = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_i} \tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), w_{i-1}^n, w_i^n) dx + \frac{1}{\Delta x} \int_{x_i}^{x_{i+\frac{1}{2}}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n) dx. \quad (2.3)$$

D'après le Lemme 1.4.3.1, le schéma numérique ci-dessus se réécrit sous une forme conservative qui s'écrit

$$w_i^{n+1} = w_i^n - \frac{\Delta t}{\Delta x} (\mathcal{F}(w_i^n, w_{i+1}^n) - \mathcal{F}(w_{i-1}^n, w_i^n)),$$

avec un flux numérique défini par

$$\begin{aligned} \mathcal{F}(w_i^n, w_{i+1}^n) &= \frac{f(w_{i+1}^n) + f(w_i^n)}{2} - \frac{\Delta x}{4\Delta t} (w_{i+1}^n - w_i^n) + \frac{1}{2\Delta t} \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n) dx \\ &- \frac{1}{2\Delta t} \int_{x_i}^{x_{i+\frac{1}{2}}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n) dx. \end{aligned}$$

Un schéma numérique ainsi rappelé définit la séquence mise à jour  $(w_i^{n+1})_{i \in \mathbb{Z}}$  or, comme le montre la Section 1.3.1, cette séquence doit satisfaire des propriétés de stabilité. Par conséquent, la proposition ci-dessous donne une condition suffisante à la stabilité entropique locale.

**Proposition 2.1.1** (Condition suffisante à la stabilité entropique locale [118]). *Soit un schéma de type Godunov (2.3) défini par un solveur de Riemann  $\tilde{w} : \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$  qui satisfait la Définition 1.4.3.1. Si le solveur de Riemann vérifie*

$$\frac{1}{\Delta x} \int_{x_i}^{x_{i+1}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n)) dx \leq \frac{\eta(w_i^n) + \eta(w_{i+1}^n)}{2} - \frac{\Delta t}{\Delta x} (G(w_{i+1}^n) - G(w_i^n)), \quad \forall i \in \mathbb{Z}, \quad (2.4)$$

alors le schéma de type Godunov (2.3) associé au solveur satisfait une inégalité d'entropie discrète et locale de la forme

$$\frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} + \frac{\mathcal{G}_{i+\frac{1}{2}} - \mathcal{G}_{i-\frac{1}{2}}}{\Delta x} \leq 0, \quad \forall i \in \mathbb{Z}, \quad (2.5)$$

avec,

$$\begin{aligned} \mathcal{G}_{i+\frac{1}{2}} &= \frac{G(w_{i+1}^n) + G(w_i^n)}{2} - \frac{\Delta x}{4\Delta t} (\eta(w_{i+1}^n) - \eta(w_i^n)) + \frac{1}{2\Delta t} \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n)) dx \\ &- \frac{1}{2\Delta t} \int_{x_i}^{x_{i+\frac{1}{2}}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n)) dx. \end{aligned} \quad (2.6)$$

*Démonstration.* Comme une condition CFL de non interaction est satisfaite, on peut scinder les intégrales figurant dans (2.3) et comme  $\eta$  est convexe, l'inégalité de Jensen donne

$$\begin{aligned} \eta(w_i^{n+1}) &\leq \frac{1}{2\Delta x} \int_{x_i}^{x_{i+\frac{1}{2}}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n)) dx + \frac{1}{2\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_i} \eta(\tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), w_{i-1}^n, w_i^n)) dx \\ &+ \frac{1}{2\Delta x} \int_{x_{i-1}}^{x_i} \eta(\tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), w_{i-1}^n, w_i^n)) dx + \frac{1}{2\Delta x} \int_{x_i}^{x_{i+1}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n)) dx \\ &- \frac{1}{2\Delta x} \int_{x_{i-1}}^{x_{i-\frac{1}{2}}} \eta(\tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), w_{i-1}^n, w_i^n)) dx - \frac{1}{2\Delta x} \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n)) dx. \end{aligned}$$

En utilisant la condition (2.4) dans l'inégalité précédente, on obtient

$$\eta(w_i^{n+1}) \leq \frac{1}{2\Delta x} \int_{x_i}^{x_{i+\frac{1}{2}}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n)) dx + \frac{1}{2\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_i} \eta(\tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), w_{i-1}^n, w_i^n)) dx$$

$$\begin{aligned}
& + \frac{1}{2} \left( \frac{\eta(w_{i-1}^n) + \eta(w_i^n)}{2} - \frac{\Delta t}{\Delta x} (G(w_i^n) - G(w_{i-1}^n)) \right) \\
& + \frac{1}{2} \left( \frac{\eta(w_i^n) + \eta(w_{i+1}^n)}{2} - \frac{\Delta t}{\Delta x} (G(w_{i+1}^n) - G(w_i^n)) \right) \\
& - \frac{1}{2\Delta x} \int_{x_{i-1}}^{x_{i-\frac{1}{2}}} \eta(\tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), w_{i-1}^n, w_i^n)) dx - \frac{1}{2\Delta x} \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n)) dx.
\end{aligned}$$

En réorganisant les termes ci-dessus à l'aide de la définition (2.6), on obtient l'inégalité (2.5). Comme la consistance de  $\mathcal{G}$  avec  $G$  se déduit immédiatement de la consistance de solveur de Riemann explicitée dans (1.43), il résulte que (2.5) est une inégalité d'entropie discrète ce qui termine la preuve.  $\square$

L'inégalité (2.4) est déjà formulée dans [118, Théorème 3.1] mais elle est souvent considérée comme un résultat à établir. C'est à dire que pour un schéma de type Godunov (2.3) donné, il faut démontrer que le solveur de Riemann associé satisfait (2.4). Si une telle preuve existe pour le cas particulier du solveur HLL (voir le Lemme 1.4.3.2), elle est malheureusement souvent technique à obtenir en général [24].

La nouveauté développée dans ce chapitre réside dans l'utilisation de (2.4) dès la phase de construction du solveur de Riemann. Ainsi, les schémas de type Godunov résultants contiennent directement les éléments suffisants à la stabilité entropique recherchée. Cette démarche lève donc une partie des difficultés liées aux preuves de stabilité et elle est maintenant appliquée à un solveur de Riemann constant par morceaux défini par la relation de consistance intégrale.

## 2.2 Solveur de Riemann à deux états intermédiaires, consistant et entropique

Pour un système hyperbolique homogène (2.1) donné, l'ensemble des  $m$  valeurs propres de la matrice  $\nabla f(w)$  est noté  $\{\mu_k(w)\}_{k \in \{1, \dots, m\}}$  et les notations suivantes sont adoptées :

$$t^n = 0, \quad x_{i+\frac{1}{2}} = 0, \quad w_i^n = w_L, \quad w_{i+1}^n = w_R. \quad (2.7)$$

De part et d'autre de l'interface ci dessus, deux réels différents sont considérés. Il sont notés  $\lambda_L$  et  $\lambda_R$  et ils vérifient

$$\lambda_L < \min_{\substack{\alpha \in \{L, R\}, \\ k \in \{1, \dots, m\}}} \mu_k(w_\alpha) \leq \max_{\substack{\alpha \in \{L, R\}, \\ k \in \{1, \dots, m\}}} \mu_k(w_\alpha) < \lambda_R. \quad (2.8)$$

Avec ces définitions, la condition CFL de non interaction s'écrit

$$0 < \max_{\alpha \in \{L, R\}} |\lambda_\alpha| \frac{\Delta t}{\Delta x} \leq \frac{1}{2}. \quad (2.9)$$

De plus, avec un troisième réel  $v^*$  donné et tel que  $\lambda_L < v^* < \lambda_R$ , il est possible de définir  $\delta_L$  et  $\delta_R$  de la manière suivante :

$$\delta_L = \frac{v^* - \lambda_L}{\lambda_R - \lambda_L} > 0, \quad \delta_R = \frac{\lambda_R - v^*}{\lambda_R - \lambda_L} > 0. \quad (2.10)$$

Enfin, les quantités  $w^{\text{HLL}}$  et  $\eta^{\text{HLL}}$  respectivement dans  $\mathbb{R}^m$  et dans  $\mathbb{R}$  sont définies par

$$w^{\text{HLL}} = \frac{\lambda_R w_R - \lambda_L w_L}{\lambda_R - \lambda_L} - \frac{f(w_R) - f(w_L)}{\lambda_R - \lambda_L}, \quad (2.11a)$$

$$\eta^{\text{HLL}} = \frac{\lambda_R \eta(w_R) - \lambda_L \eta(w_L)}{\lambda_R - \lambda_L} - \frac{G(w_R) - G(w_L)}{\lambda_R - \lambda_L}. \quad (2.11b)$$

D'après la Section 1.4.3, on rappelle que si les inégalités (2.8)-(2.9) sont satisfaites alors

$$\eta(w^{\text{HLL}}) \leq \eta^{\text{HLL}}.$$

Les solveurs de Riemann choisis sont constants par morceaux et possèdent deux états intermédiaires dans  $(\mathbb{R}^m)^2$  respectivement notés  $w_L^*$  et  $w_R^*$ . Ces solveurs de Riemann illustrés par la Figure 2.1, sont notés  $\tilde{w}(\cdot, w_L, w_R) : \mathbb{R} \rightarrow \mathbb{R}^m$  et ils s'écrivent

$$\tilde{w}(x/\Delta t, w_L, w_R) = \begin{cases} w_L & \text{si } \frac{x}{\Delta t} \leq \lambda_L, \\ w_L^* & \text{si } \lambda_L < \frac{x}{\Delta t} \leq v^*, \\ w_R^* & \text{si } v^* < \frac{x}{\Delta t} \leq \lambda_R, \\ w_R & \text{si } \lambda_R < \frac{x}{\Delta t}. \end{cases} \quad (2.12)$$

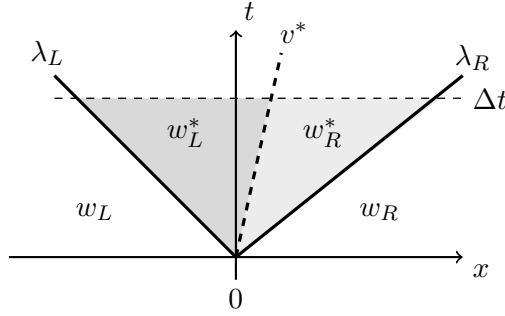


FIGURE 2.1 – Solveur de Riemann approché constitué de deux états intermédiaires constants  $w_L^*$  et  $w_R^*$  représentés dans le plan  $(x, t)$ .

L'objectif est de déterminer les états constants  $(w_L^*, w_R^*)$  de  $(\mathbb{R}^m)^2$  en vertu de la relation de consistance intégrale et de l'inégalité (2.4).

**Lemme 2.2.1.** *Soient  $w_L, w_R$  deux états constants de  $\Omega$ . Soit un schéma de type Godunov (2.3) approchant les solutions de (2.1) et défini par un solveur de Riemann  $\tilde{w}(\cdot, w_L, w_R)$  de la forme (2.12). Soit  $(\eta, G)$  un couple entropie convexe, flux d'entropie associé au système (2.1). Supposons que la condition CFL de non interaction (2.9) est satisfaite et considérons les quantités  $\delta_L, \delta_R, w^{\text{HLL}}, \eta^{\text{HLL}}$  respectivement donnés par (2.10)-(2.11). Les assertions suivantes sont vérifiées.*

i) Si,

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \tilde{w}(x/\Delta t, w_L, w_R) dx = \frac{w_L + w_R}{2} - \frac{\Delta t}{\Delta x} (f(w_R) - f(w_L)), \quad (2.13a)$$

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \eta(\tilde{w}(x/\Delta t, w_L, w_R)) dx \leq \frac{\eta(w_L) + \eta(w_R)}{2} - \frac{\Delta t}{\Delta x} (G(w_R) - G(w_L)), \quad (2.13b)$$

alors les états intermédiaires  $w_L^*$  et  $w_R^*$  du solveur de Riemann (2.12) vérifient  $m$  équations et une inégalité respectivement données par

$$\delta_L w_L^* + \delta_R w_R^* = w^{\text{HLL}}, \quad (2.14a)$$

$$\delta_L \eta(w_L^*) + \delta_R \eta(w_R^*) \leq \eta^{\text{HLL}}. \quad (2.14b)$$

Réciproquement, si les états intermédiaires  $w_L^*$  et  $w_R^*$  du solveur de Riemann (2.12) sont définis par le système d'équations (2.14a) sous la contrainte (2.14b) alors le solveur de Riemann satisfait les conditions (2.13).

ii) Le système d'équations (2.14a) contraint par l'inégalité (2.14b) admet toujours au moins une solution donnée par

$$w_L^* = w_R^* = w^{\text{HLL}}. \quad (2.15)$$

Dans ce cas, le solveur (2.12) dégénère vers le solveur HLL standard rappelé en Section 1.4.3.

*Démonstration.* Pour un solveur de la forme (2.12), si la condition CFL (2.9) est satisfaite alors le membre de gauche de l'égalité (2.13a) s'écrit

$$\begin{aligned} & \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \tilde{w}(x/\Delta t, w_L, w_R) dx \\ &= \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\lambda_L \Delta t} w_L dx + \frac{1}{\Delta x} \int_{\lambda_L \Delta t}^{v^* \Delta t} w_L^* dx + \frac{1}{\Delta x} \int_{v^* \Delta t}^{\lambda_R \Delta t} w_R^* dx + \frac{1}{\Delta x} \int_{\lambda_R \Delta t}^{\frac{\Delta x}{2}} w_R dx, \\ &= \left( \frac{\lambda_L \Delta t}{\Delta x} + \frac{1}{2} \right) w_L + \frac{\Delta t}{\Delta x} (v^* - \lambda_L) w_L^* + \frac{\Delta t}{\Delta x} (\lambda_R - v^*) w_R^* + \left( \frac{1}{2} - \frac{\lambda_R \Delta t}{\Delta x} \right) w_R. \end{aligned}$$

Ainsi, l'égalité (2.13a) se réécrit de manière équivalente sous la forme

$$\begin{aligned} \left( \frac{\lambda_L \Delta t}{\Delta x} + \frac{1}{2} \right) w_L + \frac{\Delta t}{\Delta x} (v^* - \lambda_L) w_L^* + \frac{\Delta t}{\Delta x} (\lambda_R - v^*) w_R^* + \left( \frac{1}{2} - \frac{\lambda_R \Delta t}{\Delta x} \right) w_R = \\ \frac{w_L + w_R}{2} - \frac{\Delta t}{\Delta x} (f(w_R) - f(w_L)). \end{aligned}$$

L'introduction des quantités  $\delta_L$ ,  $\delta_R$  et  $w^{\text{HLL}}$  respectivement données par (2.10)-(2.11a) dans l'égalité ci-dessus, permet d'obtenir l'expression équivalente (2.14a).

Concernant l'inégalité (2.14b), la structure du solveur de Riemann (2.12) ainsi que la condition CFL (2.9) permettent de déduire que le membre de gauche de l'inégalité (2.13b) s'écrit

$$\begin{aligned} & \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \eta(\tilde{w}(x/\Delta t, w_L, w_R)) dx \\ &= \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\lambda_L \Delta t} \eta(w_L) dx + \frac{1}{\Delta x} \int_{\lambda_L \Delta t}^{v^* \Delta t} \eta(w_L^*) dx + \frac{1}{\Delta x} \int_{v^* \Delta t}^{\lambda_R \Delta t} \eta(w_R^*) dx + \frac{1}{\Delta x} \int_{\lambda_R \Delta t}^{\frac{\Delta x}{2}} \eta(w_R) dx, \\ &= \left( \frac{\lambda_L \Delta t}{\Delta x} + \frac{1}{2} \right) \eta(w_L) + \frac{\Delta t}{\Delta x} (v^* - \lambda_L) \eta(w_L^*) + \frac{\Delta t}{\Delta x} (\lambda_R - v^*) \eta(w_R^*) + \left( \frac{1}{2} - \frac{\lambda_R \Delta t}{\Delta x} \right) \eta(w_R). \end{aligned}$$

Par conséquent, une forme équivalente de l'inégalité (2.13b) est donnée par

$$\begin{aligned} \left( \frac{\lambda_L \Delta t}{\Delta x} + \frac{1}{2} \right) \eta(w_L) + \frac{\Delta t}{\Delta x} (v^* - \lambda_L) \eta(w_L^*) + \frac{\Delta t}{\Delta x} (\lambda_R - v^*) \eta(w_R^*) + \left( \frac{1}{2} - \frac{\lambda_R \Delta t}{\Delta x} \right) \eta(w_R) \leq \\ \frac{\eta(w_L) + \eta(w_R)}{2} - \frac{\Delta t}{\Delta x} (G(w_R) - G(w_L)). \end{aligned}$$

En introduisant les quantités  $\delta_L$ ,  $\delta_R$  et  $\eta^{\text{HLL}}$  respectivement données par (2.10)-(2.11b) dans l'inégalité ci-dessus, on en déduit l'expression équivalente (2.14b) ce qui termine de montrer la première assertion.

Pour l'assertion *ii*) et comme  $\delta_L + \delta_R = 1$ , si les égalités (2.15) sont satisfaites alors l'équation (2.14a) est trivialement vérifiée. L'inégalité (2.14b) est également satisfaite puisque dans ce cas elle s'identifie à l'inégalité  $\eta(w^{\text{HLL}}) \leq \eta^{\text{HLL}}$ . Comme la dégénérescence vers le solveur HLL est immédiate d'après les deux remarques précédentes, ceci conclut la preuve.  $\square$

D'après le lemme ci-dessus, l'ensemble des éléments de  $(\mathbb{R}^m)^2$  satisfaisants les  $m + 1$  conditions données par (2.14) est non vide. Comme le système d'équations (2.14a) est de plus sous déterminé, ces conditions génériques peuvent donc servir de support à la construction de solveurs de Riemann de la forme (2.12). Ainsi, les conditions (2.14) vont maintenant être appliquées aux cas équations de Saint Venant homogènes et d'Euler.

## 2.3 Application aux équations de Saint Venant homogènes

Les équations de Saint Venant homogènes sont maintenant considérées. Elles sont obtenues en fixant  $z$  constante dans le système (1.10) et par conséquent, le système d'équations s'écrit

$$\partial_t \begin{pmatrix} h \\ hu \end{pmatrix} + \partial_x \begin{pmatrix} hu \\ hu^2 + gh^2/2 \end{pmatrix} = 0. \quad (2.16)$$

L'inconnue  $w$  se note  $(h, hu)^T$  dans  $\mathbb{R}^2$  et elle prend ses valeurs dans l'espace convexe

$$\Omega = \left\{ (h, hu) \in \mathbb{R}^2 \mid h \geq 0, hu \in \mathbb{R} \right\}. \quad (2.17)$$

La fonction de flux  $f : \Omega \rightarrow \mathbb{R}^2$  est donnée par

$$f(w) = (hu, hu^2 + gh^2/2)^T.$$

Lorsque  $w$  est sur le bord de  $\Omega$ ,  $h = 0$  et dans ce cas la convention admise [15, 166] est  $u = 0$ . Les solutions de (2.16) vérifient également une inégalité d'entropie de la forme (2.2) avec le couple entropie convexe, flux d'entropie  $(\eta, G)$  donné par

$$\eta(w) = \frac{(hu)^2}{2h} + \frac{gh^2}{2}, \quad G(w) = \frac{(hu)^3}{2h^2} + gh(hu), \quad \forall w \in \Omega. \quad (2.18)$$

Dans le cas des équations de Saint Venant homogènes ainsi rappelées, la quantité  $w^{\text{HLL}}$  définie dans (2.11a) se lit  $(h^{\text{HLL}}, (hu)^{\text{HLL}})^T$  et la définition suivante est adoptée

$$u^{\text{HLL}} = \begin{cases} \frac{(hu)^{\text{HLL}}}{h^{\text{HLL}}}, & \text{si } h^{\text{HLL}} > 0, \\ 0, & \text{sinon.} \end{cases} \quad (2.19)$$

La définition ci-dessus s'inscrit dans la convention  $u = 0$  si  $h = 0$  puisqu'un calcul direct montre que  $h^{\text{HLL}}$  est nul si et seulement si  $h_L$  et  $h_R$  sont nuls. Ces cas particuliers du bord de  $\Omega$  seront traités mais ils ne sont pas l'objet principal de ce mémoire de thèse. Avec ces notations et si  $h_L$  ou  $h_R$  est positif, il est alors possible de montrer que  $h^{\text{HLL}}$  est également positif

**Lemme 2.3.1** (Preuve de  $h^{\text{HLL}} \geq 0$ ). *Considérons l'espace convexe  $\Omega$  donné dans (2.17). Soient  $w_L$  et  $w_R$  deux éléments constants de  $\Omega$  et considérons  $w^{\text{HLL}}$  défini dans (2.11a). Si  $w_L$  et  $w_R$  sont dans l'intérieur de  $\Omega$  (resp. dans  $\Omega$ ) et si  $\lambda_L, \lambda_R$  satisfont les inégalités (2.8) alors  $h^{\text{HLL}}$  est strictement positif (resp. positif).*

*Démonstration.* D'après sa définition donnée par (2.11a),  $h^{\text{HLL}}$  se réécrit de manière équivalente sous la forme

$$\begin{aligned} h^{\text{HLL}} &= \frac{\lambda_R h_R - \lambda_L h_L}{\lambda_R - \lambda_L} - \frac{h_R u_R - h_L u_L}{\lambda_R - \lambda_L}, \\ &= \frac{h_R}{\lambda_R - \lambda_L} (\lambda_R - u_R) + \frac{h_L}{\lambda_R - \lambda_L} (u_L - \lambda_L). \end{aligned}$$

Comme  $\lambda_L, \lambda_R$  vérifient les inégalités (2.8), on a  $\lambda_L < \min(u_L, u_R) \leq \max(u_L, u_R) < \lambda_R$ . Par conséquent,  $h^{\text{HLL}}$  est une somme de termes strictement positifs (resp. positifs) si  $w_L$  et  $w_R$  sont dans l'intérieur de  $\Omega$  (resp. dans  $\Omega$ ). On en déduit ainsi le résultat annoncé.  $\square$

La preuve du lemme ci-dessus permet également de déduire que le schéma de type Godunov associé au solveur HLL présenté en Section 1.4.3 préserve l'espace  $\Omega$  si  $\lambda_L$  et  $\lambda_R$  répondent aux inégalités (2.8). Dans la suite, ces inégalités seront supposées vérifiées.

L'objectif est de concevoir un solveur de Riemann de la forme (2.12) qui vérifie les équations (2.14a) sous la contrainte (2.14b) dans le cas des équations de Saint Venant homogènes. Ainsi,  $w_L^*$  et  $w_R^*$  se lisent désormais

$$w_\alpha^* = (h_\alpha^*, h_\alpha^* u_\alpha^*)^T, \quad \forall \alpha \in \{L, R\}.$$

Dans le cas des équations de Saint Venant (2.16), le système (2.14a) donne deux équations or le solveur de Riemann est constitué de quatre inconnues :  $h_L^*, h_R^*, u_L^*$  et  $u_R^*$ . Par conséquent, il est d'abord nécessaire de compléter le système (2.14a) par deux équations complémentaires. La première proposée est l'égalité  $u_L^* = u_R^*$  qui impose la continuité de  $u$  dans les états intermédiaires du solveur de Riemann. Cette continuité a déjà été proposée pour le schéma HLLC par exemple [213]. En revanche, la seconde équation complémentaire doit maintenant assurer l'inégalité (2.14b). Comme il s'agit d'une inégalité *large*, le choix de l'égalité de ses deux membres est proposé. Ce choix préserve l'inégalité d'entropie discrète en vertu de l'inégalité de Jensen utilisée dans la preuve de la Proposition 2.1.1. De plus, ce choix permet d'obtenir un système d'équations pour lequel une résolution explicite est possible.

**Lemme 2.3.2** (Solveur de Riemann entropique pour les équations de Saint-Venant homogènes). Soient  $w_L, w_R$  deux états constants de l'intérieur de  $\Omega$  ( $h_L > 0, h_R > 0$ ). Soit un schéma de type Godunov (2.3) approchant les solutions de (2.16) et défini par un solveur de Riemann  $\tilde{w}(\cdot, w_L, w_R) : \mathbb{R} \rightarrow \mathbb{R}^2$  de la forme (2.12). Supposons que la condition CFL (2.9) est satisfaite puis considérons les quantités  $\delta_L > 0, \delta_R > 0, w^{\text{HLL}}, \eta^{\text{HLL}}$  et  $u^{\text{HLL}}$  respectivement définies dans (2.10)-(2.11)-(2.19). Si les états intermédiaires du solveur  $w_L^*$  et  $w_R^*$  sont définis par le système

$$w_L^* \delta_L + w_R^* \delta_R = w^{\text{HLL}}, \quad (2.20a)$$

$$\eta(w_L^*) \delta_L + \eta(w_R^*) \delta_R = \eta^{\text{HLL}}, \quad (2.20b)$$

$$u_L^* = u_R^*, \quad (2.20c)$$

alors deux expressions pour  $w_L^*$  et  $w_R^*$  sont possibles. Elles sont données par

$$\left\{ \begin{array}{l} u_L^* = u_R^* = u^{\text{HLL}}, \\ h_R^* = h^{\text{HLL}} \pm \sqrt{\frac{2\delta_L}{g\delta_R} (\eta^{\text{HLL}} - \eta(w^{\text{HLL}}))}, \\ h_L^* = h^{\text{HLL}} \mp \sqrt{\frac{2\delta_R}{g\delta_L} (\eta^{\text{HLL}} - \eta(w^{\text{HLL}}))}. \end{array} \right. \quad (2.21)$$

Un schéma de type Godunov (2.3) associé à un des deux solveurs définis ci-dessus

i) est consistant avec les équations de Saint Venant homogènes (2.16),

ii) vérifie une inégalité d'entropie discrète locale (2.5) pour le couple  $(\eta, G)$  défini dans (2.18) et dans laquelle le flux numérique d'entropie  $\mathcal{G}$  est donné par (2.6).

Dans les définitions (2.21), les symboles  $\pm$  et  $\mp$  sont liés. Si  $\pm$  est positif (*resp.* négatif) alors  $\mp$  est négatif (*resp.* positif). Sauf mention contraire, cette convention de notations sera conservée dans la suite de ce manuscrit. La preuve du Lemme 2.3.2 est donnée ci-dessous.

*Démonstration du Lemme 2.3.2.* Il est d'abord nécessaire de montrer que les expressions (2.21) sont les solutions du système (2.20). En utilisant la définition de  $u^{\text{HLL}}$  donnée dans (2.19), la relation de consistance intégrale (2.20a) vérifiée par les états  $w_L^*$  et  $w_R^*$  se lit composante par composante de la manière suivante :

$$\delta_L h_L^* + \delta_R h_R^* = h^{\text{HLL}}, \quad (2.22a)$$

$$\delta_L h_L^* u_L^* + \delta_R h_R^* u_R^* = h^{\text{HLL}} u^{\text{HLL}}. \quad (2.22b)$$

Comme  $u_L^*$  et  $u_R^*$  sont égaux, notons  $u^* = u_L^* = u_R^*$ . Dans ce cas, l'équation (2.22b) associée à (2.22a) donne

$$\begin{aligned} h^{\text{HLL}} u^{\text{HLL}} &= \delta_L h_L^* u_L^* + \delta_R h_R^* u_R^*, \\ &= (\delta_L h_L^* + \delta_R h_R^*) u^*, \\ &= h^{\text{HLL}} u^*. \end{aligned} \quad (2.23)$$

Puisque  $w_L$  et  $w_R$  sont dans l'intérieur de  $\Omega$  et que les inégalités (2.8) sont satisfaites, l'inégalité stricte  $h^{\text{HLL}} > 0$  est vérifiée. Ainsi, une multiplication de l'équation (2.23) par  $1/h^{\text{HLL}}$  produit le résultat attendu quant à  $u_L^*$  et  $u_R^*$ . Il faut maintenant montrer les expressions de  $h_L^*$  et de  $h_R^*$ . En appliquant la fonction  $\eta$  à l'équation (2.20a) et en utilisant l'expression de  $u_L^*$  et de  $u_R^*$ , il vient

$$\begin{aligned} \eta(w^{\text{HLL}}) &= \eta(\delta_L w_L^* + \delta_R w_R^*) = \frac{(\delta_L h_L^* u_L^* + \delta_R h_R^* u_R^*)^2}{2(\delta_L h_L^* + \delta_R h_R^*)} + \frac{g}{2} (\delta_L h_L^* + \delta_R h_R^*)^2, \\ &= \frac{(u^{\text{HLL}})^2}{2} (\delta_L h_L^* + \delta_R h_R^*) + \frac{g}{2} (\delta_L h_L^* + \delta_R h_R^*)^2. \end{aligned} \quad (2.24)$$

Par ailleurs, en considérant l'équation (2.20b) et comme  $\delta_L + \delta_R = 1$ , il suit

$$\begin{aligned}\eta^{\text{HLL}} &= \delta_L \eta(w_L^*) + \delta_R \eta(w_R^*), \\ &= \frac{(u^{\text{HLL}})^2}{2} (\delta_L h_L^* + \delta_R h_R^*) + \frac{g}{2} ((h_L^*)^2 \delta_L + (h_R^*)^2 \delta_R).\end{aligned}$$

En retranchant membre à membre (2.24) à l'équation ci-dessus et en utilisant une nouvelle fois  $\delta_L + \delta_R = 1$ , on obtient

$$\begin{aligned}\eta^{\text{HLL}} - \eta(w^{\text{HLL}}) &= \frac{g}{2} ((h_L^*)^2 \delta_L + (h_R^*)^2 \delta_R) - \frac{g}{2} (\delta_L h_L^* + \delta_R h_R^*)^2, \\ &= \frac{g}{2} \delta_L (1 - \delta_L) (h_L^*)^2 - g \delta_L \delta_R h_L^* h_R^* + \frac{g}{2} \delta_R (1 - \delta_R) (h_R^*)^2, \\ &= \frac{g}{2} \delta_R \delta_L (h_R^* - h_L^*)^2.\end{aligned}\tag{2.25}$$

Comme les conditions (2.8)-(2.9) sont satisfaites, l'inégalité  $\eta^{\text{HLL}} - \eta(w^{\text{HLL}}) \geq 0$  est assurée par la définition du schéma HLL détaillée dans la section 1.4.3. De plus, comme  $\delta_L$  et  $\delta_R$  sont tous deux strictement positifs, l'équation (2.25) est bien posée. On peut donc diviser par  $g\delta_L\delta_R/2$  puis appliquer la racine carrée de part et d'autre pour trouver

$$h_R^* - h_L^* = \pm \sqrt{2(\eta^{\text{HLL}} - \eta(w^{\text{HLL}})) / g\delta_L\delta_R}.$$

En combinant l'équation ci-dessus à l'équation de consistance (2.22a), on peut alors déduire le résultat attendu concernant  $h_L^*$  et  $h_R^*$  ce qui termine de montrer les expressions (2.21). Il est maintenant nécessaire de montrer les propriétés vérifiées par un schéma de type Godunov associé à un tel solveur de Riemann.

Pour la consistance de l'assertion *i*) et d'après la définition du solveur donnée par les expressions (2.21), on obtient immédiatement  $\tilde{w}(\cdot, w, w) = w$ . Comme l'équation  $\delta_L w_L^* + \delta_R w_R^* = w^{\text{HLL}}$  est de plus satisfaite, le Lemme 2.2.1 assure que le solveur de Riemann satisfait la relation de consistance intégrale. Par conséquent,  $\tilde{w}$  satisfait la Définition 1.4.3.1 et la consistance du schéma annoncée dans *i*) se déduit du Lemme 1.4.3.1.

Enfin, l'assertion *ii*) qui concerne l'inégalité d'entropie discrète locale se déduit par applications successives du Lemme 2.2.1 et de la Proposition 2.1.1, ce qui termine la preuve.  $\square$

Les quantités intermédiaires  $(h_\alpha^*)_{\alpha \in \{L, R\}}$  du solveur données dans l'équation (2.21) sont toujours bien définies grâce à la propriété de stabilité entropique du schéma HLL qui s'écrit  $\eta^{\text{HLL}} - \eta(w^{\text{HLL}}) \geq 0$ . En revanche, les schémas de type Godunov résultants ne préservent pas nécessairement le domaine  $\Omega$ . En conséquence, il est nécessaire de coupler les solveurs du Lemme 2.3.2 à une procédure décrite dans [166] et qui permet de préserver l'espace convexe  $\Omega$ . Une telle procédure est également essentielle à la robustesse du schéma aux *voisinages* du bord de  $\Omega$  caractérisé par  $h_L = 0$  ou  $h_R = 0$ .

**Lemme 2.3.3** (Schéma de type Godunov robuste et entropique pour les équations de Saint-Venant homogènes). *Soient  $w_L, w_R$  deux états constants de  $\Omega$ . Soit un schéma de type Godunov (2.3) approchant les solutions de (2.16) et défini par un solveur de Riemann  $\tilde{w}(\cdot, w_L, w_R) : \mathbb{R} \rightarrow \mathbb{R}^2$  de la forme (2.12). Supposons que la condition CFL (2.9) est satisfaite puis considérons les quantités  $\delta_L > 0, \delta_R > 0, w^{\text{HLL}}, \eta^{\text{HLL}}$  et  $u^{\text{HLL}}$  respectivement définies dans (2.10)-(2.11)-(2.19) et notons*

$$\tilde{h}_R^* = h^{\text{HLL}} \pm \sqrt{\frac{2\delta_L}{g\delta_R} (\eta^{\text{HLL}} - \eta(w^{\text{HLL}}))},\tag{2.26a}$$

$$\tilde{h}_L^* = h^{\text{HLL}} \mp \sqrt{\frac{2\delta_R}{g\delta_L} (\eta^{\text{HLL}} - \eta(w^{\text{HLL}}))}.\tag{2.26b}$$



Si les états intermédiaires  $w_L^*$ ,  $w_R^*$  du solveur de Riemann s'écrivent

$$\begin{aligned} u_L^* = u_R^* &= u^{\text{HLL}}, \\ h_R^* &= \min\left(\max(\widetilde{h}_R^*, 0), \frac{h^{\text{HLL}}}{\delta_R}\right), \\ h_L^* &= \min\left(\max(\widetilde{h}_L^*, 0), \frac{h^{\text{HLL}}}{\delta_L}\right), \end{aligned} \quad (2.27)$$

alors les assertions suivantes sont vérifiées.

- i) Un schéma de type Godunov (2.3) associé à un tel solveur est consistant avec les équations de Saint Venant homogènes (2.16) et il préserve l'ensemble convexe  $\Omega$ .
- ii) Si de plus,  $\widetilde{h}_L^* > 0$  et  $\widetilde{h}_R^* > 0$  alors un schéma de type Godunov (2.3) associé à un tel solveur vérifie une inégalité d'entropie discrète locale (2.5) pour le couple  $(\eta, G)$  défini dans (2.18). Le flux numérique d'entropie  $\mathcal{G}$  de cette inégalité est donné par (2.6).

*Démonstration.* D'abord, si  $\widetilde{h}_L^*$  et  $\widetilde{h}_R^*$  donnés par (2.26) sont tous deux strictement positifs alors les expressions (2.27) coïncident avec les expressions (2.21) et dans ce cas la consistance a été montrée dans le Lemme 2.3.2. Supposons maintenant que la quantité  $\widetilde{h}_L^*$  donnée par (2.26b) est telle que  $\widetilde{h}_L^* \leq 0$ . Dans ce cas, la technique de limitation  $\min(\max(\cdot, \cdot), \cdot)$  impose  $h_L^* = 0$  et  $h_R^* = h^{\text{HLL}}/\delta_R \geq 0$ . Ainsi, l'équation de consistance  $\delta_L h_L^* + \delta_R h_R^* = h^{\text{HLL}}$  est maintenue. Comme des calculs identiques peuvent être conduits dans le cas où  $\widetilde{h}_R^* \leq 0$ , on en déduit la consistance annoncée dans l'assertion i).

Ensuite, pour montrer la préservation de  $\Omega$  dans le cas des équations de Saint-Venant (2.16), il est suffisant de montrer que  $h_L^* \geq 0$  et  $h_R^* \geq 0$  or d'après les expressions (2.27) ces inégalités sont immédiatement satisfaites.

Enfin, si les quantités  $\widetilde{h}_L^*$  et  $\widetilde{h}_R^*$  données par (2.26) sont toutes deux strictement positives alors les expressions (2.27) coïncident avec les expressions (2.21) et dans ce cas l'inégalité d'entropie a été montrée dans le Lemme 2.3.2 ce qui conclut la preuve.  $\square$

La procédure de robustesse décrite dans (2.27) préserve la relation de consistance (2.14a) lorsque  $\widetilde{h}_L^* \leq 0$  ou  $\widetilde{h}_R^* \leq 0$  mais pas nécessairement la condition de stabilité entropique (2.14b). En effet, en utilisant  $\delta_L + \delta_R = 1$  et les expressions (2.27), il est clair que  $\widetilde{h}_L^*$  et  $\widetilde{h}_R^*$  ne peuvent pas être tous deux négatifs puis un calcul direct montre que la quantité  $\eta(w_L^*)\delta_L + \eta(w_R^*)\delta_R$  se lit

$$\eta(w_L^*)\delta_L + \eta(w_R^*)\delta_R = \begin{cases} \eta^{\text{HLL}}, & \text{si } \widetilde{h}_L^* \geq 0 \text{ et } \widetilde{h}_R^* \geq 0, \\ \eta(w^{\text{HLL}}) + \frac{g(h^{\text{HLL}})^2}{2} \frac{\delta_L}{\delta_R}, & \text{si } \widetilde{h}_L^* \leq 0 \text{ et } \widetilde{h}_R^* \geq 0, \\ \eta(w^{\text{HLL}}) + \frac{g(h^{\text{HLL}})^2}{2} \frac{\delta_R}{\delta_L}, & \text{si } \widetilde{h}_L^* \geq 0 \text{ et } \widetilde{h}_R^* \leq 0. \end{cases}$$

Par conséquent, sans choix particulier de  $\delta_L$  et de  $\delta_R$ , la stabilité entropique discrète pourrait être localement perdue aux *voisines* du bord de  $\Omega$ .

La section suivante montre comment les conditions (2.14) peuvent s'articuler au design d'un solveur dédié aux équations d'Euler.

## 2.4 Application aux équations d'Euler

Considérons les équations d'Euler pour un gaz parfait décrites dans la section 1.2.3. Dans ce cas, l'inconnue  $w$  s'écrit  $(\rho, \rho u, \rho E)^T$  dans  $\mathbb{R}^3$  et le système d'équations est

$$\partial_t \begin{pmatrix} \rho \\ \rho u \\ \rho E \end{pmatrix} + \partial_x \begin{pmatrix} \rho u \\ \rho u^2 + p \\ (\rho E + p)u \end{pmatrix} = 0. \quad (2.28)$$

L'inconnue  $w$  prend ses valeurs dans l'espace

$$\Omega = \left\{ (\rho, \rho u, \rho E) \in \mathbb{R}^3 \mid \rho > 0, \rho u \in \mathbb{R}, E - u^2/2 > 0 \right\}. \quad (2.29)$$

Pour  $\gamma$  dans  $]1, 3]$ , la fonction  $p : \Omega \rightarrow \mathbb{R}^+$  est donnée par

$$p : w \mapsto (\gamma - 1) \left( \rho E - \frac{\rho u^2}{2} \right), \quad \forall w \in \Omega, \quad (2.30)$$

et la fonction de flux  $f : \Omega \rightarrow \mathbb{R}^3$  vérifie

$$f(w) = (\rho u, \rho u^2 + p, (\rho E + p)u)^\top.$$

Les solutions du système (2.28) vérifient de plus une inégalité d'entropie de la forme (2.2). Parmi les entropies possibles décrites par exemple dans [22, 88, 138], le couple entropie convexe, flux d'entropie  $(\eta, G)$  retenu est donné par

$$\eta(w) = -\rho \ln(p/\rho^\gamma), \quad G(w) = -(\rho u) \ln(p/\rho^\gamma), \quad \forall w \in \Omega. \quad (2.31)$$

Dans le cas des équations d'Euler, les composantes de  $w^{\text{HLL}}$  défini dans (2.11a) sont notées  $(\rho^{\text{HLL}}, (\rho u)^{\text{HLL}}, (\rho E)^{\text{HLL}})^\top$ . En admettant également dans ce cas que la solution du problème du Riemann demeure dans  $\Omega$  alors les inégalités strictes  $\rho^{\text{HLL}} > 0$  et  $(\rho E)^{\text{HLL}} > 0$  peuvent être montrées.

**Lemme 2.4.1** (Preuve de  $\rho^{\text{HLL}} > 0, (\rho E)^{\text{HLL}} > 0$ ). *Considérons l'espace  $\Omega$  donné par (2.29). Soient  $w_L$  et  $w_R$  deux éléments constants de  $\Omega$  et considérons ensuite  $w^{\text{HLL}}$  défini dans (2.11a). Considérons enfin le problème de Riemann pour les équations d'Euler donné par le système (2.28) associé à la condition initiale*

$$w_0(x) = \begin{cases} w_L & \text{si } x < 0, \\ w_R & \text{sinon.} \end{cases}$$

*Si  $\lambda_L$  et  $\lambda_R$  vérifient les inégalités (2.8) et si pour tout  $x$  dans l'intervalle  $[\lambda_L \Delta t, \lambda_R \Delta t]$ , la solution du problème de Riemann  $(\rho, \rho u, \rho E)^\top(x, \Delta t)$  est dans  $\Omega$  alors  $\rho^{\text{HLL}} > 0$  et  $(\rho E)^{\text{HLL}} > 0$ .*

*Démonstration.* En notant  $w$  la solution du problème de Riemann d'un système hyperbolique homogène quelconque, la preuve du Lemme 1.4.3.2 donne l'expression de  $w^{\text{HLL}}$  suivante :

$$w^{\text{HLL}} = \frac{1}{(\lambda_R - \lambda_L)\Delta t} \int_{\lambda_L \Delta t}^{\lambda_R \Delta t} w(x, \Delta t) dx.$$

Par conséquent, en écrivant la première et la dernière composante de l'égalité ci-dessus pour le cas particulier des équations d'Euler (2.28), il suit

$$\begin{aligned} \rho^{\text{HLL}} &= \frac{1}{(\lambda_R - \lambda_L)\Delta t} \int_{\lambda_L \Delta t}^{\lambda_R \Delta t} \rho(x, \Delta t) dx, \\ (\rho E)^{\text{HLL}} &= \frac{1}{(\lambda_R - \lambda_L)\Delta t} \int_{\lambda_L \Delta t}^{\lambda_R \Delta t} (\rho E)(x, \Delta t) dx. \end{aligned}$$

Comme  $\rho$  et  $\rho E$  sont strictement positifs et comme  $\lambda_L < \lambda_R$ , on en déduit le résultat annoncé.  $\square$

Si les conditions du lemme ci-dessus sont satisfaites alors, il est également possible de déduire que le schéma de type Godunov associé au solveur HLL présenté en Section 1.4.3 préserve l'espace  $\Omega$ . Dans la suite, les conditions énoncées dans le Lemme 2.4.1 sont supposées vérifiées et les quantités suivantes peuvent donc être définies :

$$u^{\text{HLL}} = \frac{(\rho u)^{\text{HLL}}}{\rho^{\text{HLL}}}, \quad E^{\text{HLL}} = \frac{(\rho E)^{\text{HLL}}}{\rho^{\text{HLL}}} > 0. \quad (2.32)$$

L'objectif de cette section est d'adapter les conditions (2.14) à la conception d'un solveur de Riemann de la forme (2.12) dédié aux équations d'Euler (2.28). Par conséquent, les états intermédiaires  $w_L^*$  et  $w_R^*$  du solveur de Riemann (2.12) sont à déterminer et ils s'écrivent dans ce cas

$$w_\alpha^* = (\rho_\alpha^*, \rho_\alpha^* u_\alpha^*, \rho_\alpha^* E_\alpha^*)^\top, \quad \forall \alpha \in \{L, R\}.$$

Comme le système (2.14) est sous déterminé, le premier lemme ci-dessous introduit deux équations de fermeture qui imposent la continuité de  $u$  et de  $p$  dans les états intermédiaires du solveur de Riemann. De telles continuités ont déjà été proposées pour le schéma HLLC par exemple [213].

**Lemme 2.4.2.** *Soient  $w_L, w_R$  deux états constants de  $\Omega$  et considérons  $w \mapsto p(w)$  donnée par (2.30). Soit un schéma de type Godunov (2.3) approchant les solutions de (2.28) et défini par un solveur de Riemann  $\tilde{w}(\cdot, w_L, w_R) : \mathbb{R} \rightarrow \mathbb{R}^3$  de la forme (2.12) Supposons que la condition CFL (2.9) est satisfaite puis considérons les quantités  $\delta_L > 0, \delta_R > 0, w^{\text{HLL}}$  et  $u^{\text{HLL}}$  respectivement définies dans (2.10)-(2.11a)-(2.32). Si les états intermédiaires du solveur notés  $w_L^*$  et  $w_R^*$  sont définis par les équations*

$$w_L^* \delta_L + w_R^* \delta_R = w^{\text{HLL}}, \quad (2.33a)$$

$$u_L^* = u_R^* \quad (2.33b)$$

$$p(w_L^*) = p(w_R^*), \quad (2.33c)$$

alors,

$$u_L^* = u_R^* = u^{\text{HLL}}, \quad (2.34a)$$

$$p(w_L^*) = p(w_R^*) = p(w^{\text{HLL}}). \quad (2.34b)$$

*Démonstration.* Écrivons composante par composante la relation de consistance intégrale (2.33a) vérifiée par les états  $w_L^*, w_R^*$  pour les équations d'Euler. En utilisant les définitions de  $u^{\text{HLL}}$  et de  $E^{\text{HLL}}$  données dans (2.32), on a

$$\delta_L \rho_L^* + \delta_R \rho_R^* = \rho^{\text{HLL}}, \quad (2.35a)$$

$$\delta_L \rho_L^* u_L^* + \delta_R \rho_R^* u_R^* = \rho^{\text{HLL}} u^{\text{HLL}}, \quad (2.35b)$$

$$\delta_L \rho_L^* E_L^* + \delta_R \rho_R^* E_R^* = \rho^{\text{HLL}} E^{\text{HLL}}. \quad (2.35c)$$

Comme  $u_L^*$  et  $u_R^*$  sont égaux, notons  $u^* = u_L^* = u_R^*$ . Dans ce cas, l'équation (2.35b) associée à (2.35a) donnent

$$\begin{aligned} \rho^{\text{HLL}} u^{\text{HLL}} &= \delta_L \rho_L^* u_L^* + \delta_R \rho_R^* u_R^*, \\ &= (\delta_L \rho_L^* + \delta_R \rho_R^*) u^*, \\ &= \rho^{\text{HLL}} u^*. \end{aligned} \quad (2.36)$$

Comme  $w_L$  et  $w_R$  sont dans  $\Omega$  et que les inégalités (2.8) sont satisfaites, on a  $\rho^{\text{HLL}} > 0$ . Ainsi, une multiplication de l'équation (2.36) par  $1/\rho^{\text{HLL}}$  produit le résultat attendu quant à  $u_L^*$  et  $u_R^*$ .

En utilisant ainsi les égalités (2.34a) on déduit que  $p(w_L^*)$  et  $p(w_R^*)$  s'écrivent

$$p(w_\alpha^*) = (\gamma - 1) \left( \rho_\alpha^* E_\alpha^* - \frac{\rho_\alpha^* (u^{\text{HLL}})^2}{2} \right), \quad \forall \alpha \in \{L, R\}. \quad (2.37)$$

Avec le résultat ci-dessus, on peut former la quantité  $\delta_L p(w_L^*) + \delta_R p(w_R^*)$  et en utilisant les équations (2.35a)-(2.35c), on obtient alors

$$\begin{aligned} \delta_L p(w_L^*) + \delta_R p(w_R^*) &= (\gamma - 1) \left( \delta_L \rho_L^* E_L^* + \delta_R \rho_R^* E_R^* - \frac{\delta_L \rho_L^* + \delta_R \rho_R^*}{2} (u^{\text{HLL}})^2 \right), \\ &= (\gamma - 1) \left( \rho^{\text{HLL}} E^{\text{HLL}} - \frac{\rho^{\text{HLL}}}{2} (u^{\text{HLL}})^2 \right), \\ &= p(w^{\text{HLL}}). \end{aligned}$$

Comme  $\delta_L + \delta_R = 1$  et que l'égalité  $p(w_L^*) = p(w_R^*)$  est imposée, on en déduit le résultat.  $\square$

Les équations de fermeture du lemme précédent permettent de déduire  $u_L^*$ ,  $u_R^*$ ,  $p(w_L^*)$  et  $p(w_R^*)$ . De plus, si les quantités  $\rho_L^*$  et  $\rho_R^*$  sont connues, alors le calcul de  $E_L^*$  et de  $E_R^*$  est possible d'après les équations (2.37). Par conséquent, il reste seulement à déterminer  $\rho_L^*$  et  $\rho_R^*$  d'après l'équation sous contrainte

$$\rho_L^* \delta_L + \rho_R^* \delta_R = \rho^{\text{HLL}}, \quad (2.38a)$$

$$\delta_L \eta(w_L^*) + \delta_R \eta(w_R^*) \leq \eta^{\text{HLL}}. \quad (2.38b)$$

Comme les non linéarités de l'entropie  $\eta$  complexifie la résolution de l'équation sous contrainte (2.38), la contrainte (2.38b) va être substituée par une équation quadratique plus simple. Cette nouvelle équation est obtenue à partir d'une majoration de  $\delta_L \eta(w_L^*) + \delta_R \eta(w_R^*)$  présentée ci-dessous.

**Lemme 2.4.3** (Majoration quadratique de  $\delta_L \eta(w_L^*) + \delta_R \eta(w_R^*)$ ). *Soient  $w_L, w_R$  deux états constants de  $\Omega$ . Considérons l'entropie  $w \mapsto \eta(w)$  donnée par (2.31). Soit un solveur de Riemann  $\tilde{w}(\cdot, w_L, w_R) : \mathbb{R} \rightarrow \mathbb{R}^3$  de la forme (2.12) et défini par le système d'équations (2.33). Considérons les quantités  $\delta_L > 0$ ,  $\delta_R > 0$ ,  $w^{\text{HLL}}$  respectivement définies dans (2.10)-(2.11a). Si  $\rho_L^*$  et  $\rho_R^*$  sont strictement positifs, alors*

$$\eta(w_L^*) \delta_L + \eta(w_R^*) \delta_R \leq \frac{\gamma \delta_L}{\rho^{\text{HLL}}} (\rho_L^* - \rho^{\text{HLL}})^2 + \frac{\gamma \delta_R}{\rho^{\text{HLL}}} (\rho_R^* - \rho^{\text{HLL}})^2 + \eta(w^{\text{HLL}}). \quad (2.39)$$

*Démonstration.* Remarquons tout d'abord que si  $\rho_L^* = \rho_R^*$ , alors l'équation (2.35a) impose  $\rho_L^* = \rho_R^* = \rho^{\text{HLL}}$ . Dans ce cas, l'inégalité (2.39) est trivialement vérifiée au sens de l'égalité. Étudions à présent le cas  $\rho_L^* \neq \rho_R^*$ . En utilisant la définition de l'entropie  $\eta(w)$  donnée dans (2.31) ainsi que les équations (2.34b)-(2.35a), il suit

$$\delta_L \eta(w_L^*) + \delta_R \eta(w_R^*) = -\delta_L \rho_L^* \ln \left( \frac{p(w^{\text{HLL}})}{(\rho_L^*)^\gamma} \right) - \delta_R \rho_R^* \ln \left( \frac{p(w^{\text{HLL}})}{(\rho_R^*)^\gamma} \right). \quad (2.40)$$

En considérant l'équation (2.35a) dans la quantité  $\eta(w^{\text{HLL}})$ , on obtient de plus

$$\eta(w^{\text{HLL}}) = -\delta_L \rho_L^* \ln \left( \frac{p(w^{\text{HLL}})}{(\rho^{\text{HLL}})^\gamma} \right) - \delta_R \rho_R^* \ln \left( \frac{p(w^{\text{HLL}})}{(\rho^{\text{HLL}})^\gamma} \right).$$

En soustrayant membres à membres l'équation ci-dessus à l'équation (2.40), on peut éliminer  $p(w^{\text{HLL}})$  et trouver

$$\frac{\delta_L \eta(w_L^*) + \delta_R \eta(w_R^*) - \eta(w^{\text{HLL}})}{\gamma \rho^{\text{HLL}}} = \delta_L \frac{\rho_L^*}{\rho^{\text{HLL}}} \ln \left( \frac{\rho_L^*}{\rho^{\text{HLL}}} \right) + \delta_R \frac{\rho_R^*}{\rho^{\text{HLL}}} \ln \left( \frac{\rho_R^*}{\rho^{\text{HLL}}} \right). \quad (2.41)$$

Quitte à permuter, on peut supposer que  $\rho_L^* < \rho_R^*$ . Dans ce cas, un raisonnement par l'absurde immédiat combiné à l'équation (2.35a), montre que  $\rho_L^* < \rho^{\text{HLL}}$ . Comme on suppose de plus que  $\rho_L^* > 0$  et  $\rho_R^* > 0$ , on en déduit alors l'existence de deux réels  $r_L^*, r_R^*$  tels que

$$\frac{\rho_L^*}{\rho^{\text{HLL}}} = 1 - r_L^*, \quad \frac{\rho_R^*}{\rho^{\text{HLL}}} = 1 + r_R^*, \quad \text{avec, } r_L^* \geq 0, r_R^* \in [0, 1[. \quad (2.42)$$

En reformulant l'équation (2.41) avec  $r_L^*$  et  $r_R^*$ , on trouve

$$\frac{\delta_L \eta(w_L^*) + \delta_R \eta(w_R^*) - \eta(w^{\text{HLL}})}{\gamma \rho^{\text{HLL}}} = \delta_L (1 - r_L^*) \ln(1 - r_L^*) + \delta_R (1 + r_R^*) \ln(1 + r_R^*).$$

En utilisant un développement de Taylor avec reste intégral, on a

$$\ln(1 - r_L^*) = -r_L^* - \int_0^{r_L^*} \frac{(r_L^* - s)}{(1 - s)^2} ds, \quad (1 + r_R^*) \ln(1 + r_R^*) = r_R^* + \frac{(r_R^*)^2}{2} - \frac{1}{2} \int_0^{r_R^*} \left( \frac{r_R^* - s}{1 + s} \right)^2 ds.$$

Puisque  $0 \leq r_L^* < 1$  et  $0 \leq r_R^*$ , les deux restes des développements précédents sont signés et on en déduit

$$\begin{aligned} \delta_L (1 - r_L^*) \ln(1 - r_L^*) + \delta_R (1 + r_R^*) \ln(1 + r_R^*) &\leq -\delta_L r_L^* (1 - r_L^*) + \delta_R \left( r_R^* + \frac{(r_R^*)^2}{2} \right), \\ &\leq \delta_R r_R^* - \delta_L r_L^* + \delta_L (r_L^*)^2 + \delta_R (r_R^*)^2. \end{aligned}$$

Si on reformule l'équation (2.35a) avec les quantités  $r_L^*$  et  $r_R^*$  définies dans (2.42), on trouve  $\delta_L r_L^* = \delta_R r_R^*$ . En utilisant ce résultat dans l'inégalité ci dessus, on obtient

$$\begin{aligned} \frac{\delta_L \eta(w_L^*) + \delta_R \eta(w_R^*) - \eta(w^{\text{HLL}})}{\gamma \rho^{\text{HLL}}} &= \delta_L (1 - r_L^*) \ln(1 - r_L^*) + \delta_R (1 + r_R^*) \ln(1 + r_R^*), \\ &\leq \delta_L (r_L^*)^2 + \delta_R (r_R^*)^2. \end{aligned}$$

En réutilisant la définition (2.42) pour récrire l'inégalité précédente avec  $\rho_L^*$  et  $\rho_R^*$ , on déduit l'inégalité (2.39) ce qui termine la preuve.  $\square$

Le majorant de l'inégalité (2.39) n'est clairement pas le plus optimal possible mais il est retenu en raison de sa simplicité. Imposer ce majorant égal à  $\eta^{\text{HLL}}$  permet d'obtenir une dernière équation qui par construction, préserve l'inégalité  $\delta_L \eta(w_L^*) + \delta_R \eta(w_R^*) \leq \eta^{\text{HLL}}$ . Ainsi le solveur de Riemann (2.12) proposé pour les équations d'Euler (2.28) est défini par les équations

$$w_L^* \delta_L + w_R^* \delta_R = w^{\text{HLL}}, \quad (2.43a)$$

$$u_L^* = u_R^*, \quad (2.43b)$$

$$p(w_L^*) = p(w_R^*), \quad (2.43c)$$

$$\frac{\gamma \delta_L}{\rho^{\text{HLL}}} (\rho_L^* - \rho^{\text{HLL}})^2 + \frac{\gamma \delta_R}{\rho^{\text{HLL}}} (\rho_R^* - \rho^{\text{HLL}})^2 + \eta(w^{\text{HLL}}) = \eta^{\text{HLL}}. \quad (2.43d)$$

Avec le système d'équations ci-dessus et les résultats du Lemme 2.4.2, il est désormais possible de terminer le design du solveur et d'en déduire un schéma de type Godunov entropique.

**Lemme 2.4.4** (Schéma de type Godunov entropique pour les équations d'Euler). *Soient  $w_L, w_R$  deux états constants de  $\Omega$ . Considérons la fonction  $w \mapsto p(w)$  et l'entropie  $w \mapsto \eta(w)$  respectivement données par (2.30)-(2.31). Soit un schéma de type Godunov (2.3) approchant les solutions de (2.28) et défini par un solveur de Riemann  $\tilde{w}(\cdot, w_L, w_R) : \mathbb{R} \rightarrow \mathbb{R}^3$  de la forme (2.12). Supposons que la condition CFL (2.9) est satisfaite puis considérons les quantités  $\delta_L > 0, \delta_R > 0, w^{\text{HLL}}, \eta^{\text{HLL}}$  et  $u^{\text{HLL}}$  respectivement définies dans (2.10)-(2.11)-(2.32). Si les états intermédiaires du solveur notés  $w_L^*$  et  $w_R^*$  s'écrivent*

$$u_L^* = u_R^* = u^{\text{HLL}}, \quad (2.44a)$$

$$p(w_L^*) = p(w_R^*) = p(w^{\text{HLL}}), \quad (2.44b)$$

$$\rho_L^* = \rho^{\text{HLL}} \mp \sqrt{\frac{\rho^{\text{HLL}}}{\gamma} \frac{\delta_R}{\delta_L} (\eta^{\text{HLL}} - \eta(w^{\text{HLL}}))}, \quad (2.44c)$$

$$\rho_R^* = \rho^{\text{HLL}} \pm \sqrt{\frac{\rho^{\text{HLL}}}{\gamma} \frac{\delta_L}{\delta_R} (\eta^{\text{HLL}} - \eta(w^{\text{HLL}}))}, \quad (2.44d)$$

alors les assertions suivantes sont vérifiées.

- i) Un schéma de type Godunov (2.3) associé à un tel solveur est consistant avec les équations d'Euler (2.28).
- ii) Si de plus,  $\rho_L^* > 0$  et  $\rho_R^* > 0$  alors un schéma de type Godunov (2.3) associé à un tel solveur vérifie une inégalité d'entropie discrète locale (2.5) pour le couple  $(\eta, G)$  défini dans (2.31). Le flux numérique d'entropie  $\mathcal{G}$  de cette inégalité est donné par (2.6).
- iii) Si la séquence  $(w_i^n)_{i \in \mathbb{Z}}$  définie un état de contact stationnaire caractérisé par  $(u_i^n)_{i \in \mathbb{Z}}$  nulle et  $(p(w_i^n))_{i \in \mathbb{Z}}$  constante alors un schéma de type Godunov (2.3) associé à un tel solveur préserve exactement les variables  $u$  et  $p$ . Cette propriété s'écrit

$$\text{si } \begin{cases} u_i^n = 0, \\ p(w_i^n) = \text{cste}, \end{cases} \quad \forall i \in \mathbb{Z}, \quad \text{alors } \begin{cases} u_i^{n+1} = 0, \\ p(w_i^{n+1}) = \text{cste}, \end{cases} \quad \forall i \in \mathbb{Z}.$$

*Démonstration.* Considérons un solveur de Riemann de la forme (2.12), défini par le système d'équations (2.43) et montrons que les expressions (2.44) sont les solutions de ce système d'équations. Les résultats (2.44a)-(2.44b) concernant respectivement  $u_L^*$ ,  $u_R^*$  et  $p(w_L^*)$ ,  $p(w_R^*)$  ont été prouvés dans le Lemme 2.4.2. Il faut maintenant montrer les expressions de  $\rho_L^*$  et de  $\rho_R^*$  définis par l'équation (2.43a) et par l'équation quadratique (2.43d). Ainsi,  $\rho_L^*$  et  $\rho_R^*$  vérifient les deux équations suivantes :

$$\delta_L \rho_L^* + \delta_R \rho_R^* = \rho^{\text{HLL}}, \quad (2.45a)$$

$$\frac{\gamma \delta_L}{\rho^{\text{HLL}}} (\rho_L^* - \rho^{\text{HLL}})^2 + \frac{\gamma \delta_R}{\rho^{\text{HLL}}} (\rho_R^* - \rho^{\text{HLL}})^2 + \eta(w^{\text{HLL}}) = \eta^{\text{HLL}}. \quad (2.45b)$$

En introduisant l'équation (2.45a) dans l'équation (2.45b), on trouve

$$\frac{\gamma \delta_R \delta_L}{\rho^{\text{HLL}}} (\rho_R^* - \rho_L^*)^2 = \eta^{\text{HLL}} - \eta(w^{\text{HLL}}). \quad (2.46)$$

Comme les conditions (2.8)-(2.9) sont satisfaites, l'inégalité large  $\eta^{\text{HLL}} - \eta(w^{\text{HLL}}) \geq 0$  est une nouvelle fois assurée par la propriété du schéma HLL détaillée dans le Lemme 1.4.3.2. De plus, comme  $\delta_L$  et  $\delta_R$  sont tous deux strictement positifs, on peut multiplier l'équation (2.46) par  $\rho^{\text{HLL}}/\gamma\delta_L\delta_R$  puis appliquer la racine carrée de part et d'autre pour en déduire

$$\rho_R^* - \rho_L^* = \pm \sqrt{\rho^{\text{HLL}} (\eta^{\text{HLL}} - \eta(w^{\text{HLL}})) / \gamma \delta_L \delta_R}.$$

Puisque  $\delta_L + \delta_R = 1$ , l'équation précédente couplée à (2.45a) permet de déduire le résultat attendu quant à l'expression de  $\rho_L^*$ ,  $\rho_R^*$  ce qui termine de montrer que (2.44) sont les solutions du système (2.43).

Pour la consistance  $i$ ), et d'après la définition du solveur donnée par les expressions (2.44), l'égalité  $\tilde{w}(\cdot, w, w) = w$  est satisfaite. Comme les solveurs sont de plus construits à partir de l'équation (2.14a) équivalente à la relation de consistance intégrale (2.13a), il suit que  $\tilde{w}$  répond à la Définition 1.4.3.1. Par conséquent et d'après le Lemme 1.4.3.1, les schémas de type Godunov associés sont consistants avec les équations d'Euler (2.28).

Concernant l'assertion  $ii$ ), d'après le Lemme 2.4.3, l'équation (2.45b) permet d'assurer l'inégalité  $\delta_L \eta(w_L^*) + \delta_R \eta(w_R^*) \leq \eta^{\text{HLL}}$ . En conséquence et par l'intermédiaire du Lemme 2.2.1, les solveurs satisfont la Proposition 2.1.1, ce qui permet de déduire la stabilité entropique attendue.

Enfin, pour montrer la préservation des contacts stationnaires de l'assertion  $iii$ ), il est suffisant de montrer que si

$$u_L = u_R = 0, \quad (2.47a)$$

$$p(w_L) = p(w_R) = p_0, \quad (2.47b)$$

alors  $u_L^* = u_R^* = 0$  et  $\rho_\alpha E_\alpha = \rho_\alpha^* E_\alpha^*$  pour tout  $\alpha$  dans  $\{L, R\}$ . Tout d'abord, les égalités (2.44a) donnent  $u_L^* = u_R^* = u^{\text{HLL}}$  or d'après la définition de  $u^{\text{HLL}}$  (2.32) il suit

$$\begin{aligned} u^{\text{HLL}} &= \frac{(\rho u)^{\text{HLL}}}{\rho^{\text{HLL}}}, \\ &= \frac{1}{\rho^{\text{HLL}}(\lambda_R - \lambda_L)} \left( \lambda_R \rho_R u_R - \lambda_L \rho_L u_L - \rho_R u_R^2 + \rho_L u_L^2 - (p(w_R) - p(w_L)) \right), \\ &= 0. \end{aligned} \quad (2.48)$$

Par conséquent, l'équation ci-dessus permet de déduire que si les égalités (2.47) sont satisfaites alors  $u_L^* = u_R^* = 0$ . De plus, et d'après l'expression de  $p(w)$  donnée par (2.30), lorsque les conditions (2.47) sont vérifiées, un calcul immédiat montre que

$$\rho_\alpha E_\alpha = \frac{p_0}{\gamma - 1}, \quad \forall \alpha \in \{L, R\}. \quad (2.49)$$

L'utilisation du résultat ci-dessus et des égalités (2.48)-(2.47a) dans la définition de  $w^{\text{HLL}}$  (2.11) donne

$$\begin{aligned}
\frac{p(w^{\text{HLL}})}{\gamma - 1} &= (\rho E)^{\text{HLL}} - \frac{\rho^{\text{HLL}}(u^{\text{HLL}})^2}{2}, \\
&= \frac{\lambda_R \rho_R E_R - \lambda_L \rho_L E_L}{\lambda_R - \lambda_L} - \frac{(\rho_R E_R + p(w_R))u_R - (\rho_L E_L + p(w_L))u_L}{\lambda_R - \lambda_L}, \\
&= \frac{\lambda_R \rho_R E_R - \lambda_L \rho_L E_L}{\lambda_R - \lambda_L}, \\
&= \frac{p_0}{\gamma - 1}.
\end{aligned}$$

Puisque les égalités  $u_L^* = u_R^* = 0$  ont été démontrées, l'association de l'égalité ci-dessus et des résultats (2.44b)-(2.49) permet de déduire que si les conditions (2.47) sont satisfaites alors

$$\rho_\alpha^* E_\alpha^* = \frac{p(w_\alpha^*)}{\gamma - 1} = \frac{p(w^{\text{HLL}})}{\gamma - 1} = \frac{p_0}{\gamma - 1} = \rho_\alpha E_\alpha, \quad \forall \alpha \in \{L, R\}.$$

Ces dernières égalités terminent de montrer l'assertion *iii*) ainsi que la preuve du lemme.  $\square$

Les schémas de type Godunov associés aux solveurs du lemme ci-dessus sont entropiques mais ils reposent sur l'hypothèse  $\rho_L^* > 0$  et  $\rho_R^* > 0$  qui pourrait de pas être vérifiée. Dans ce cas, la stabilité entropique des schémas proposés pourrait ne plus être garantie. Il est cependant possible de corriger ce défaut de robustesse en s'appuyant sur le schéma HLL. En notant  $\widetilde{w}_R^*$  et  $\widetilde{w}_L^*$  deux états définis par les expressions (2.44), une version robuste des solveurs du Lemme 2.4.4 est donnée par

$$(w_L^*, w_R^*) = \begin{cases} (\widetilde{w}_L^*, \widetilde{w}_R^*), & \text{si } (\widetilde{w}_L^*, \widetilde{w}_R^*) \in \Omega^2, \\ (w^{\text{HLL}}, w^{\text{HLL}}), & \text{sinon.} \end{cases}$$

D'après le Lemme 2.2.1, la procédure ci-dessus préserve la consistance ainsi que la stabilité entropique et, d'après le Lemme 2.4.1, elle assure de plus la préservation du domaine  $\Omega$ .

La section suivante présente des expériences numériques qui illustrent les schémas numériques développés.

## 2.5 Résultats de simulations numériques

Les solveurs de Riemann proposés dans ce chapitre sont constants par morceaux et possèdent deux états intermédiaires, Par conséquent, les schémas de type Godunov associés peuvent être comparés numériquement aux schémas existants qui possèdent une structure de solveur similaire.

Dans le cas des équations de Saint Venant, les résultats sont comparés à un schéma de relaxation de Suliciu [32] puis dans le cas des équations d'Euler, le schéma HLLC [213] est utilisé. Dans les deux cas, une comparaison avec le schéma de [118, Section b.ii] est également réalisée. Ce schéma nécessite la définition de deux constantes pour lesquelles les auteurs ne proposent pas de formule de calcul explicite. Ces deux constantes sont sélectionnées empiriquement et leur valeur est fixée à  $10^{-7}$ . Ce schéma est de plus symétrisé en imposant l'égalité  $\lambda_R = -\lambda_L$ .

Les conditions CFL utilisées pour ces schémas comparatifs sont les mêmes que celles employées pour les schémas numériques proposés. Les premiers cas tests concernent les équations de Saint Venant.

### 2.5.1 Les équations de Saint Venant homogènes

Dans cette section, on considère les deux schémas décrits dans le Lemme 2.3.2 ainsi que leurs versions robustes données par le Lemme 2.3.3. La constante de gravitation  $g$  vaut 9,81. Pour toutes les

interfaces du maillage ayant de part et d'autre un état constant gauche  $w_L$  et un état constant droit  $w_R$ , on considère  $u^{\text{HLL}}$  donné par (2.19) et les trois réels  $\lambda_R, \lambda_L, v^*$  sont tels que

$$\lambda_L = \min_{\alpha \in \{L,R\}} u_\alpha \pm \sqrt{gh_\alpha}, \quad \lambda_R = \max_{\alpha \in \{L,R\}} u_\alpha \pm \sqrt{gh_\alpha}, \quad v^* = u^{\text{HLL}}, \quad \forall (L,R). \quad (2.50)$$

Le pas de temps  $\Delta t$  est choisi selon la condition CFL (2.9) et avec un nombre CFL égale à 0,5.

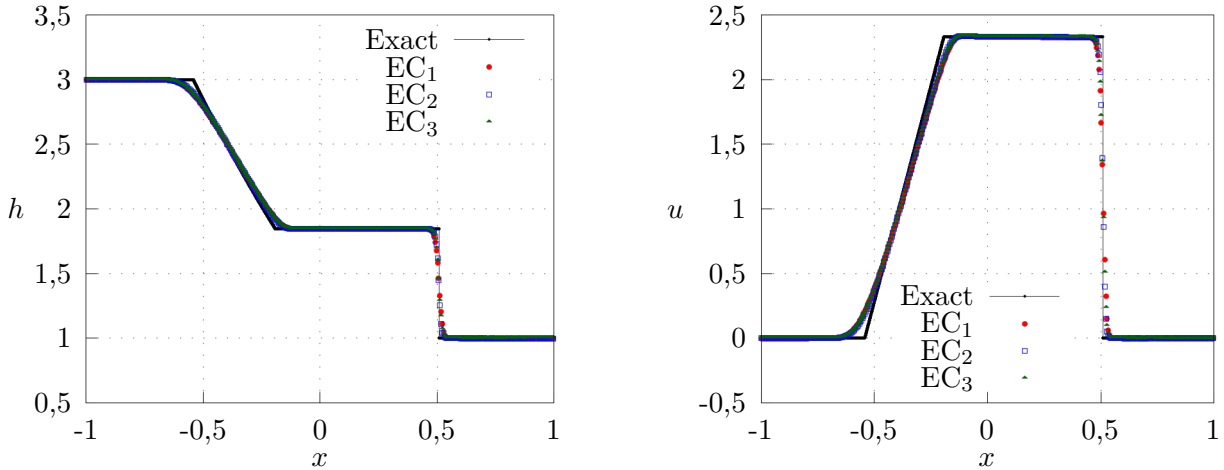
Le premier cas test étudie l'influence du choix des quantités intermédiaires  $h^*$  décrites dans le Lemme 2.3.2. En considérant  $\delta_L > 0, \delta_R > 0, w^{\text{HLL}}$  et  $\eta^{\text{HLL}}$  respectivement donnés par (2.10)-(2.11), trois configurations sont définies. Elles sont désignées par  $\text{EC}_1, \text{EC}_2$  et  $\text{EC}_3$  avec  $\text{EC}_3$  un choix aléatoire entre  $\text{EC}_1$  et  $\text{EC}_2$  noté  $\text{random}(\text{EC}_1, \text{EC}_2)$ . Ces trois configurations s'écrivent

$$\begin{aligned} \text{EC}_1 : & \begin{cases} h_R^* &= h^{\text{HLL}} + \sqrt{\frac{2\delta_L}{g\delta_R}(\eta^{\text{HLL}} - \eta(w^{\text{HLL}}))}, \\ h_L^* &= h^{\text{HLL}} - \sqrt{\frac{2\delta_R}{g\delta_L}(\eta^{\text{HLL}} - \eta(w^{\text{HLL}}))}. \end{cases} \\ \text{EC}_2 : & \begin{cases} h_R^* &= h^{\text{HLL}} - \sqrt{\frac{2\delta_L}{g\delta_R}(\eta^{\text{HLL}} - \eta(w^{\text{HLL}}))}, \\ h_L^* &= h^{\text{HLL}} + \sqrt{\frac{2\delta_R}{g\delta_L}(\eta^{\text{HLL}} - \eta(w^{\text{HLL}}))}. \end{cases} \\ \text{EC}_3 : & \text{random}(\text{EC}_1, \text{EC}_2). \end{aligned} \quad (2.51)$$

Le domaine  $[-1, 1]$  est discrétisé avec 400 cellules et la donnée initiale est

$$h_0(x) = \begin{cases} 3 & \text{si } x < 0,5, \\ 1 & \text{sinon,} \end{cases} \quad u_0(x) = 0. \quad (2.52)$$

Des conditions de bords de type Neumann homogène sont imposées sur les deux frontières du domaine. La solution exacte se compose d'une onde de détente et d'une onde de choc. La date finale de la simulation est  $t = 0,1$  et la Figure 2.2 montre les résultats comparés.





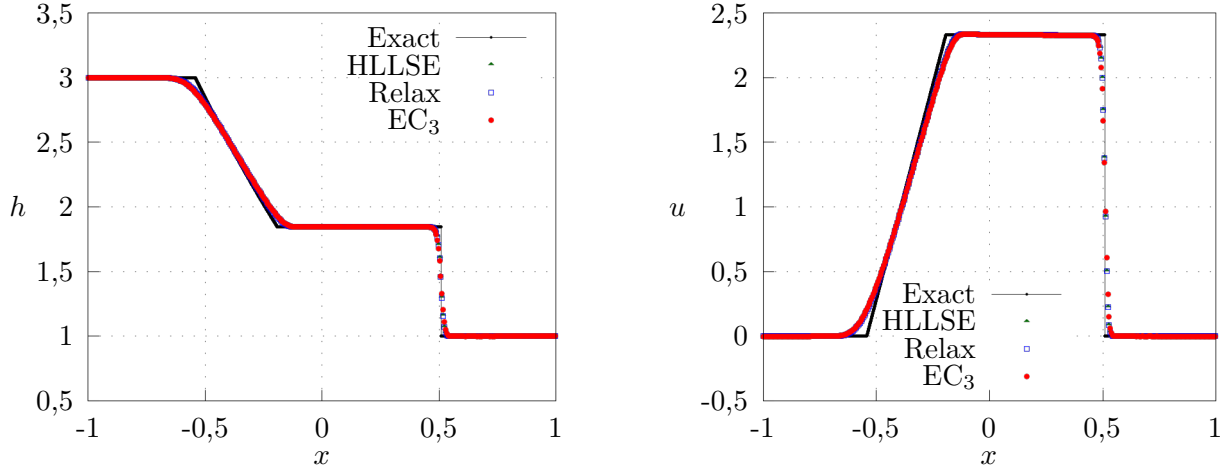


FIGURE 2.2 – Résultats numériques à la date  $t = 0,1$  pour la donnée initiale (2.52), avec la légende  $(EC_i)_{i \in \{1,2,3\}}$  : solveurs du Lemme 2.3.2 avec les configurations (2.51), Relax : schéma de relaxation de Suliciu [32] et HLLSE : solveur [118, Section b.ii].

Les trois configurations  $(EC_i)_{i \in \{1,2,3\}}$  donnent des résultats très similaires et conformes à la solution exacte. Par conséquent, seule la configuration  $EC_3$  est conservée dans la suite.

Le second cas test illustre la procédure  $\min(\max(\cdot, \cdot), \cdot)$  du Lemme 2.3.3 et pour lequel un choix particulier de  $\delta_L$ ,  $\delta_R$  est adopté. Tout d'abord, en considérant  $\lambda_L$ ,  $\lambda_R$ ,  $v^*$  donnés par (2.50),  $\delta_L$  et  $\delta_R$  sont respectivement définis par les égalités  $\delta_L = (v^* - \lambda_L)/(\lambda_R - \lambda_L)$  et  $\delta_R = (\lambda_R - v^*)/(\lambda_R - \lambda_L)$ . Ensuite, les valeurs de  $h_L^*$  et de  $h_R^*$  sont calculées à l'aide de la configuration  $EC_3$ . Si  $h_L^* < 0$  (*resp.*  $h_R^* > 0$ ) alors les égalités  $\delta_L = 0$ ,  $\delta_R = 1$  (*resp.*  $\delta_L = 1$ ,  $\delta_R = 0$ ) sont imposées puis la procédure (2.27) est appliquée. Ce procédé de sélection fait basculer les solveurs étudiés dans cette section vers le solveur HLL [118] aux *voisinages* du bord de  $\Omega$  ce qui assure la stabilité entropique dans de telles zones. Le domaine  $[-1, 1]$  est discrétisé avec 400 cellules et la donnée initiale est

$$h_0(x) = 0,1, \quad u_0(x) = \begin{cases} 10 & \text{si } x < 0,5, \\ 0 & \text{sinon.} \end{cases} \quad (2.53)$$

Des conditions de bords de type Neumann homogène sont imposées sur les deux frontières du domaine. La solution exacte est constituée de deux ondes de chocs. La condition CFL est la même que celle employée pour le cas test précédent. La date finale de la simulation est  $t = 0,1$  et la Figure 2.3 rapporte les résultats comparés.

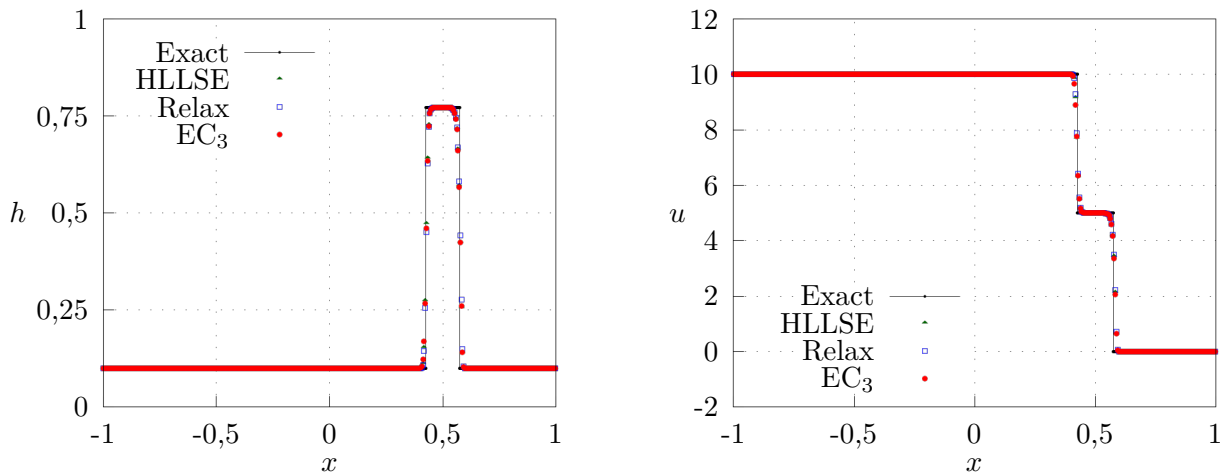


FIGURE 2.3 – Résultats numériques à la date  $t = 0,1$  pour la donnée initiale (2.53), avec la légende  $EC_3$  : configuration  $EC_3$  couplée à la procédure de limitation détaillée dans le Lemme 2.3.3, Relax : schéma de relaxation de Suliciu [32], HLLSE : solveur [118, Section b.ii].

Les performances sont très satisfaisantes. En revanche, sans les techniques de limitations du Lemme 2.3.3, ce problème ne peut pas être simulé avec les solveurs du Lemme 2.3.2.

## 2.5.2 Les équations d'Euler

Les équations d'Euler (2.28) pour un gaz parfait diatomique ( $\gamma = 1,4$ ) sont considérées dans cette section et les schémas décrits dans le Lemme 2.4.4 sont illustrés. Pour toutes les interfaces du maillage ayant de part et d'autre un état constant gauche  $w_L$  et un état constant droit  $w_R$ , on considère  $u^{\text{HLL}}$  donné par (2.32) et les quantités  $\lambda_R$ ,  $\lambda_L$  et  $v^*$  sont égales à

$$\lambda_L = \min_{\alpha \in \{L,R\}} (u_\alpha \pm \sqrt{\gamma p(w_\alpha)/\rho_\alpha}, u_\alpha), \quad \lambda_R = \max_{\alpha \in \{L,R\}} (u_\alpha \pm \sqrt{\gamma p(w_\alpha)/\rho_\alpha}, u_\alpha), \quad v^* = u^{\text{HLL}}, \quad \forall (L, R). \quad (2.54)$$

Le pas de temps  $\Delta t$  est choisi selon la condition CFL (2.9) et avec un nombre CFL égale à 0,5. Pour chaque cas test, un des deux solveurs donnés par (2.44) est sélectionné aléatoirement.

Le premier cas test est le problème du tube à choc de Sod [198]. Le domaine spatial  $[0, 1]$  est discrétisé avec 400 cellules. La condition initiale est

$$\rho_0(x) = \begin{cases} 1 & \text{si } x < 0,5, \\ 0,125 & \text{sinon,} \end{cases} \quad u_0(x) = 0, \quad p_0(x) = \begin{cases} 1 & \text{si } x < 0,5, \\ 0,1 & \text{sinon.} \end{cases} \quad (2.55)$$

Les conditions de bord aux deux extrémités du domaine sont de type Neumann homogène. La solution exacte est composée d'une onde de détente connectée à un contact lui même connecté à une onde de choc. La date finale de la simulation est  $t = 0,2$  et la Figure 2.4 rapporte les résultats comparés.

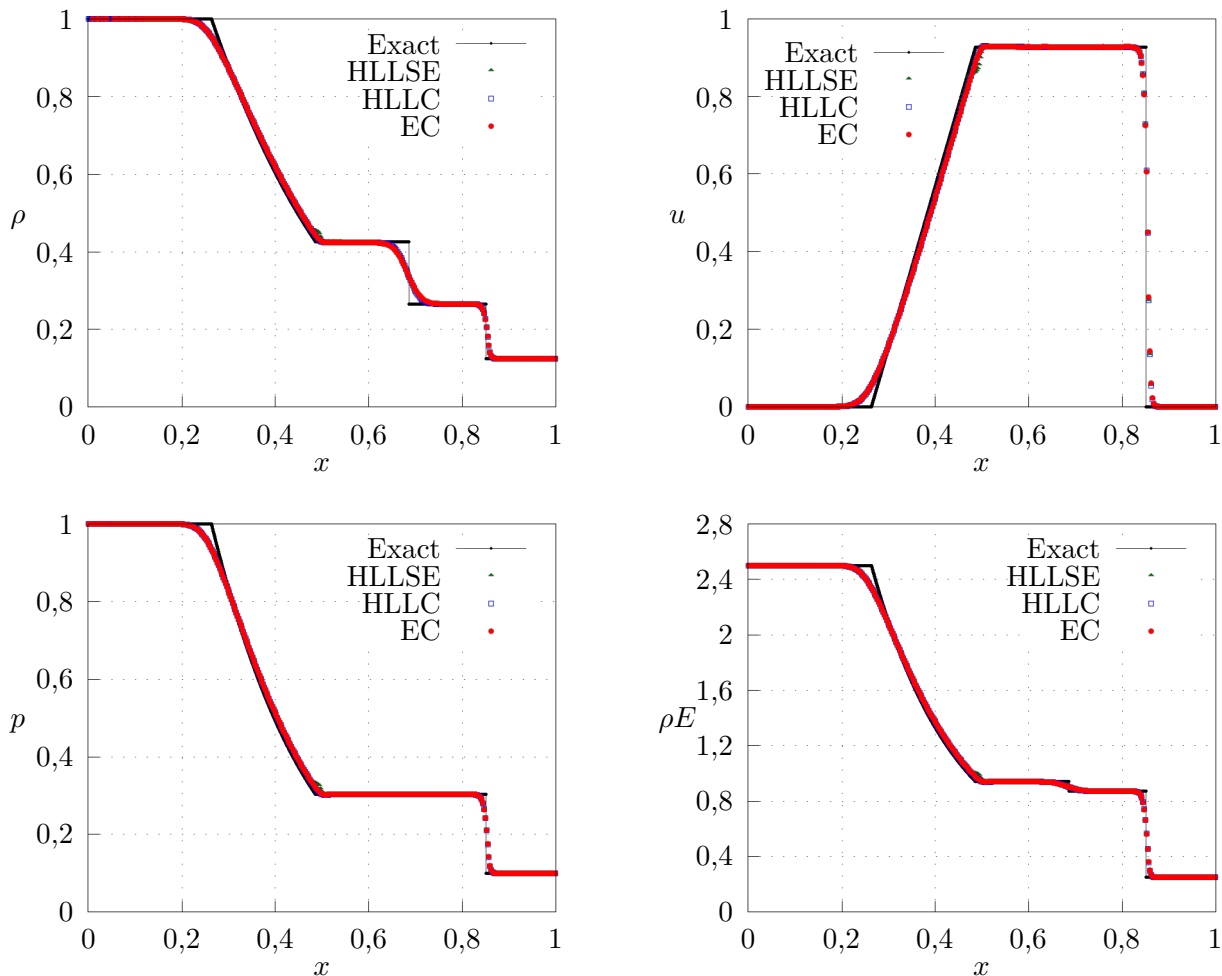


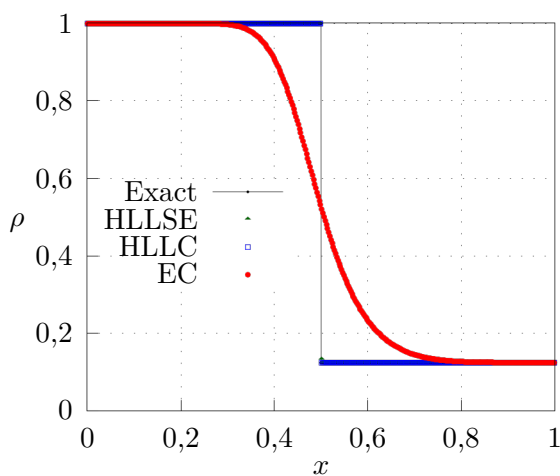
FIGURE 2.4 – Résultats numériques pour le cas test de Sod (2.55), à la date  $t = 0,2$ , avec la légende EC : solveurs du Lemme 2.4.4, HLLC et HLLSE : schémas respectivement décrits dans [213]-[118, Section b.ii].

Les résultats des schémas du Lemme 2.4.4 sont très satisfaisants et conformes aux standards de la littérature.

Le second cas test concerne la simulation des ondes de contact stationnaires. Comme précédemment, le triplet  $\lambda_L, \lambda_R, v^*$ , vérifient (2.54) et la condition CFL sur le pas de temps est donnée par (2.9). La condition initiale est

$$\rho_0(x) = \begin{cases} 1 & \text{si } x < 0,5, \\ 0,125 & \text{sinon,} \end{cases} \quad u_0(x) = 0, \quad p_0(x) = 1. \quad (2.56)$$

Des conditions de bord de Neumann homogènes sont fixées aux deux extrémités du domaine. La solution exacte est une onde contact stationnaire dont la discontinuité est localisée en  $x = 0$ . La date finale de la simulation est  $t = 1,0$ . La Figure 2.5 montre les résultats graphiques et les erreurs numériques.



Erreurs sur $(\rho, u, p)$ pour la préservation des contacts stationnaires.			
	$\rho$		
	$L^1$	$L^2$	$L^\infty$
EC	5,92E-02	1,49E-02	0,46E-00
HLLC	0,00E-00	0,00E-00	0,00E-00
HLLSE	4,34E-05	1,68E-07	8,18E-03
	$u$		
	$L^1$	$L^2$	$L^\infty$
EC	4,16E-15	2,35E-29	8,07E-015
HLLC	0,00E-00	0,00E-00	0,00E-00
HLLSE	6,27E-05	4,39E-09	1,01E-04
	$p$		
	$L^1$	$L^2$	$L^\infty$
EC	1,20E-015	1,97E-30	2,66E-015
HLLC	0,00E-00	0,00E-00	0,00E-00
HLLSE	6,74E-05	4,80E-09	8,17E-05

FIGURE 2.5 – Sur la gauche, résultats numériques à la date  $t = 1,0$  pour le problème de préservation des ondes de contact (2.56) avec la légende EC : solveurs du Lemme 2.4.4, HLLC, HLLSE : schémas respectivement détaillés dans [213]-[118, Section b.ii]. Sur la droite, erreurs numériques commises sur les variables  $\rho, u$  et  $p$  à la date  $t = 1,0$ .

Le schéma HLLC [213] préserve exactement la discontinuité de contact définie par la condition initiale (2.56). Le schéma [118, Section b.ii] commet une erreur très faible sur la variable  $\rho$  et nettement inférieure à l'erreur du premier ordre commise par les schémas du Lemme 2.4.4. Pour cette variable, les schémas proposés dans le Lemme 2.4.4 sont plus sujets à la diffusion numérique que les schémas comparables déjà existants. En revanche, les schémas du Lemme 2.4.4 préservent exactement les quantités  $u$  et  $p$  ce qui n'est pas le cas du schéma décrit dans [118, Section b.ii].

## 2.6 Synthèse et conclusion

Des schémas numériques de type Godunov pour les équations d'Euler et de Saint Venant homogènes ont été introduits dans ce chapitre. Si aucun élément de la séquence  $(w_i^n)_{i \in \mathbb{Z}}$  n'est pas dans un *voisinage* immédiat du bord de  $\Omega$ , les schémas développés satisfont une inégalité d'entropie locale totalement discrète. Dans le cas contraire, la stabilité entropique pourrait être localement perdue. Ces schémas numériques sont constitués de solveurs de Riemann ayant deux états intermédiaires constants. Ces solveurs de Riemann sont définis par des équations quadratiques assurant par construction l'inégalité recherchée. Il est établi que ces équations quadratiques admettent toujours des solutions. Par

conséquent, les schémas numériques obtenus sont simples et entièrement déterminés ce qui n'est pas le cas par exemple d'un schéma comparable détaillé dans [118, Section b.ii]. Les deux solutions des équations quadratiques donnent naissances à deux solveurs distincts. Ces deux solveurs produisent des résultats numériques similaires, tous conformes aux standards de la littérature, mais parfois sujets à la diffusion numérique.

## Chapitre 3

# Solveurs de Riemann *well-balanced* et à stabilité entropique locale pour les systèmes de Saint Venant et de Ripa

Ce chapitre est consacré au développement de schémas de type Godunov pour approcher les solutions des systèmes de Saint Venant et de Ripa respectivement introduits en Section 1.2.1 et 1.2.2. Pour  $z : \mathbb{R} \rightarrow \mathbb{R}$  une fonction régulière donnée, ces deux systèmes d'EDP avec terme source s'écrivent

$$\begin{cases} \partial_t w + \partial_x f(w) = S(w)\partial_x z, & t > 0, \quad x \in \mathbb{R}, \\ w(x, t=0) = w_0(x). \end{cases} \quad (3.1)$$

L'inconnue  $w$  prend ses valeurs dans un espace convexe  $\Omega$  inclus dans  $\mathbb{R}^m$  et en notant  $\hat{w} = (w, z)^T$  dans  $\Omega \times \mathbb{R}$ , les solutions des systèmes de Saint Venant et de Ripa vérifient une inégalité d'entropie qui s'écrit

$$\partial_t \hat{\eta}(\hat{w}) + \partial_x \hat{G}(\hat{w}) \leq 0. \quad (3.2)$$

L'objectif de ce chapitre est de proposer des conditions suffisantes de stabilité entropique pour l'approximation numérique des systèmes de Saint Venant et de Ripa. Si l'inégalité (2.4) donnée dans le Chapitre 2 convient lorsque la fonction  $z$  est constante (voir la Proposition 2.1.1), elle devient malheureusement inopérante dans le cas d'une fonction  $z$  régulière quelconque. En effet, l'introduction d'une telle fonction  $z$  fait perdre la convexité des entropies or c'est cette convexité qui justifie l'inégalité (2.4). Par conséquent, les conditions suffisantes de stabilité entropique dans le cas d'une fonction  $z$  régulière quelconque demande une attention particulière. De plus, comme les systèmes considérés dans ce chapitre possèdent un terme source, ces conditions de stabilité doivent nécessairement être complétées par la propriété *well-balanced* qui assure la préservation des états stationnaires.

### 3.1 Schémas entropiques et *well-balanced* pour les équations de Saint Venant

#### 3.1.1 Condition suffisante de stabilité entropique locale

L'inconnue  $w$  du système de Saint Venant est notée  $(h, hu)^T$  dans l'espace  $\Omega \subset \mathbb{R}^2$  tel que  $\Omega = \{(h, hu) \in \mathbb{R}^2 \mid h \geq 0, hu \in \mathbb{R}\}$ . Pour  $z : \mathbb{R} \rightarrow \mathbb{R}$  une fonction régulière donnée, indépendante du temps, le système d'équations s'écrit

$$\partial_t \begin{pmatrix} h \\ hu \end{pmatrix} + \partial_x \begin{pmatrix} hu \\ hu^2 + gh^2/2 \end{pmatrix} = \begin{pmatrix} 0 \\ -gh \end{pmatrix} \partial_x z. \quad (3.3)$$

Par conséquent, la fonction de flux  $f : \Omega \rightarrow \mathbb{R}^2$  et le terme source  $S : \Omega \rightarrow \mathbb{R}^2$  des équations de Saint Venant (3.3) sont respectivement donnés par

$$f(w) = (hu, hu^2 + gh^2/2)^T,$$

$$S(w) = (0, -gh)^T.$$

Pour alléger les notations, la quantité  $\hat{w} = (w, z)^T$  est introduite et par construction, cette quantité prend ses valeurs dans un espace convexe  $\hat{\Omega}$  qui s'écrit

$$\hat{\Omega} = \left\{ (h, hu, z) \in \mathbb{R}^3 \mid h \geq 0, hu \in \mathbb{R}, z \in \mathbb{R} \right\}.$$

De plus, les solutions du système de Saint Venant (3.3) vérifient une inégalité d'entropie (3.2) pour le couple entropie, flux d'entropie  $(\hat{\eta}, \hat{G})$  défini par :

$$\hat{\eta}(\hat{w}) = \frac{(hu)^2}{2h} + \frac{gh^2}{2} + ghz, \quad \hat{G}(\hat{w}) = \frac{(hu)^3}{2h^2} + g(hu)(h+z), \quad \forall \hat{w} \in \hat{\Omega}. \quad (3.4)$$

Pour décrire le couple  $(\hat{\eta}, \hat{G})$  donné ci-dessus, des fonctions  $\eta$  et  $G$  à valeurs dans  $\mathbb{R}$  sont considérées. Ces fonctions s'écrivent

$$\eta(w) = \frac{(hu)^2}{2h} + \frac{gh^2}{2}, \quad G(w) = \frac{(hu)^3}{2h^2} + gh(hu), \quad \forall w \in \Omega. \quad (3.5)$$

Les fonctions  $(\eta, G)$  ainsi définies forment un couple entropie, flux d'entropie du système de Saint Venant (3.3) dans le cas particulier où la fonction  $z$  est constante. Un calcul direct montre de plus que la fonction  $w \mapsto \eta(w)$  est strictement convexe. Avec ces quantités ainsi définies, il suit immédiatement

$$\hat{\eta}(\hat{w}) = \eta(w) + ghz, \quad \hat{G}(\hat{w}) = G(w) + g(hu)z, \quad \forall \hat{w} \in \hat{\Omega}.$$

D'après la section 1.4.4, l'approximation numérique des solutions du système de Saint Venant (3.3) nécessite d'abord une discrétisation de la fonction  $z$  donnée sous la forme

$$z_i = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} z(x) dx, \quad \forall i \in \mathbb{Z}.$$

Ensuite, les solutions du système (3.3) sont approchées par un schéma de type Godunov. Ce schéma est défini par la donnée d'un solveur de Riemann  $\tilde{w} : \mathbb{R} \times \hat{\Omega} \times \hat{\Omega} \rightarrow \Omega$  qui satisfait la relation de consistance intégrale (1.63). En complétant cette donnée par la notation  $\hat{w}_i^n = (h_i^n, h_i^n u_i^n, z_i)^T$  dans  $\hat{\Omega}$  et par

$$\xi_{i+\frac{1}{2}}^n : (x, t) \mapsto \frac{x - x_{i+\frac{1}{2}}}{t^n + t}, \quad \forall (i, n, x, t) \in \mathbb{Z} \times \mathbb{N} \times ]x_i, x_{i+1}[ \times ]0, \Delta t],$$

un schéma numérique de type Godunov pour le système de Saint Venant (3.3) s'écrit pour tout  $i$  dans  $\mathbb{Z}$

$$w_i^{n+1} = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_i} \tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), \hat{w}_{i-1}^n, \hat{w}_i^n) dx + \frac{1}{\Delta x} \int_{x_i}^{x_{i+\frac{1}{2}}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n) dx. \quad (3.6)$$

Enfin, pour  $(\hat{S}_{i+\frac{1}{2}}^n)_{i \in \mathbb{Z}}$  une séquence de  $\mathbb{R}^2$  qui vérifie la condition de consistance (1.62), la Définition 1.4.4.1 assure l'existence d'une fonction  $\hat{\sigma} : (\hat{\Omega})^2 \rightarrow \mathbb{R}$  telle que

$$\hat{S}_{i+\frac{1}{2}}^n = \left( \begin{array}{c} 0 \\ \hat{\sigma}(\hat{w}_i^n, \hat{w}_{i+1}^n) \end{array} \right) \frac{z_{i+1} - z_i}{\Delta x}, \quad \hat{\sigma}(\hat{w}, \hat{w}) = -gh, \quad \forall (i, \hat{w}) \in \mathbb{Z} \times \hat{\Omega}. \quad (3.7)$$

La fonction  $\hat{\sigma}$  est formellement un paramètre libre qui peut être choisi arbitrairement sous réserve de satisfaire la relation de consistance  $\hat{\sigma}(\hat{w}, \hat{w}) = -gh$ . D'après l'expression (3.7), cette fonction  $\hat{\sigma}$  définit totalement la séquence  $(\hat{S}_{i+\frac{1}{2}}^n)_{i \in \mathbb{Z}}$  et puisqu'il n'y a pas de confusion possible, la notation suivante est désormais adoptée :

$$\hat{s}_{i+\frac{1}{2}} = \hat{\sigma}(\hat{w}_i^n, \hat{w}_{i+1}^n) \frac{z_{i+1} - z_i}{\Delta x}, \quad \forall i \in \mathbb{Z}.$$

En s'appuyant sur la donnée d'une telle séquence  $(\hat{S}_{i+1/2}^n)_{i \in \mathbb{Z}}$ , le Lemme 1.4.4.1 permet une réécriture équivalente du schéma de type Godunov (3.6) sous la forme

$$w_i^{n+1} = w_i^n - \frac{\Delta t}{\Delta x} (\hat{\mathcal{F}}(\hat{w}_i^n, \hat{w}_{i+1}^n) - \hat{\mathcal{F}}(\hat{w}_{i-1}^n, \hat{w}_i^n)) + \frac{\Delta t}{2} (\hat{S}_{i+1/2}^n + \hat{S}_{i-1/2}^n), \quad \forall i \in \mathbb{Z},$$

avec,

$$\begin{aligned} \hat{\mathcal{F}}(\hat{w}_i^n, \hat{w}_{i+1}^n) &= \frac{f(w_{i+1}^n) + f(w_i^n)}{2} - \frac{\Delta x}{4\Delta t} (w_{i+1}^n - w_i^n) + \frac{1}{2\Delta t} \int_{x_{i+1/2}}^{x_{i+1}} \tilde{w}(\xi_{i+1/2}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n) dx \\ &- \frac{1}{2\Delta t} \int_{x_i}^{x_{i+1/2}} \tilde{w}(\xi_{i+1/2}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n) dx. \end{aligned} \quad (3.8)$$

L'objectif de cette section est de proposer une condition suffisante de stabilité entropique locale *i.e* à chaque interface du maillage et pour le couple  $(\hat{\eta}, \hat{G})$  défini par (3.4). Comme le système de Saint Venant (3.3) possède un terme source, cette proposition doit aussi s'articuler avec la préservation des solutions stationnaires. Cette préservation est assurée par la propriété *well-balanced* qui est maintenant définie.

**Définition 3.1.1.1** (Équilibre local et schéma *well-balanced* pour les équations de Saint Venant (3.3)). *Considérons les équations de Saint Venant (3.3) complétées par la notation*

$$\hat{B}(\hat{w}) = \frac{u^2}{2} + g(h+z), \quad \forall \hat{w} \in \hat{\Omega}.$$

À la date  $t^n$ , considérons  $\hat{w}_i^n$  et  $\hat{w}_{i+1}^n$  deux états constants dans  $\hat{\Omega}$  respectivement à gauche et à droite de l'interface  $x_{i+1/2}$ .

i) Les deux états  $\hat{w}_i^n$  et  $\hat{w}_{i+1}^n$  sont dits à l'équilibre local pour les équilibres mouvants si

$$h_i^n u_i^n = h_{i+1}^n u_{i+1}^n, \quad \hat{B}(\hat{w}_i^n) = \hat{B}(\hat{w}_{i+1}^n). \quad (3.9)$$

ii) Les deux états  $\hat{w}_i^n$  et  $\hat{w}_{i+1}^n$  sont dits à l'équilibre local pour le lac au repos si

$$u_i^n = u_{i+1}^n = 0, \quad h_i^n + z_i = h_{i+1}^n + z_{i+1}. \quad (3.10)$$

iii) Un schéma numérique de type Godunov (3.6) est dit *well-balanced* pour les équilibres mouvants s'il préserve exactement une séquence  $(\hat{w}_i^n)_{i \in \mathbb{Z}}$  qui satisfait à chaque interface du maillage un équilibre local (3.9). Cette propriété s'écrit

$$\text{si } \forall i \in \mathbb{Z} \begin{cases} h_i^n u_i^n = h_{i+1}^n u_{i+1}^n, \\ \hat{B}(\hat{w}_i^n) = \hat{B}(\hat{w}_{i+1}^n), \end{cases} \quad \text{alors } w_i^{n+1} = w_i^n, \quad \forall i \in \mathbb{Z}.$$

Un schéma numérique *well-balanced* pour le lac au repos s'énonce mutadis mutandis en considérant les équilibres locaux (3.10). Cette propriété s'écrit dans ce cas

$$\text{si } \forall i \in \mathbb{Z} \begin{cases} u_i^n = u_{i+1}^n = 0, \\ h_i^n + z_i = h_{i+1}^n + z_{i+1}, \end{cases} \quad \text{alors } w_i^{n+1} = w_i^n, \quad \forall i \in \mathbb{Z}.$$

D'après les égalités (3.9)-(3.10), les équilibres locaux des équations de Saint Venant (3.3) sont caractérisés par la continuité de  $hu$  pour tout doublet  $(z_i, z_{i+1})$  dans  $(\mathbb{R})^2$ . Lorsque  $z_i$  et  $z_{i+1}$  sont égaux, la continuité de  $h$  est également vérifiée pour le cas du lac au repos (3.10). Ainsi, dans ce mémoire de thèse, on suppose que les équilibres mouvants locaux (3.9) vérifient également la continuité de  $h$  lorsque  $z_i$  et  $z_{i+1}$  sont égaux. Cette hypothèse s'écrit

$$\text{si } \begin{cases} z_i = z_{i+1} \\ h_i^n u_i^n = h_{i+1}^n u_{i+1}^n, \\ \hat{B}(\hat{w}_i^n) = \hat{B}(\hat{w}_{i+1}^n), \end{cases} \quad \text{alors } h_i^n = h_{i+1}^n. \quad (3.11)$$

En dehors des singularités, le Lemme A.1-ii) donné en Annexe A établit que les solutions stationnaires régulières du système de Saint Venant (3.3) vérifient un analogue continu des égalités (3.11). Ces inégalités signifient formellement que si  $z$  est localement constante alors les solutions stationnaires régulières sont nécessairement les constantes de  $\hat{\Omega}$ .

Avec la Définition 3.1.1.1, il est maintenant possible de donner une condition suffisante à la propriété *well-balanced*. Cette condition est par exemple détaillée dans [25, 165, 166] et elle est rappelée dans la proposition suivante qui donne également une condition suffisante de stabilité entropique locale pour le couple  $(\hat{\eta}, \hat{G})$  défini par (3.4).

**Proposition 3.1.1.1** (Solveur de Riemann entropique et *well-balanced* pour le système de Saint Venant). *Considérons le couple de fonctions  $(\eta, G)$  à valeurs dans  $\mathbb{R}$  donné par (3.5). Soient  $(\hat{S}_{i+1/2}^n)_{i \in \mathbb{Z}}$  une suite de  $\mathbb{R}^2$  et un solveur de Riemann  $\tilde{w} : \mathbb{R} \times \hat{\Omega} \times \hat{\Omega} \rightarrow \Omega$  satisfaisant respectivement les conditions de consistances des Définitions 1.4.4.1 et 1.4.4.2. Notons  $\mathcal{F}_{i+1/2}^h$  le flux numérique de la variable  $h$  donné par la première composante de  $\hat{\mathcal{F}}(\hat{w}_i^n, \hat{w}_{i+1}^n)$  défini dans (3.8). Supposons qu'une condition CFL de non interaction est satisfaite.*

i) *Si le solveur de Riemann vérifie la relation de consistence intégrale*

$$\frac{1}{\Delta x} \int_{x_i}^{x_{i+1}} \tilde{w}(\xi_{i+1/2}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n) dx = \frac{w_i^n + w_{i+1}^n}{2} - \frac{\Delta t}{\Delta x} (f(w_{i+1}^n) - f(w_i^n)) + \Delta t \hat{S}_{i+1/2}^n, \quad \forall i \in \mathbb{Z}, \quad (3.12)$$

*alors le schéma de type Godunov (3.6) est consistant avec les équations de Saint Venant (3.3).*

ii) *Si le solveur de Riemann satisfait l'inégalité*

$$\begin{aligned} \frac{1}{\Delta x} \int_{x_i}^{x_{i+1}} \eta(\tilde{w}(\xi_{i+1/2}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n)) dx \leq \\ \frac{\eta(w_i^n) + \eta(w_{i+1}^n)}{2} - \frac{\Delta t}{\Delta x} (G(w_{i+1}^n) - G(w_i^n)) - g \frac{\Delta t}{\Delta x} (z_{i+1} - z_i) \mathcal{F}_{i+1/2}^h, \quad \forall i \in \mathbb{Z}, \end{aligned} \quad (3.13)$$

*alors le schéma de type Godunov (3.6) vérifie une inégalité d'entropie discrète locale pour le couple  $(\hat{\eta}, \hat{G})$  défini dans (3.4). Cette inégalité s'écrit*

$$\frac{\hat{\eta}(\hat{w}_i^{n+1}) - \hat{\eta}(\hat{w}_i^n)}{\Delta t} + \frac{\hat{G}_{i+1/2} - \hat{G}_{i-1/2}}{\Delta x} \leq 0, \quad \forall i \in \mathbb{Z}, \quad (3.14)$$

*avec,*

$$\begin{aligned} \hat{G}_{i+1/2} &= \frac{G(w_{i+1}^n) + G(w_i^n)}{2} + g \mathcal{F}_{i+1/2}^h \frac{z_{i+1} + z_i}{2} - \frac{\Delta x}{4\Delta t} (\eta(w_{i+1}^n) - \eta(w_i^n)) \\ &+ \frac{1}{2\Delta t} \int_{x_{i+1/2}}^{x_{i+1}} \eta(\tilde{w}(\xi_{i+1/2}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n)) dx \\ &- \frac{1}{2\Delta t} \int_{x_i}^{x_{i+1/2}} \eta(\tilde{w}(\xi_{i+1/2}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n)) dx. \end{aligned} \quad (3.15)$$

iii) *Si le solveur de Riemann satisfait*

$$\tilde{w}(\xi_{i+1/2}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n) = \begin{cases} w_i^n & \text{si } x - x_{i+1/2} < 0, \\ w_{i+1}^n & \text{sinon,} \end{cases} \quad \forall i \in \mathbb{Z}, \quad (3.16)$$

*dès que la séquence  $(\hat{w}_i^n)_{i \in \mathbb{Z}}$  vérifie, à chaque interface du maillage, un équilibre local pour les équilibres mouvants (3.9) (resp. pour le lac au repos (3.10)) alors le schéma de type Godunov (3.6) est *well-balanced* pour les équilibres mouvants (resp. pour le lac au repos).*



*Démonstration.* Concernant l'assertion *i*), d'après l'expression de  $\hat{S}_{i+1/2}^n$  donnée par (3.7) et comme  $z$  est régulière, il est clair que  $\hat{S}_{i+1/2}^n$  est consistant avec  $(0, -gh\partial_x z)^\top$  et il vérifie de plus

$$\hat{S}(\hat{w}, \hat{w}) = 0, \quad \forall \hat{w} \in \hat{\Omega}. \quad (3.17)$$

Par conséquent, un solveur de Riemann qui vérifie la relation de consistance intégrale (3.12) donne un schéma consistant.

Pour montrer l'assertion *ii*), considérons la fonction  $w \mapsto \eta(w)$  définie par (3.5). Comme cette fonction est convexe et qu'elle satisfait les conditions décrites dans [173], l'inégalité de Jensen appliquée au schéma de type Godunov (3.6) donne

$$\begin{aligned} \eta(w_i^{n+1}) &\leq \\ &\frac{1}{2\Delta x} \int_{x_i}^{x_{i+\frac{1}{2}}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n)) dx + \frac{1}{2\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_i} \eta(\tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), \hat{w}_{i-1}^n, \hat{w}_i^n)) dx \\ &+ \frac{1}{2\Delta x} \int_{x_{i-1}}^{x_i} \eta(\tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), \hat{w}_{i-1}^n, \hat{w}_i^n)) dx + \frac{1}{2\Delta x} \int_{x_i}^{x_{i+1}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n)) dx \\ &- \frac{1}{2\Delta x} \int_{x_{i-1}}^{x_{i-\frac{1}{2}}} \eta(\tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), \hat{w}_{i-1}^n, \hat{w}_i^n)) dx - \frac{1}{2\Delta x} \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n)) dx. \end{aligned}$$

En utilisant l'inégalité (3.13) dans l'inégalité ci-dessus, il suit

$$\begin{aligned} \eta(w_i^{n+1}) &\leq \\ &\frac{1}{2\Delta x} \int_{x_i}^{x_{i+\frac{1}{2}}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n)) dx + \frac{1}{2\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_i} \eta(\tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), \hat{w}_{i-1}^n, \hat{w}_i^n)) dx \\ &+ \frac{1}{2} \left( \frac{\eta(w_{i-1}^n) + \eta(w_i^n)}{2} - \frac{\Delta t}{\Delta x} (G(w_i^n) - G(w_{i-1}^n)) - \frac{\Delta t}{\Delta x} g(z_i - z_{i-1}) \mathcal{F}_{i-\frac{1}{2}}^h \right) \\ &+ \frac{1}{2} \left( \frac{\eta(w_i^n) + \eta(w_{i+1}^n)}{2} - \frac{\Delta t}{\Delta x} (G(w_{i+1}^n) - G(w_i^n)) - \frac{\Delta t}{\Delta x} g(z_{i+1} - z_i) \mathcal{F}_{i+\frac{1}{2}}^h \right) \\ &- \frac{1}{2\Delta x} \int_{x_{i-1}}^{x_{i-\frac{1}{2}}} \eta(\tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), \hat{w}_{i-1}^n, \hat{w}_i^n)) dx - \frac{1}{2\Delta x} \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n)) dx. \end{aligned} \quad (3.18)$$

Par ailleurs, comme il n'y a pas de contribution du terme source pour la première équation du système de Saint Venant (3.3), le schéma numérique pour la variable  $h$  s'écrit

$$h_i^{n+1} = h_i^n - \frac{\Delta t}{\Delta x} (\mathcal{F}_{i+\frac{1}{2}}^h - \mathcal{F}_{i-\frac{1}{2}}^h), \quad \forall i \in \mathbb{Z}. \quad (3.19)$$

En multipliant l'équation (3.19) par  $gz_i$  puis en additionnant le résultat à l'inégalité (3.18), on obtient une inégalité de la forme

$$\eta(w_i^{n+1}) + gh_i^{n+1} z_i \leq \eta(w_i^n) + gh_i^n z_i - \frac{\Delta t}{\Delta x} (\hat{\mathcal{G}}_{i+\frac{1}{2}} - \hat{\mathcal{G}}_{i-\frac{1}{2}}), \quad \forall i \in \mathbb{Z},$$

avec  $\hat{\mathcal{G}}_{i+\frac{1}{2}}$  défini par (3.15). Comme un calcul immédiat montre que  $\hat{\eta}(\hat{w}) = \eta(w) + ghz$  et puisque la fonction  $z$  est indépendante du temps, cette dernière inégalité se réécrit sous la forme de (3.14). Avant de conclure la preuve de l'assertion *ii*), il faut maintenant s'assurer que la quantité  $\hat{\mathcal{G}}_{i+\frac{1}{2}}$  donnée par (3.15) est consistante avec  $\hat{G}$  définie par (3.5).

En considérant l'égalité de consistance  $\tilde{w}(\xi, \hat{w}, \hat{w}) = w$  donnée par la Définition 1.4.4.2-iii) ainsi que la consistance du flux numérique  $\mathcal{F}_{i+\frac{1}{2}}^h$ , il suit

$$\hat{G}(\hat{w}, \hat{w}) = \frac{G(w) + G(w)}{2} + g\mathcal{F}^h(w, w)z - \frac{\Delta x}{4\Delta t} (\eta(w) - \eta(w))$$

$$\begin{aligned}
& + \frac{1}{2\Delta t} \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}, \hat{w})) \, dx - \frac{1}{2\Delta t} \int_{x_i}^{x_{i+\frac{1}{2}}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}, \hat{w})) \, dx, \\
& = G(w) + g(hu)z + \frac{1}{2\Delta t} \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \eta(w) \, dx - \frac{1}{2\Delta t} \int_{x_i}^{x_{i+\frac{1}{2}}} \eta(w) \, dx, \\
& = G(w) + g(hu)z.
\end{aligned}$$

Puisqu'un calcul direct montre que  $\hat{G}(\hat{w}) = G(w) + g(hu)z$ , cette dernière égalité termine de montrer l'assertion *ii*).

Enfin, pour l'assertion *iii*), supposons que la séquence  $(\hat{w}_i^n)_{i \in \mathbb{Z}}$  est telle qu'un équilibre local est satisfait à chaque interface du maillage. Dans ce cas, en utilisant la condition (3.16) dans un schéma de type Godunov (3.6), il suit

$$\begin{aligned}
w_i^{n+1} & = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_i} \tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), \hat{w}_{i-1}^n, \hat{w}_i^n) \, dx + \frac{1}{\Delta x} \int_{x_i}^{x_{i+\frac{1}{2}}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n) \, dx, \\
& = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_i} w_i^n \, dx + \frac{1}{\Delta x} \int_{x_i}^{x_{i+\frac{1}{2}}} w_i^n \, dx, \\
& = w_i^n.
\end{aligned}$$

D'après la Définition 3.1.1.1-*iii*), cette dernière égalité donne la propriété *well-balanced* annoncée ce qui conclut la preuve.  $\square$

Comme l'entropie  $\hat{w} \mapsto \hat{\eta}(\hat{w})$  définie dans (3.4) n'est pas convexe, la preuve de l'inégalité d'entropie discrète (3.14) s'appuie sur la condition suffisante (3.13) mais également sur l'égalité (3.19) issue du schéma numérique. Cette preuve trouve des similitudes avec un calcul évoqué dans [32, Section 4.4] mais l'originalité de la Proposition 3.1.1.1 tient à la définition de l'inégalité (3.13). Cette inégalité est une extension de (2.4) au cas d'une fonction  $z$  quelconque et son utilisation dans la définition d'un solveur de Riemann assure la stabilité entropique discrète locale (3.14).

Ainsi, la section suivante est consacrée à la construction d'un solveur de Riemann constant par morceaux à deux états intermédiaires symétriques satisfaisant la relation de consistance intégrale (3.12) et l'inégalité d'entropie locale (3.13).

### 3.1.2 Solveur de Riemann à deux états intermédiaires symétriques, consistant et entropique

Pour définir un solveur de Riemann à deux états intermédiaires symétriques, les notations suivantes sont tout d'abord adoptées

$$\begin{aligned}
t^n & = 0, \\
x_{i+\frac{1}{2}} & = 0, \\
(w_i^n, z_i) & = (w_L, z_L) = \hat{w}_L, \\
(w_{i+1}^n, z_{i+1}) & = (w_R, z_R) = \hat{w}_R, \\
\hat{s}_{i+\frac{1}{2}} & = \hat{s}_{LR}.
\end{aligned} \tag{3.20}$$

Ensuite, en considérant un réel  $\lambda > 0$ , une condition CFL est donnée par les inégalités

$$\lambda \geq \max_{\alpha \in \{L, R\}} |u_\alpha \pm \sqrt{gh_\alpha}| \quad \text{et} \quad \frac{\lambda \Delta t}{\Delta x} \leq \frac{1}{2}. \tag{3.21}$$

Les inégalités ci-dessus sont supposées vérifiées et par conséquent, il est possible de définir un solveur de Riemann constant par morceaux à deux états intermédiaires symétriques. Un tel solveur est noté

$\tilde{w}(\cdot, \hat{w}_L, \hat{w}_R) : \mathbb{R} \rightarrow \mathbb{R}^2$  et il s'écrit

$$\tilde{w}(x/\Delta t, \hat{w}_L, \hat{w}_R) = \begin{cases} w_L & \text{si } \frac{x}{\Delta t} \leq -\lambda, \\ w_L^* & \text{si } -\lambda < \frac{x}{\Delta t} \leq 0, \\ w_R^* & \text{si } 0 < \frac{x}{\Delta t} \leq \lambda, \\ w_R & \text{si } \lambda < \frac{x}{\Delta t}. \end{cases} \quad (3.22)$$

Comme l'illustre la Figure 3.1, les solveurs ci-dessus sont des versions simplifiées des solveurs (2.12) utilisés dans le Chapitre 2. Les solveurs de Riemann (2.12) pourraient être employés dans ce chapitre à condition d'imposer  $v^* = 0$ .

Enfin, le doublet  $(w^{\text{HLL}}, \eta^{\text{HLL}})$  dans  $\mathbb{R}^2 \times \mathbb{R}$  est considéré et ses éléments se lisent respectivement

$$w^{\text{HLL}} = (h^{\text{HLL}}, (hu)^{\text{HLL}})^T = \frac{w_R + w_L}{2} - \frac{f(w_R) - f(w_L)}{2\lambda}, \quad (3.23a)$$

$$\eta^{\text{HLL}} = \frac{\eta(w_R) + \eta(w_L)}{2} - \frac{G(w_R) - G(w_L)}{2\lambda}. \quad (3.23b)$$

Avec ces notations, le Lemme 2.3.1 permet de déduire que si les inégalités (3.21) sont satisfaites et si  $h_L$  et  $h_R$  sont strictement positifs (*resp.* positifs) alors  $h^{\text{HLL}}$  est strictement positif (*resp.* positif). De plus, l'inégalité  $\eta(w^{\text{HLL}}) \leq \eta^{\text{HLL}}$  est toujours vérifiée dès lors que  $\lambda$  satisfait (3.21).

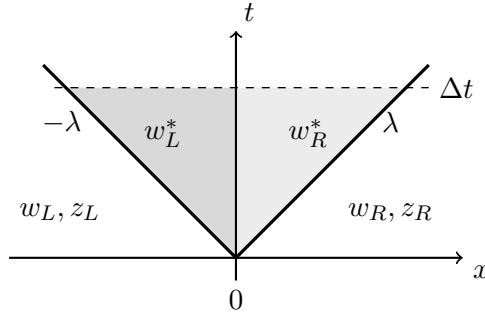


FIGURE 3.1 – Solvateur de Riemann approché constitué de deux états intermédiaires symétriques constants  $w_L^*$  et  $w_R^*$  représentés dans le plan  $(x, t)$ .

Ainsi, les états intermédiaires  $w_L^*$  et  $w_R^*$  du solveur de Riemann (3.22) doivent être déterminés et ils sont notés

$$w_\alpha^* = (h_\alpha^*, h_\alpha^* u_\alpha^*)^T, \quad \forall \alpha \in \{L, R\}.$$

On définit maintenant un solveur de Riemann de la forme (3.22) en considérant la relation de consistance intégrale (3.12) et une équation supplémentaire obtenue en posant l'égalité des deux membres de (3.13). Le lemme ci-dessous présente les équations vérifiées par  $w_L^*$  et  $w_R^*$ .

**Lemme 3.1.2.1.** *Soient  $\hat{w}_L, \hat{w}_R$  deux états constants de l'intérieur de  $\hat{\Omega}$ ,  $\hat{s}_{LR}$  une discrétisation consistante de  $\frac{1}{\Delta x} \int_{-\Delta x/2}^{\Delta x/2} (-gh \partial_x z) dx$  au sens de l'expression (3.7) et  $\tilde{w}(\cdot, \hat{w}_L, \hat{w}_R) : \mathbb{R} \rightarrow \mathbb{R}^2$  un solveur de Riemann de la forme (3.22). Considérons les couples  $(\hat{\eta}, \hat{G})$  et  $(\eta, G)$  respectivement donnés par (3.4), (3.5). Supposons que la condition CFL de non interaction (3.21) est satisfaite.*

*Considérons  $w^{\text{HLL}}$  défini dans (3.23a) et notons*

$$h^{\text{HLL}} \hat{u}^{\text{HLL}} = (hu)^{\text{HLL}} + \frac{\Delta x \hat{s}_{LR}}{2\lambda}, \quad (3.24a)$$

$$\hat{w}^{\text{HLL}} = (h^{\text{HLL}}, h^{\text{HLL}} \hat{u}^{\text{HLL}}, \frac{z_L + z_R}{2})^T, \quad (3.24b)$$

$$\hat{\eta}^{\text{HLL}} = \frac{\hat{\eta}(\hat{w}_R) + \hat{\eta}(\hat{w}_L)}{2} - \frac{\hat{G}(\hat{w}_R) - \hat{G}(\hat{w}_L)}{2\lambda}. \quad (3.24c)$$

Il y a équivalence entre les égalités

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \tilde{w}(x/\Delta t, \hat{w}_L, \hat{w}_R) dx = \frac{w_L + w_R}{2} - \frac{\Delta t}{\Delta x} (f(w_R) - f(w_L)) + \Delta t (0, \hat{s}_{LR})^T, \quad (3.25a)$$

$$\begin{aligned} \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \eta(\tilde{w}(x/\Delta t, \hat{w}_L, \hat{w}_R)) dx = \\ \frac{\eta(w_L) + \eta(w_R)}{2} - \frac{\Delta t}{\Delta x} (G(w_R) - G(w_L)) - g \frac{\Delta t}{\Delta x} (z_R - z_L) \mathcal{F}_{LR}^h, \end{aligned} \quad (3.25b)$$

et les égalités

$$\frac{h_L^* + h_R^*}{2} = h^{\text{HLL}}, \quad (3.26a)$$

$$\frac{h_L^* u_L^* + h_R^* u_R^*}{2} = h^{\text{HLL}} \hat{u}^{\text{HLL}}, \quad (3.26b)$$

$$\frac{h_L^* h_R^*}{8h^{\text{HLL}}} (u_R^* - u_L^*)^2 + \frac{g}{8} (h_R^* - h_L^* + z_R - z_L)^2 = \hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8} (z_R - z_L)^2. \quad (3.26c)$$

Avant de procéder à la démonstration du lemme ci-dessus, il est important de souligner que la quantité  $h^{\text{HLL}} \hat{u}^{\text{HLL}}$  définie dans (3.24a) est paramétrée par la donnée de  $\hat{s}_{LR}$ . Cette donnée doit satisfaire la condition de consistante rappelée dans (3.7) et des expressions possibles seront proposées dans la suite de ce mémoire de thèse. Avec cette remarque, la démonstration du Lemme 3.1.2.1 peut désormais être établie.

*Démonstration du Lemme 3.1.2.1.* Dans un premier temps, on montre que le système d'équations (3.25) implique le système (3.26). Puisqu'une condition CFL de non interaction est vérifiée et en notant  $\nu = \lambda \Delta t / \Delta x$ , il est possible de récrire de manière équivalente le membre de gauche de l'équation (3.25a) de la façon suivante :

$$\begin{aligned} \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \tilde{w}(x/\Delta t, \hat{w}_L, \hat{w}_R) dx \\ = \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{-\lambda \Delta t} w_L dx + \frac{1}{\Delta x} \int_{-\lambda \Delta t}^0 w_L^* dx + \frac{1}{\Delta x} \int_0^{\lambda \Delta t} w_R^* dx + \frac{1}{\Delta x} \int_{\lambda \Delta t}^{\frac{\Delta x}{2}} w_R dx, \\ = (-\nu + 1/2)w_L + \nu w_L^* + \nu w_R^* + (1/2 - \nu)w_R. \end{aligned}$$

Par conséquent, l'équation (3.25a) est équivalente à

$$(-\nu + 1/2)w_L + \nu w_L^* + \nu w_R^* + (1/2 - \nu)w_R = \frac{w_L + w_R}{2} - \nu \frac{f(w_R) - f(w_L)}{\lambda} + \nu (0, \Delta x \hat{s}_{LR} / \lambda, 0)^T.$$

En écrivant composante par composante l'équation ci-dessus puis en divisant de part et d'autre par  $2\nu$ , on en déduit les équations (3.26a)-(3.26b).

Il s'agit maintenant de récrire l'équation (3.25b). En développant l'intégrale du membre de gauche de (3.25b) et en conservant la notation  $\nu = \lambda \Delta t / \Delta x$ , il suit

$$\begin{aligned} \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \eta(\tilde{w}(x/\Delta t, \hat{w}_L, \hat{w}_R)) dx \\ = \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{-\lambda \Delta t} \eta(w_L) dx + \frac{1}{\Delta x} \int_{-\lambda \Delta t}^0 \eta(w_L^*) dx + \frac{1}{\Delta x} \int_0^{\lambda \Delta t} \eta(w_R^*) dx + \frac{1}{\Delta x} \int_{\lambda \Delta t}^{\frac{\Delta x}{2}} \eta(w_R) dx, \\ = (-\nu + 1/2)\eta(w_L) + \nu\eta(w_L^*) + \nu\eta(w_R^*) + (1/2 - \nu)\eta(w_R). \end{aligned}$$

Avec l'expression ci-dessus, l'équation (3.25b) est équivalente à

$$(-\nu + 1/2)\eta(w_L) + \nu\eta(w_L^*) + \nu\eta(w_R^*) + (1/2 - \nu)\eta(w_R) = \frac{\eta(w_L) + \eta(w_R)}{2} - \nu \frac{G(w_R) - G(w_L)}{\lambda} - g\nu \frac{z_R - z_L}{\lambda} \mathcal{F}_{LR}^h.$$

En multipliant cette égalité par  $1/(2\nu) > 0$  et en utilisant la quantité  $\eta^{\text{HLL}}$  définie dans (3.23b), on obtient une nouvelle forme équivalente de l'équation (3.25b) donnée par

$$\frac{\eta(w_L^*) + \eta(w_R^*)}{2} = \eta^{\text{HLL}} - g \frac{z_R - z_L}{2\lambda} \mathcal{F}_{LR}^h. \quad (3.27)$$

Pour récrire le flux numérique  $\mathcal{F}_{LR}^h$ , la notation  $[\cdot] = \cdot_R - \cdot_L$  est maintenant adoptée. D'après la forme du solveur de Riemann (3.22), cette quantité s'écrit

$$\mathcal{F}_{LR}^h = \frac{(hu)_R + (hu)_L}{2} - \frac{\lambda}{2}[h] + \frac{\lambda}{2}[h^*].$$

Par conséquent, et en utilisant la définition de  $h^{\text{HLL}}$  énoncée dans (3.23a), il suit

$$\begin{aligned} -\frac{[z]}{2\lambda} \mathcal{F}_{LR}^h &= -\frac{[z][h^*]}{4} + \frac{[z]}{4} \left( [h] - \frac{(hu)_L + (hu)_R}{\lambda} \right), \\ &= -\frac{[z][h^*]}{4} + \frac{z_R h_R + z_L h_L}{2} - \frac{z_L + z_R}{2} h^{\text{HLL}} - \frac{[huz]}{2\lambda}. \end{aligned}$$

À présent, et en introduisant l'expression ci-dessus dans l'équation (3.27) à l'aide de  $\hat{\eta}^{\text{HLL}}$  défini par (3.24c) puis en explicitant les expressions de  $\eta(w_L^*)$  et de  $\eta(w_R^*)$ , on obtient

$$\frac{h_R^*(u_R^*)^2 + h_L^*(u_L^*)^2}{4} + g \frac{(h_L^*)^2 + (h_R^*)^2}{4} + \frac{g}{4}[z][h^*] = \hat{\eta}^{\text{HLL}} - g \frac{z_L + z_R}{2} h^{\text{HLL}}. \quad (3.28)$$

De plus, en élevant au carré l'équation (3.26a) puis en multipliant le résultat par  $g/2$ , il est possible d'écrire

$$\frac{g}{8}(h_L^* + h_R^*)^2 = \frac{g}{2}(h^{\text{HLL}})^2.$$

Ainsi, en soustrayant l'équation ci-dessus à l'équation (3.28), il suit

$$\frac{h_R^*(u_R^*)^2 + h_L^*(u_L^*)^2}{4} + \frac{g}{8}[h^*]^2 + \frac{g}{4}[z][h^*] = \hat{\eta}^{\text{HLL}} - \frac{g}{2}(h^{\text{HLL}})^2 - g \frac{z_L + z_R}{2} h^{\text{HLL}}. \quad (3.29)$$

L'étape de simplification suivante consiste à récrire la quantité  $h_L^*(u_L^*)^2 + h_R^*(u_R^*)^2$  pour montrer que l'équation (3.29) est équivalente à

$$\frac{h_R^* h_L^* [u^*]^2}{8h^{\text{HLL}}} + \frac{g}{8}[h^*]^2 + \frac{g}{4}[h^*][z] = \hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}). \quad (3.30)$$

En utilisant les équations (3.26a)-(3.26b) il vient

$$\begin{aligned} \frac{h_L^*(u_L^*)^2 + h_R^*(u_R^*)^2}{4} &= \frac{h_L^*(u_L^*)^2 + h_R^*(u_R^*)^2}{4} \frac{h_L^* + h_R^*}{2h^{\text{HLL}}}, \\ &= \frac{(h_L^* u_L^*)^2 + 2h_L^* u_L^* h_R^* u_R^* + (h_R^* u_R^*)^2 + h_R^* h_L^* [u^*]^2}{8h^{\text{HLL}}}, \\ &= \frac{1}{2h^{\text{HLL}}} \left( \frac{h_L^* u_L^* + h_R^* u_R^*}{2} \right)^2 + \frac{h_L^* h_R^* [u^*]^2}{8h^{\text{HLL}}}, \\ &= \frac{(h^{\text{HLL}} \hat{w}^{\text{HLL}})^2}{2h^{\text{HLL}}} + \frac{h_L^* h_R^* [u^*]^2}{8h^{\text{HLL}}}. \end{aligned} \quad (3.31)$$

En introduisant le résultat ci-dessus dans l'équation (3.29), il suit finalement que l'équation (3.25b) est équivalente à l'équation (3.30). En ajoutant  $g[z]^2/8$  de part et d'autre de l'équation (3.30), on en déduit l'équation (3.26c) ce qui termine de montrer que le système (3.25) implique le système (3.26). La réciproque est immédiate.  $\square$

Les résultats du lemme ci-dessus sont donnés avec  $\hat{w}_L$  et  $\hat{w}_R$  dans l'intérieur de  $\hat{\Omega}$  ce qui signifie formellement  $h_L \gg 0$  et  $h_R \gg 0$ . L'étude des cas  $h_L$  ou  $h_R$  proches de zéro est appelée *transition en zones sèches et zones mouillées (sèches-mouillées)* et elle demande une attention particulière [15, 164, 166]. En effet, la quantité  $\hat{u}^{\text{HLL}}$  donnée dans (3.24a) est définie à une multiplication près par  $h^{\text{HLL}}$ . Comme  $h^{\text{HLL}}$  est nul si et seulement si  $h_L = 0$  et  $h_R = 0$ , la convention admise dans [15, 164, 166] impose de définir  $\hat{u}^{\text{HLL}}$  dans  $\mathbb{R}$  par une disjonction de cas qui s'écrit

$$\hat{u}^{\text{HLL}} = \begin{cases} 0, & \text{si } h_L = 0 \text{ et } h_R = 0, \\ \frac{(hu)^{\text{HLL}} + \Delta x \hat{s}_{LR}/2\lambda}{h^{\text{HLL}}}, & \text{sinon.} \end{cases} \quad (3.32)$$

D'une manière générale, les transitions en zones sèches-mouillées seront traitées mais comme elles ne constituent pas le sujet principal de ce mémoire de thèse, les résultats intermédiaires sont donnés en considérant  $\hat{w}_L$  et  $\hat{w}_R$  dans l'intérieur de  $\hat{\Omega}$ . Dans ce cas, les définitions (3.24a) et (3.32) sont équivalentes.

Les équations (3.25) sont retenues car d'après la Proposition 3.1.1.1-i)-ii), elles sont suffisantes à la consistance et à la stabilité entropique du schéma de type Godunov associé au solveur de Riemann. À partir des ces équations, le Lemme 3.1.2.1 établi un système de trois équations (3.26) qui se composent de quatre inconnues  $h_L^*$ ,  $h_R^*$ ,  $u_L^*$ ,  $u_R^*$  et d'un paramètre  $\hat{s}_{LR}$ . Par conséquent, le système d'équations (3.26) qui définit  $w_L^*$  et  $w_R^*$  est sous-déterminé.

Le paramètre  $\hat{s}_{LR}$  intervient notamment dans la quantité  $\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + g(z_R - z_L)^2/8$  qui définit le membre de droite de l'équation quadratique (3.26c). D'après son expression, il est clair que le caractère bien posé de l'équation (3.26c) requière l'inégalité

$$\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g(z_R - z_L)^2}{8} \geq 0. \quad (3.33)$$

Bien que cette inégalité ne peut pas être établie à ce stade, il est néanmoins possible de donner deux propriétés.

**Lemme 3.1.2.2** (Propriétés de  $\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + g(z_R - z_L)^2/8$ ). *Soient  $\hat{w}_L$ ,  $\hat{w}_R$  deux états constants de l'intérieur de  $\hat{\Omega}$ , Considérons les couples  $(\hat{\eta}, \hat{G})$  et  $(\eta, G)$  respectivement donnés par (3.4), (3.5) ainsi que les quantités  $(\hat{w}^{\text{HLL}}, \hat{\eta}^{\text{HLL}})$  et  $(w^{\text{HLL}}, \eta^{\text{HLL}})$  respectivement définies dans (3.24) et dans (3.23). En supposant que la condition CFL de non interaction (3.21) est satisfaite, les assertions suivantes sont vérifiées.*

i) Si  $\hat{s}_{LR}$  est consistant avec  $\frac{1}{\Delta x} \int_{-\Delta x/2}^{\Delta x/2} (-gh \partial_x z) dx$  au sens de l'expression (3.7) alors

$$\left( \hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8}(z_R - z_L)^2 \right) \Big|_{z_L=z_R} = \eta^{\text{HLL}} - \eta(w^{\text{HLL}}) \geq 0. \quad (3.34)$$

ii) Pour toute quantité  $\hat{s}_{LR}$  donnée, l'estimation suivante est satisfaite :

$$\begin{aligned} \hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8}(z_R - z_L)^2 = & \\ & \frac{h_R h_L (u_R - u_L)^2}{4(h_L + h_R)} + \frac{g}{8}(h_R - h_L + z_R - z_L)^2 \\ & - \frac{(u_R - u_L)^2 h_R h_L (h_L u_R - h_R u_L)}{4\lambda(h_L + h_R)^2} - \frac{h_L u_L + h_R u_R}{4\lambda} \left( g(z_R - z_L) + \frac{\Delta x \hat{s}_{LR}}{h^{\text{HLL}}} \right) + \mathcal{O}\left(\frac{1}{\lambda^2}\right). \end{aligned} \quad (3.35)$$

*Démonstration.* Pour l'assertion i), si  $\hat{s}_{LR}$  est consistant alors lorsque  $z_L = z_R$ , l'égalité  $\hat{s}_{LR}|_{z_L=z_R} = 0$  est satisfaite. Dans ce cas, un calcul immédiat qui utilise la démonstration du Lemme 3.1.2.1 permet de montrer que

$$\left( \hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8}[z]^2 \right) \Big|_{[z]=0} = \eta^{\text{HLL}} - \eta(w^{\text{HLL}}).$$

Comme la condition CFL (3.21) est satisfaite, l'inégalité  $\eta^{\text{HLL}} - \eta(w^{\text{HLL}}) \geq 0$  est assurée par la définition du schéma HLL rappelée en Section 1.4.3.

Concernant l'assertion *ii*), en utilisant la notation  $[\cdot] = \cdot_R - \cdot_L$  puis en développant les expressions des quantités  $\hat{u}^{\text{HLL}}$ ,  $\hat{\eta}^{\text{HLL}}$ ,  $\hat{\eta}(\hat{w}^{\text{HLL}})$  définies dans (3.24), il suit

$$\begin{aligned}
\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8}[z]^2 &= \hat{\eta}^{\text{HLL}} - \frac{((hu)^{\text{HLL}} + \Delta x \hat{s}_{LR}/2\lambda)^2}{2h^{\text{HLL}}} - \frac{g}{2}(h^{\text{HLL}})^2 - gh^{\text{HLL}} \frac{z_L + z_R}{2} + \frac{g}{8}[z]^2, \\
&= \eta^{\text{HLL}} + g \frac{h_L z_L + h_R z_R}{2} - g \frac{[hu]z}{2\lambda} - \eta(w^{\text{HLL}}) - gh^{\text{HLL}} \frac{z_L + z_R}{2} \\
&\quad - \frac{(hu)^{\text{HLL}}}{h^{\text{HLL}}} \frac{\Delta x \hat{s}_{LR}}{2\lambda} - \frac{1}{8h^{\text{HLL}}} \left( \frac{\Delta x \hat{s}_{LR}}{\lambda} \right)^2 + \frac{g}{8}[z]^2, \\
&= \eta^{\text{HLL}} - \eta(w^{\text{HLL}}) + \frac{g}{4}[z][h] + \frac{g}{8}[z]^2 \\
&\quad - \frac{g}{4\lambda}[z](h_L u_L + h_R u_R) - \frac{(hu)^{\text{HLL}}}{h^{\text{HLL}}} \frac{\Delta x \hat{s}_{LR}}{2\lambda} - \frac{1}{8h^{\text{HLL}}} \left( \frac{\Delta x \hat{s}_{LR}}{\lambda} \right)^2.
\end{aligned} \tag{3.36}$$

L'objectif à ce stade est de faire un développement limité du terme de droite de l'expression ci-dessus en considérant que  $\lambda$  tend vers l'infini.

En notant  $q = hu$  et en développant les expressions de  $w^{\text{HLL}}$  et de  $\eta^{\text{HLL}}$  respectivement données dans (3.23a)-(3.23b), on obtient tout d'abord

$$\eta^{\text{HLL}} - \eta(w^{\text{HLL}}) = \frac{q_L^2}{4h_L} + \frac{q_R^2}{4h_R} + \frac{g(h_L^2 + h_R^2)}{4} - \frac{[G]}{2\lambda} - \frac{g}{8} \left( h_R + h_L - \frac{[q]}{\lambda} \right)^2 - \frac{((hu)^{\text{HLL}})^2}{2h^{\text{HLL}}}. \tag{3.37}$$

Ensuite, il est nécessaire de développer la quantité  $((hu)^{\text{HLL}})^2/(2h^{\text{HLL}})$  dans l'expression ci-dessus. En utilisant les définitions de  $h^{\text{HLL}}$  et de  $(hu)^{\text{HLL}}$  données par (3.23a), on a immédiatement

$$(hu)^{\text{HLL}} = \frac{q_L + q_R}{2} - \frac{1}{2\lambda}[hu^2 + gh^2/2], \tag{3.38a}$$

$$\frac{1}{2h^{\text{HLL}}} = \frac{1}{(h_L + h_R) \left( 1 - \frac{[q]}{\lambda(h_R + h_L)} \right)}. \tag{3.38b}$$

En élevant au carré l'expression (3.38a) et en interprétant la quantité  $1/(1 - \frac{[q]}{\lambda(h_R + h_L)})$  comme une série géométrique, il vient

$$((hu)^{\text{HLL}})^2 = \left( \frac{q_L + q_R}{2} \right)^2 - \frac{1}{2\lambda}(q_L + q_R)[hu^2 + \frac{g}{2}h^2] + \mathcal{O}\left(\frac{1}{\lambda^2}\right), \tag{3.39a}$$

$$\frac{1}{1 - \frac{[q]}{\lambda(h_R + h_L)}} = 1 + \frac{[q]}{\lambda(h_L + h_R)} + \mathcal{O}\left(\frac{1}{\lambda^2}\right). \tag{3.39b}$$

Par conséquent, en s'appuyant sur les deux égalités ci-dessus pour développer la quantité  $((hu)^{\text{HLL}})^2/(2h^{\text{HLL}})$  puis en introduisant le résultat dans l'expression de  $\eta^{\text{HLL}} - \eta(w^{\text{HLL}})$  donnée par (3.37), il suit

$$\begin{aligned}
\eta^{\text{HLL}} - \eta(w^{\text{HLL}}) &= \frac{q_L^2}{4h_L} + \frac{q_R^2}{4h_R} - \frac{(q_L + q_R)^2}{4(h_L + h_R)} + \frac{g}{8}[h]^2 - \frac{[G]}{2\lambda} + \frac{g}{4}(h_L + h_R) \frac{[q]}{\lambda} \\
&\quad - \frac{(q_L + q_R)^2[q]}{4\lambda(h_L + h_R)^2} + \frac{(q_L + q_R)[hu^2 + \frac{g}{2}h^2]}{2\lambda(h_L + h_R)} + \mathcal{O}\left(\frac{1}{\lambda^2}\right), \\
&= \frac{h_R h_L [u]^2}{4(h_L + h_R)} + \frac{g}{8}[h]^2 - \frac{[u]^2 h_R h_L (h_L u_R - h_R u_L)}{4\lambda(h_L + h_R)^2} + \mathcal{O}\left(\frac{1}{\lambda^2}\right).
\end{aligned}$$

En utilisant les expressions (3.38), il est également possible de développer  $(hu)^{\text{HLL}}/h^{\text{HLL}}$  et  $1/h^{\text{HLL}}$  dans l'équation (3.36) puis en introduisant l'estimation ci-dessus dans le résultat, on en déduit

$$\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8}[z]^2 = \frac{h_R h_L [u]^2}{4(h_L + h_R)} + \frac{g}{8}[h + z]^2 - \frac{[u]^2 h_R h_L (h_L u_R - h_R u_L)}{4\lambda(h_L + h_R)^2} - \frac{h_L u_L + h_R u_R}{4\lambda} \left( g[z] + \frac{\Delta x \hat{s}_{LR}}{h^{\text{HLL}}} \right) + \mathcal{O}\left(\frac{1}{\lambda^2}\right).$$

Puisque cette dernière expression correspond à l'estimation annoncée, ceci conclut la preuve.  $\square$

Le lemme ci-dessus souligne le rôle du paramètre  $\hat{s}_{LR}$  dans l'obtention de l'inégalité (3.33). Lorsque  $z_L$  et  $z_R$  sont égaux, la consistance de  $\hat{s}_{LR}$  et la condition CFL (3.21) suffisent à obtenir l'inégalité souhaitée. En revanche, et d'après l'estimation (3.35), la satisfaction de l'inégalité (3.33) dans le cas  $z_R \neq z_L$  dépend de la quantité  $\hat{s}_{LR}$ . Ainsi, cette quantité doit être choisie pour assurer la consistance d'une part et pour obtenir l'inégalité (3.33) d'autre part.

Une expression de  $\hat{s}_{LR}$  et une équation de fermeture au système sous-déterminé (3.26) sont proposées dans la prochaine section afin d'obtenir les conditions énoncées précédemment ainsi que la propriété *well-balanced*.

### 3.1.3 Schéma entropique et *well-balanced* pour le lac au repos

Le Lemme 3.1.2.1 a permis d'établir le système d'équations sous-déterminé (3.26). D'après la Proposition 3.1.1.1-ii), ces équations donnent un solveur de Riemann (3.22) qui contient intrinsèquement des conditions suffisantes à une stabilité entropique discrète pour le couple  $(\hat{\eta}, \hat{G})$  défini par (3.4). Pour totalement déterminer le solveur, il est maintenant nécessaire de compléter le système (3.26) par une équation de fermeture et par une expression de  $\hat{s}_{LR}$ . Comme le système (3.26) est quadratique, sa fermeture doit permettre l'obtention de l'inégalité (3.33) et le lemme suivant propose des expressions possibles.

**Lemme 3.1.3.1.** *Soient  $\hat{w}_L, \hat{w}_R$  deux états constants de l'intérieur de  $\hat{\Omega}$  et  $\tilde{w}(\cdot, \hat{w}_L, \hat{w}_R) : \mathbb{R} \rightarrow \mathbb{R}^2$  un solveur de Riemann de la forme (3.22). Il existe  $\lambda > 0$  suffisamment grand tel que si la condition CFL (3.21) est satisfaite et si les états intermédiaires du solveur notés  $w_L^*$  et  $w_R^*$  sont définis par le système d'équations (3.26) complété par*

$$u_L^* = u_R^*, \tag{3.40a}$$

$$\hat{s}_{LR} = \hat{s}_{LR}^{\text{WBAR}} = -gh^{\text{HLL}} \frac{z_R - z_L}{\Delta x} - \sqrt{\frac{g}{h^{\text{HLL}}} \frac{h_L u_L + h_R u_R}{h^{\text{HLL}}} \frac{(z_R - z_L)^2}{\Delta x}}, \tag{3.40b}$$

alors deux expressions pour  $w_L^*$  et  $w_R^*$  sont possibles. En considérant les quantités  $\hat{w}^{\text{HLL}}$  et  $\hat{\eta}^{\text{HLL}}$  définies dans (3.24b)-(3.24c), ces expressions sont

$$u_L^* = u_R^* = \hat{u}^{\text{HLL}}, \tag{3.41a}$$

$$h_R^* = h^{\text{HLL}} + \frac{-(z_R - z_L) \pm \sqrt{8(\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}))/g + (z_R - z_L)^2}}{2}, \tag{3.41b}$$

$$h_L^* = h^{\text{HLL}} - \frac{-(z_R - z_L) \pm \sqrt{8(\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}))/g + (z_R - z_L)^2}}{2}. \tag{3.41c}$$

Un schéma de type Godunov (3.6) défini par la donnée d'un tel solveur de Riemann

- i) est consistant avec les équations de Saint Venant (3.3),
- ii) satisfait une inégalité d'entropie discrète locale (3.14) pour le couple  $(\hat{\eta}, \hat{G})$  défini dans (3.4). Le flux numérique d'entropie  $\hat{G}$  de cette inégalité est donné par (3.15).



La notation  $\hat{s}_{LR}^{\text{WBAR}}$  dans (3.40b) signifie *Well-Balanced At Rest*. Cette convention d'écriture est introduite pour distinguer l'expression (3.40b) d'une discrétisation  $\hat{s}_{LR}$  quelconque. Avec cette précision, la démonstration du Lemme 3.1.3.1 peut maintenant être présentée.

*Démonstration du Lemme 3.1.3.1.* Il est d'abord nécessaire de montrer que les expressions (3.41) sont les solutions du système (3.26), (3.40). Dans un premier temps, on montre les égalités  $u_L^* = u_R^* = \hat{u}^{\text{HLL}}$ .

Comme  $u_L^*$  et  $u_R^*$  sont égaux d'après (3.40a), notons  $u^* = u_L^* = u_R^*$ . Dans ce cas, l'équation (3.26b) associée à (3.26a) donne

$$\begin{aligned} h^{\text{HLL}} \hat{u}^{\text{HLL}} &= \frac{h_L^* u_L^* + h_R^* u_R^*}{2}, \\ &= \frac{h_L^* + h_R^*}{2} u^*, \\ &= h^{\text{HLL}} u^*. \end{aligned} \tag{3.42}$$

Puisque  $\hat{w}_L$  et  $\hat{w}_R$  sont dans l'intérieur de  $\hat{\Omega}$  et que la condition CFL (3.21) est satisfaite, l'inégalité stricte  $h^{\text{HLL}} > 0$  est vérifiée. Ainsi, une multiplication de l'équation (3.42) par  $1/h^{\text{HLL}}$  produit le résultat attendu quant à  $u_L^*$  et  $u_R^*$ .

Ensuite, pour montrer les expressions de  $h_R^*$  et de  $h_L^*$  données dans (3.41b)-(3.41c) il faut maintenant étudier l'équation (3.26c). En utilisant l'égalité de  $u_L^*$  et  $u_R^*$  ainsi que la notation  $[\cdot] = \cdot_R - \cdot_L$ , cette équation se lit

$$\frac{g}{8} [h^* + z]^2 = \hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8} [z]^2. \tag{3.43}$$

Puisque l'équation ci-dessus est quadratique, il est nécessaire de montrer que son membre de droite est positif. Procédons par disjonction de cas en considérant tout d'abord  $[z] = 0$ . Dans ce cas, puisque  $\hat{s}_{LR}^{\text{WBAR}}$  est consistant et que la condition CFL (3.21) est satisfaite, le Lemme 3.1.2.2-*i*) donne immédiatement

$$\left( \hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8} [z]^2 \right) \Big|_{[z]=0} = \eta^{\text{HLL}} - \eta(w^{\text{HLL}}) \geq 0.$$

Dans le cas  $[z] \neq 0$ , le Lemme 3.1.2.2-*ii*) donne l'estimation suivante pour tout  $\hat{s}_{LR}$  :

$$\begin{aligned} \hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8} [z]^2 &= \\ \frac{h_R h_L [u]^2}{4(h_L + h_R)} + \frac{g}{8} [h + z]^2 - \frac{[u]^2 h_R h_L (h_L u_R - h_R u_L)}{4\lambda (h_L + h_R)^2} - \frac{h_L u_L + h_R u_R}{4\lambda} \left( g[z] + \frac{\Delta x \hat{s}_{LR}}{h^{\text{HLL}}} \right) + \mathcal{O}\left(\frac{1}{\lambda^2}\right). \end{aligned} \tag{3.44}$$

Ainsi, et comme  $h_L > 0$  et  $h_R > 0$ , si  $[u] \neq 0$  ou  $[h + z] \neq 0$  alors il existe  $\lambda$  suffisamment grand tel que l'inégalité  $\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + g[z]^2/8 \geq 0$  est satisfaite. Les cas particuliers  $[u] = 0$  et  $[h + z] = 0$  annulent les termes d'ordre zéro du développement limité (3.44). Il est alors nécessaire d'étudier les termes d'ordre un et de montrer qu'ils demeurent positifs. Pour les cas  $[u] = 0$  et  $[h + z] = 0$ , en notant  $u = u_L = u_R$  et en utilisant l'expression de  $\hat{s}_{LR}^{\text{WBAR}}$  donnée par (3.40b), on a

$$\begin{aligned} \left( g[z] + \frac{\Delta x \hat{s}_{LR}^{\text{WBAR}}}{h^{\text{HLL}}} \right) \Big|_{\substack{[u]=0, \\ [h+z]=0}} &= -\sqrt{g} \frac{(h_L + h_R)}{(h^{\text{HLL}})^{\frac{5}{2}}} u [z]^2, \\ &= -2^{\frac{5}{2}} \sqrt{g} (h_L + h_R)^{-\frac{3}{2}} u [z]^2 + \mathcal{O}\left(\frac{1}{\lambda}\right). \end{aligned}$$

En insérant l'équation ci-dessus dans le développement limité (3.44), il suit

$$\begin{aligned} \left( \hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8} [z]^2 \right) \Big|_{\substack{[u]=0, \\ [h+z]=0}} &= -\frac{(h_L + h_R)u}{4\lambda} \left( g[z] + \frac{\Delta x \hat{s}_{LR}^{\text{WBAR}}}{h^{\text{HLL}}} \Big|_{\substack{[u]=0, \\ [h+z]=0}} \right) + \mathcal{O}\left(\frac{1}{\lambda^2}\right), \\ &= \sqrt{\frac{2g}{h_L + h_R}} \frac{u^2 [z]^2}{\lambda} + \mathcal{O}\left(\frac{1}{\lambda^2}\right). \end{aligned} \tag{3.45}$$

D'après l'équation précédente, et si  $u \neq 0$ , il existe de nouveau  $\lambda > 0$  suffisamment grand tel que l'inégalité  $\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + g[z]^2/8 \geq 0$  est satisfaite pour les cas  $[h+z] = 0$  et  $u_L = u_R \neq 0$ . Avant de conclure, il faut encore étudier le cas  $[h+z] = 0$  et  $u_L = u_R = 0$  qui s'identifie à un équilibre local pour le lac au repos (3.10). Dans ce dernier cas, un calcul direct à partir de  $\hat{w}^{\text{HLL}}$  et de  $\hat{s}_{LR}^{\text{WBAR}}$  respectivement définis par (3.24b) et par (3.40b) donne

$$\begin{aligned} h^{\text{HLL}} \Big|_{\substack{u_L=u_R=0, \\ [h+z]=0}} &= \frac{h_L + h_R}{2}, \\ \hat{s}_{LR}^{\text{WBAR}} \Big|_{\substack{u_L=u_R=0, \\ [h+z]=0}} &= -g \frac{h_L + h_R}{2} \frac{z_R - z_L}{\Delta x}, \\ (h^{\text{HLL}} \hat{u}^{\text{HLL}}) \Big|_{\substack{u_L=u_R=0, \\ [h+z]=0}} &= -\frac{g}{4\lambda} [h^2] - \frac{g}{4\lambda} (h_L + h_R) [z] = -\frac{g}{4\lambda} (h_L + h_R) [h+z] = 0. \end{aligned} \quad (3.46)$$

En considérant les trois équations précédentes dans l'expression de  $\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + g[z]^2/8$ , on obtient

$$\left( \hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8} [z]^2 \right) \Big|_{\substack{u_L=u_R=0, \\ [h+z]=0}} = \frac{h_R h_L [u]^2}{4(h_L + h_R)} + \frac{g}{8} [h+z]^2 = 0. \quad (3.47)$$

Ainsi, dans tous les cas, il existe toujours  $\lambda > 0$  suffisamment grand tel que l'inégalité suivante est vérifiée :

$$\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g[z]^2}{8} \geq 0. \quad (3.48)$$

Par conséquent l'équation quadratique (3.43) est bien posée. Après avoir divisé par  $g/8$  l'équation (3.43), il est donc possible d'appliquer la racine carrée de part et d'autre et d'en déduire

$$[h^* + z] = \pm \sqrt{8(\hat{\eta}^{\text{HLL}} - \hat{\eta}(w^{\text{HLL}}))/g + [z]^2}.$$

En associant l'équation ci-dessus à l'équation (3.26a), on peut alors en déduire les expressions des quantités  $h_R^*$  et de  $h_L^*$  respectivement annoncées dans (3.41b)-(3.41c).

Concernant les assertions *i*) et *ii*), puisque  $\hat{s}_{LR}^{\text{WBAR}}$  est consistant et qu'un calcul direct montre que  $\tilde{w}(\cdot, \hat{w}, \hat{w}) = w$ , on en déduit que le solveur de Riemann satisfait la Définition 1.4.4.2-*iii*). De plus, comme les états intermédiaires du solveur sont définis par le système (3.26), le Lemme 3.1.2.1 assure que le solveur de Riemann vérifie la relation de consistance intégrale (3.12) et l'inégalité (3.13). Par conséquent, la consistance du schéma de type Godunov *i*) et la stabilité entropique *ii*) se déduisent de la Proposition 3.1.1.1 ce qui termine la preuve.  $\square$

D'après la preuve du lemme ci-dessus, la discrétisation consistante  $\hat{s}_{LR}^{\text{WBAR}}$  donnée par (3.40b) permet d'obtenir les égalités suivantes :

$$\hat{u}^{\text{HLL}} \Big|_{[z]=0} = u^{\text{HLL}}, \quad \left( \hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8} [z]^2 \right) \Big|_{[z]=0} = \eta^{\text{HLL}} - \eta(w^{\text{HLL}}).$$

Par conséquent, les expressions (3.41) dégénèrent vers (2.21) lorsque  $z_L$  et  $z_R$  sont égaux. Les solveurs de Riemann du Lemme 3.1.3.1 sont donc les extensions directes des solveurs obtenus en Section 2.3 pour le cas particulier d'une fonction  $z$  constante.

Les formulations explicites de  $w_L^*$  et  $w_R^*$  données dans (3.41) définissent entièrement un solveur de Riemann (3.22). Il est néanmoins nécessaire d'y adjoindre des procédures de limitation qui assurent la robustesse du schéma dans les transitions en zones sèches-mouillées caractérisées par  $h_L$  ou  $h_R$  proches de zéro. Par conséquent, une procédure de limitation est introduite dans le théorème suivant qui établit également la propriété *well-balanced* pour le lac au repos.

**Théorème 3.1.3.1** (Schéma de type Godunov robuste, entropique, et *well-balanced* pour le lac au repos). *Soient  $\hat{w}_L$ ,  $\hat{w}_R$  deux états constants de  $\hat{\Omega}$  et  $\tilde{w}(\cdot, \hat{w}_L, \hat{w}_R) : \mathbb{R} \rightarrow \mathbb{R}^2$  un solveur de Riemann de la forme (3.22). Supposons que  $\lambda > 0$  est tel que la condition CFL de non interaction (3.21) est*

satisfaite et tel que le système d'équations (3.26), (3.40) admette des solutions réelles. Considérons les quantités  $\hat{w}^{\text{HLL}}$  et  $\hat{\eta}^{\text{HLL}}$  définies dans (3.24b)-(3.24c) ainsi que les quantités  $(\tilde{h}_R^*, \tilde{h}_L^*)$  dans  $\mathbb{R}^2$  telles que

$$\tilde{h}_R^* = h^{\text{HLL}} + \frac{-(z_R - z_L) \pm \sqrt{8(\hat{\eta}^{\text{HLL}} - \hat{w}^{\text{HLL}})/g + (z_R - z_L)^2}}{2}, \quad (3.49a)$$

$$\tilde{h}_L^* = h^{\text{HLL}} - \frac{-(z_R - z_L) \pm \sqrt{8(\hat{\eta}^{\text{HLL}} - \hat{w}^{\text{HLL}})/g + (z_R - z_L)^2}}{2}. \quad (3.49b)$$

Si  $\hat{s}_{LR}$  vérifie

$$\Delta x \hat{s}_{LR} = \begin{cases} 0, & \text{si } h_L = 0 \text{ et } h_R = 0, \\ gh_R^2/2, & \text{si } h_R u_R = 0 \text{ et } h_L = 0 \text{ et } h_R + z_R \leq z_L, \\ -gh_L^2/2, & \text{si } h_L u_L = 0 \text{ et } h_R = 0 \text{ et } h_L + z_L \leq z_R, \\ -g(h_L + h_R)(z_R - z_L)/2, & \text{si } h_L = 0 \text{ ou } h_R = 0, \\ \Delta x \hat{s}_{LR}^{\text{WBAR}}, & \text{sinon,} \end{cases} \quad (3.50)$$

avec  $\hat{s}_{LR}^{\text{WBAR}}$  défini dans (3.40b) et si les états intermédiaires du solveur notés  $w_L^*$  et  $w_R^*$  s'écrivent

$$\begin{pmatrix} h_L^* \\ u_L^* \\ h_R^* \\ u_R^* \end{pmatrix} = \begin{cases} (0, 0, 0, 0)^T, & \text{si } h_L = 0 \text{ et } h_R = 0, \\ (0, 0, 2h^{\text{HLL}}, \hat{u}^{\text{HLL}})^T, & \text{si } h_L = 0 \text{ et } h_R > 0, \\ (2h^{\text{HLL}}, \hat{u}^{\text{HLL}}, 0, 0)^T, & \text{si } h_L > 0 \text{ et } h_R = 0, \\ \begin{pmatrix} \min(\max(\tilde{h}_L^*, 0), 2h^{\text{HLL}}) \\ \hat{u}^{\text{HLL}} \\ \min(\max(\tilde{h}_R^*, 0), 2h^{\text{HLL}}) \\ \hat{u}^{\text{HLL}} \end{pmatrix}, & \text{sinon,} \end{cases} \quad (3.51)$$

alors un schéma de type Godunov (3.6) défini par la donnée d'un tel solveur de Riemann

- i) est consistant avec les équations de Saint Venant (3.3),
- ii) préserve l'ensemble convexe  $\hat{\Omega}$ , i.e : si  $(\hat{w}_i^n)_{i \in \mathbb{Z}} \subset \hat{\Omega}$  alors,  $(\hat{w}_i^{n+1})_{i \in \mathbb{Z}} \subset \hat{\Omega}$ ,
- iii) est robuste pour les transitions dans les zones sèches-mouillées,
- iv) est well-balanced pour le lac au repos, ce qui s'écrit

$$\text{si } \forall i \in \mathbb{Z} \begin{cases} u_i^n = 0, \\ h_i^n + z_i = \text{cste}, \end{cases} \text{ alors } w_i^{n+1} = w_i^n, \forall i \in \mathbb{Z}.$$

Si de plus,  $\tilde{h}_L^* > 0$  et  $\tilde{h}_R^* > 0$  alors un schéma de type Godunov (3.6) associé à un tel solveur de Riemann satisfait une inégalité d'entropie discrète locale (3.14) pour le couple  $(\hat{\eta}, \hat{G})$  défini par (3.4). Le flux numérique d'entropie  $\hat{G}$  de cette inégalité est donné par (3.15).

Les propriétés *well-balanced* et de stabilité entropique décrites ci-dessus sont également établies dans [32] à l'aide d'un schéma de relaxation qui nécessite néanmoins la résolution d'une équation cubique. Le résultat proposé dans le Théorème 3.1.3.1 s'affranchit de cette contrainte et son originalité repose sur la résolution explicite d'une équation quadratique.

Cette résolution explicite donne deux schémas numériques distingués par le symbole  $\pm$  dans les définitions (3.51). Ces deux schémas sont entropiques, *well-balanced* pour le lac au repos et préserve l'espace  $\hat{\Omega}$  en vertu des procédures de limitations  $\min(\max(\cdot, \cdot), \cdot)$ . Cependant, si ces procédures de

limitations sont actives, le système d'équations (3.26) n'est plus nécessairement vérifié et par conséquent la stabilité entropique pourrait être localement perdue dans les transitions en zones sèches-mouillées.

Ces transitions s'opèrent également à l'aide des disjonctions de cas dans les expressions (3.50) et (3.51). D'après [164, Section 3.1.2.4], ces disjonctions sont robustes mais il est clair que leurs expressions ne sont pas continues avec le schéma défini par le Lemme 3.1.3.1.

Avec ces remarques, la preuve du Théorème 3.1.3.1 peut désormais être établie.

*Démonstration du Théorème 3.1.3.1.* En ce qui concerne *i*), il est seulement nécessaire de montrer la consistance en dehors des transitions en zones sèches. Ainsi, si  $\widetilde{h}_L^*$  et  $\widetilde{h}_R^*$  donnés par (3.49) sont tous deux strictement positifs alors les expressions (3.51) coïncident avec les expressions (3.41). Dans ce cas la consistance est montrée dans le Lemme 3.1.3.1. Si  $\widetilde{h}_L^* \leq 0$  ou  $\widetilde{h}_R^* \leq 0$ , alors les procédures de limitation  $\min(\max(\cdot, \cdot), \cdot)$  s'activent or la preuve du Lemme 2.3.3 montre que dans ce cas la relation de consistance intégrale est préservée. Par conséquent, on en déduit la consistance annoncée.

Ensuite, pour montrer la préservation du domaine  $\widehat{\Omega}$  énoncée dans *ii*), il est suffisant de montrer que  $h_L^* \geq 0$  et  $h_R^* \geq 0$  or d'après les expressions (3.51) ces inégalités sont immédiatement assurées.

Concernant l'assertion *iii*) et d'après [164, Section 3.1.2.4], les disjonctions proposées dans (3.50) pour  $\widehat{s}_{LR}$  et dans (3.51) pour les états  $w_L^*$ ,  $w_R^*$  assurent la robustesse du schéma dans le cas d'une transition en zones sèches.

Pour montrer la propriété *well-balanced* de l'assertion *iv*), il est suffisant, d'après la Proposition 3.1.1.1, de montrer que si  $\widehat{w}_L$  et  $\widehat{w}_R$  forment un équilibre local pour le lac au repos (3.10) alors  $w_L^* = w_L$  et  $w_R^* = w_R$ . En utilisant la notation  $[\cdot] = \cdot_R - \cdot_L$ , si  $\widehat{w}_L$  et  $\widehat{w}_R$  forment un équilibre local (3.10) alors  $u_L = u_R = 0$ ,  $[h + z] = 0$  puis  $\widehat{s}_{LR}$ ,  $w_L^*$  et  $w_R^*$  sont respectivement donnés par (3.40b) et par (3.41).

Comme  $u_L^* = u_R^* = \widehat{u}^{\text{HLL}}$ , le premier objectif est de montrer que  $\widehat{u}^{\text{HLL}} = 0$ . Lorsque les conditions locales du lac au repos (3.10) sont satisfaites, il a été établi dans le Lemme 3.1.3.1 que  $h^{\text{HLL}}\widehat{u}^{\text{HLL}} = 0$ . Par conséquent, il suit

$$\begin{aligned} 0 &= (h^{\text{HLL}}\widehat{u}^{\text{HLL}})\Big|_{\substack{u_L=u_R=0, \\ [h+z]=0}}, & (3.52) \\ &= h^{\text{HLL}}\Big|_{\substack{u_L=u_R=0, \\ [h+z]=0}}, \widehat{u}^{\text{HLL}}\Big|_{\substack{u_L=u_R=0, \\ [h+z]=0}}, \\ &= \frac{h_L + h_R}{2}\widehat{u}^{\text{HLL}}\Big|_{\substack{u_L=u_R=0, \\ [h+z]=0}}. \end{aligned}$$

D'après l'équation ci-dessus, lorsque les conditions locales du lac au repos (3.10) sont vérifiées  $\widehat{u}^{\text{HLL}} = 0$  et par conséquent  $u_L^* = u_R^* = 0$ .

Il reste à montrer les égalités  $h_L^* = h_L$  et  $h_R^* = h_R$ . Lorsque les conditions locales du lac au repos (3.10) sont vérifiées, la preuve du Lemme 3.1.3.1 donne dans (3.47) l'égalité  $\widehat{\eta}^{\text{HLL}} - \widehat{\eta}(\widehat{w}^{\text{HLL}}) + g[z]^2/8 = 0$ . En associant cette égalité aux expressions de  $h_R^*$  et de  $h_L^*$  respectivement données dans (3.41b) et (3.41c), il vient

$$\begin{aligned} h_R^*\Big|_{\substack{u_L=u_R=0, \\ [h+z]=0}} &= \frac{h_L + h_R - [z]}{2} = \frac{h_L + h_R + [h]}{2} = h_R, \\ h_L^*\Big|_{\substack{u_L=u_R=0, \\ [h+z]=0}} &= \frac{h_L + h_R + [z]}{2} = \frac{h_L + h_R - [h]}{2} = h_L. \end{aligned}$$

Ces deux derniers résultats terminent de montrer la propriété *well-balanced* de l'assertion *iv*).

Enfin, concernant l'inégalité d'entropie discrète, si  $\widetilde{h}_L^*$  et  $\widetilde{h}_R^*$  donnés par (3.49) sont tous deux strictement positifs alors les expressions (3.51) coïncident avec les expressions (3.41). Dans ce cas l'inégalité d'entropie discrète est une conséquence directe de la définition du solveur de Riemann et cette inégalité est démontrée dans le Lemme 3.1.3.1-*ii*) ce qui conclut la preuve.  $\square$

Les schémas du Théorème 3.1.3.1 sont obtenus à partir du système (3.26) fermé par les équations additionnelles (3.40). D'autres équations de fermeture sont maintenant proposées afin d'obtenir un schéma numérique entropique et *well-balanced* pour les équilibres mouvants.

### 3.1.4 Schéma entropique et *well-balanced* pour tous les états d'équilibres réguliers

Avec la donnée d'états intermédiaires  $w_L^*$  et  $w_R^*$  définis par le système d'équations (3.26), le Lemme 3.1.2.1 et la Proposition 3.1.1.1 assurent que le schéma de type Godunov associé au solveur de Riemann (3.22) vérifie une stabilité entropique locale pour le couple  $(\hat{\eta}, \hat{G})$  (3.4). Comme le système (3.26) est sous déterminé, il s'agit maintenant de proposer des équations de fermeture permettant de restituer la propriété *well-balanced* pour les équilibres mouvants.

Toutes les notations utilisées dans les sections 3.1.2 et 3.1.3 demeurent et en particulier la notation  $[\cdot] = \cdot_R - \cdot_L$ . Ces notations sont tout d'abord complétées par

$$\hat{B}(\hat{w}) = \frac{(hu)^2}{2h^2} + g(h+z), \quad [\hat{B}] = \hat{B}(\hat{w}_R) - \hat{B}(\hat{w}_L).$$

Avec les notations ci-dessus, les équilibres mouvants locaux (3.9) se lisent maintenant

$$h_L u_L = h_R u_R, \quad \text{et} \quad \hat{B}(\hat{w}_L) = \hat{B}(\hat{w}_R). \quad (3.53)$$

De plus, pour toute quantité  $X$ , si les conditions d'équilibre local écrites ci-dessus sont vérifiées alors  $X$  devient  $X^{\text{eq}}$ . Cette notation s'écrit

$$X|_{[q]=0, [\hat{B}]=0} = X^{\text{eq}}.$$

Puisque les équilibres mouvants locaux (3.53) sont caractérisés par  $h_L u_L = h_R u_R$ , la notation  $q^{\text{eq}}$  dans  $\mathbb{R}$  est adoptée sous la forme des égalités suivantes :

$$(h_L u_L)^{\text{eq}} = (h_R u_R)^{\text{eq}} = q^{\text{eq}}.$$

Enfin,  $\bar{h}$  désigne la moyenne arithmétique de  $h_L$  et de  $h_R$  qui s'écrit  $\bar{h} = (h_L + h_R)/2$  dans  $\mathbb{R}^+$ .

En associant ces notations à un couple  $\lambda > 0$ ,  $\Delta t > 0$  et à  $w^{\text{HLL}}$  (3.23a), il est désormais possible de définir les quantités réelles qui prennent des valeurs particulières lorsque  $\hat{w}_L$  et  $\hat{w}_R$  forment un équilibre local (3.53). Ces quantités réelles sont notées  $\zeta_{LR}$ ,  $q_{LR}$ ,  $\beta_{LR}$  et  $F_{LR}^2$  et elles sont chacune définies par

$$\zeta_{LR} = [hu]^2 + \bar{h}[\hat{B}]^2/g, \quad (3.54a)$$

$$q_{LR} = \begin{cases} 0, & \text{si } \zeta_{LR} = 0 \text{ et } [z] = 0, \\ \frac{\min(h_L u_L, h_R u_R)(g[z]\Delta t)^2}{(g[z]\Delta t)^2 + \zeta_{LR}}, & \text{sinon,} \end{cases}, \quad (3.54b)$$

$$\beta_{LR} = \begin{cases} 0, & \text{si } \zeta_{LR} = 0 \text{ et } [u] = 0 \text{ et } [h+z] = 0, \\ \frac{(\bar{h}[u])^2 + g\bar{h}[h+z]^2}{(\bar{h}[u])^2 + g\bar{h}[h+z]^2 + \zeta_{LR}}, & \text{sinon,} \end{cases}, \quad (3.54c)$$

$$F_{LR}^2 = \frac{q_{LR}^2}{gh_L h_R h^{\text{HLL}}}. \quad (3.54d)$$

Le lemme ci-après donne des propriétés satisfaites par  $\zeta_{LR}$ ,  $q_{LR}$  et  $F_{LR}^2$ .

**Lemme 3.1.4.1.** *Soient  $\hat{w}_L$ ,  $\hat{w}_R$  deux états constants de l'intérieur de  $\hat{\Omega}$  ainsi que  $\lambda > 0$  et  $\Delta t > 0$  satisfaisant la condition CFL (3.21). Considérons  $w^{\text{HLL}}$  donné dans (3.23a) ainsi que  $\zeta_{LR}$ ,  $q_{LR}$  et  $F_{LR}^2$  respectivement défini par (3.54a), (3.54b) et (3.54d). Les assertions suivantes sont vérifiées.*

- i)  $\zeta_{LR} = 0$  si et seulement si  $\hat{w}_L$ ,  $\hat{w}_R$  vérifient un équilibre local (3.53).
- ii) Si  $[z] \neq 0$  alors  $q_{LR}^{\text{eq}} = q^{\text{eq}}$ .
- iii) Si  $[z] \neq 0$  alors  $(F_{LR}^2)^{\text{eq}} = 2(q^{\text{eq}})^2/(gh_L h_R (h_L + h_R))$ .

*Démonstration.* Concernant l'assertion *i*) et comme  $\hat{w}_L, \hat{w}_R$  sont dans l'intérieur de  $\hat{\Omega}$ , l'inégalité stricte  $\bar{h} > 0$  est satisfaite. Par conséquent  $\zeta_{LR}$  est une somme de termes positifs et le résultat est immédiat.

Pour l'assertion *ii*), il s'agit de montrer une égalité dans le cas d'un équilibre local (3.53). Comme  $[z] \neq 0$ , l'utilisation de l'égalité  $\zeta_{LR}^{\text{eq}} = 0$  donne

$$q_{LR}^{\text{eq}} = \left( \frac{\min(h_L u_L, h_R u_R) (g[z] \Delta t)^2}{(g[z] \Delta t)^2 + \zeta_{LR}} \right)^{\text{eq}} = (\min(h_L u_L, h_R u_R))^{\text{eq}} = q^{\text{eq}}.$$

Enfin, pour l'assertion *iii*), en écrivant la définition de  $h^{\text{HLL}}$  donnée par (3.23a), il suit

$$(h^{\text{HLL}})^{\text{eq}} = \frac{h_L + h_R}{2} - \frac{[hu]^{\text{eq}}}{2\lambda} = \frac{h_L + h_R}{2}.$$

En utilisant l'équation ci-dessus et le résultat *ii*) dans la définition de  $F_{LR}^2$  donnée dans (3.54d), on en déduit le résultat attendu ce qui termine la preuve.  $\square$

En s'appuyant sur les définitions (3.54), une fermeture du système d'équations (3.26) est maintenant accessible. Cette fermeture est présentée dans la proposition ci-dessous qui utilise les symboles usuels  $\pm$  pour désigner  $+$  ou  $-$ . Deux symboles indépendants  $\pm_1$  et  $\pm_2$  sont nécessaires et le second est complété par  $\mp_2$  tel que si  $\pm_2$  est positif (*resp.* négatif) alors  $\mp_2$  est négatif (*resp.* positif).

**Proposition 3.1.4.1.** *Soient  $\hat{w}_L, \hat{w}_R$  deux états constants de l'intérieur de  $\hat{\Omega}$  et  $\tilde{w}(\cdot, \hat{w}_L, \hat{w}_R) : \mathbb{R} \rightarrow \mathbb{R}^2$  un solveur de Riemann de la forme (3.22). Il existe  $\lambda > 0$  suffisamment grand et  $\Delta t > 0$  suffisamment petit tels que si la condition CFL (3.21) est satisfaite et si les états intermédiaires du solveur notés  $w_L^*$  et  $w_R^*$  sont définis par le système d'équations (3.26) complété par*

$$h_R^* h_L^* (u_R^* - u_L^*)^2 = \frac{q_{LR}^2}{h_R h_L} (h_R^* - h_L^*)^2, \quad (3.55a)$$

$$\hat{s}_{LR} = \hat{s}_{LR}^{\text{FWB}} = \hat{s}_{LR}^{\text{WBAR}} + \beta_{LR} \left( \frac{g \bar{h} F_{LR}^2 (h_R - h_L)^3}{4 h_L h_R \Delta x} + \sqrt{g} \frac{h_L u_L + h_R u_R (z_R - z_L)^2}{(h^{\text{HLL}})^{3/2} \Delta x} \right), \quad (3.55b)$$

alors deux expressions pour  $h_L^*$  et  $h_R^*$  sont possibles. En considérant les quantités  $\hat{w}^{\text{HLL}}$  et  $\hat{\eta}^{\text{HLL}}$  définies dans (3.24b)-(3.24c), ces expressions sont

$$h_L^* = h^{\text{HLL}} - \frac{[h^*]}{2}, \quad (3.56a)$$

$$h_R^* = h^{\text{HLL}} + \frac{[h^*]}{2}, \quad (3.56b)$$

$$\text{avec, } [h^*] = - \frac{(z_R - z_L)}{1 + F_{LR}^2} \pm_1 \sqrt{\frac{8 \hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}})}{1 + F_{LR}^2} + \frac{(z_R - z_L)^2}{(1 + F_{LR}^2)^2}}. \quad (3.56c)$$

Si de plus,  $h_R^*$  et  $h_L^*$  définis ci-dessus sont strictement positifs alors les quantités  $u_L^*$  et  $u_R^*$  s'écrivent respectivement

$$u_L^* = \hat{u}^{\text{HLL}} \mp_2 \frac{q_{LR} [h^*]}{2 h^{\text{HLL}} \sqrt{h_L h_R}} \sqrt{\frac{h_R^*}{h_L^*}}, \quad (3.57a)$$

$$u_R^* = \hat{u}^{\text{HLL}} \pm_2 \frac{q_{LR} [h^*]}{2 h^{\text{HLL}} \sqrt{h_L h_R}} \sqrt{\frac{h_L^*}{h_R^*}}. \quad (3.57b)$$

Enfin, un schéma de type Godunov (3.6) défini par la donnée d'un tel solveur de Riemann

- i*) est consistant avec les équations de Saint Venant (3.3),
- ii*) satisfait une inégalité d'entropie discrète locale pour le couple  $(\hat{\eta}, \hat{G})$  (3.4) et dans laquelle le flux numérique d'entropie  $\hat{G}$  est donné par (3.15).

Les symboles  $\pm_1$  et  $\pm_2$  respectivement employés dans les expressions (3.56) et (3.57) peuvent s'interpréter comme des paramètres libres qui seront fixés par la suite afin de garantir la propriété *well-balanced*.

La propriété *well-balanced* sera également obtenue avec l'expression (3.55b) pour laquelle la notation  $\hat{s}_{LR}^{\text{FWB}}$  signifie *Fully Well-Balanced*. Cette convention d'écriture est introduite pour distinguer l'expression (3.55b) d'une discrétisation  $\hat{s}_{LR}$  quelconque. D'après cette expression,  $\hat{s}_{LR}^{\text{FWB}}$  est défini comme une correction de la quantité  $\hat{s}_{LR}^{\text{WBAR}}$  et cette propriété sera utilisée dans la suite cette section.

Avec ces remarques, et avant de démontrer la Proposition 3.1.4.1, il est maintenant nécessaire d'établir deux résultats intermédiaires. Le premier concerne la quantité  $\hat{s}_{LR}^{\text{FWB}}$  définie dans (3.55b)

**Lemme 3.1.4.2** (Propriétés de  $\hat{s}_{LR}^{\text{FWB}}$ ). *Soient  $\hat{w}_L, \hat{w}_R$  deux états constants de l'intérieur de  $\hat{\Omega}$ . Si la condition CFL (3.21) est satisfaite et si  $\hat{s}_{LR} = \hat{s}_{LR}^{\text{FWB}}$  dans l'expression de  $h^{\text{HLL}}\hat{u}^{\text{HLL}}$  donnée par (3.24a) alors*

$$(h^{\text{HLL}}\hat{u}^{\text{HLL}})^{\text{eq}} = \left( (hu)^{\text{HLL}} + \frac{\Delta x \hat{s}_{LR}^{\text{FWB}}}{2\lambda} \right)^{\text{eq}} = q^{\text{eq}}.$$

*Démonstration.* La preuve se mène par disjonction de trois cas qui concernent

- i)  $[z] = 0$ ,
- ii)  $[z] \neq 0$  et  $q^{\text{eq}} = 0$ ,
- iii)  $[z] \neq 0$  et  $q^{\text{eq}} \neq 0$ .

Dans le cas *i*), et d'après la définition (3.55b), il est clair que  $\hat{s}_{LR}^{\text{FWB}} = 0$ . De plus, et d'après l'hypothèse (3.11), un équilibre local défini par  $[z] = 0$  est nécessairement caractérisé par  $\hat{w}_L = \hat{w}_R$ . Par conséquent, en utilisant la définition de  $h^{\text{HLL}}\hat{u}^{\text{HLL}}$  donnée par (3.24a), il suit

$$\begin{aligned} (h^{\text{HLL}}\hat{u}^{\text{HLL}})^{\text{eq}}|_{[z]=0} &= ((hu)^{\text{HLL}})^{\text{eq}}, \\ &= \frac{(h_L u_L)^{\text{eq}} + (h_R u_R)^{\text{eq}}}{2} - \frac{[hu^2 + gh^2/2]^{\text{eq}}}{2\lambda}, \\ &= q^{\text{eq}}. \end{aligned}$$

Comme l'égalité ci-dessus correspond au résultat souhaité, il est maintenant possible d'étudier le second cas.

Le cas *ii*) est caractérisé par  $[z] \neq 0$  et  $q^{\text{eq}} = 0$ . Dans ce cas, un calcul direct montre que l'équilibre local mouvant (3.53) dégénère vers un équilibre local du lac au repos qui s'écrit  $u_L = u_R = 0$  et  $[h + z] = 0$ . Selon la définition de  $\beta_{LR}$  donnée par (3.54c), si les conditions d'un équilibre local du lac au repos sont satisfaites alors  $\beta_{LR} = 0$ . Ainsi, d'après l'équation (3.55b), on a dans ce cas

$$(\hat{s}_{LR}^{\text{FWB}})^{\text{eq}}|_{q^{\text{eq}}=0} = (\hat{s}_{LR}^{\text{WBAR}})^{\text{eq}}|_{q^{\text{eq}}=0} = -g \frac{h_L + h_R}{2} \frac{z_R - z_L}{\Delta x}.$$

Par conséquent, en utilisant le résultat (3.52) du Théorème 3.1.3.1-*iv*), on en déduit  $h^{\text{HLL}}\hat{u}^{\text{HLL}} = 0$  et comme  $h^{\text{HLL}} > 0$ , cette égalité conduit au résultat attendu.

Considérons maintenant le cas *iii*) défini par  $[z] \neq 0$  et  $q^{\text{eq}} \neq 0$ . Puisque  $\hat{w}_L$  et  $\hat{w}_R$  définissent un équilibre local mouvant (3.53) avec  $q^{\text{eq}}$  non nul, des calculs usuels qui s'appuient sur (3.54c) et sur (3.53) montrent respectivement que

$$\beta_{LR}^{\text{eq}}|_{q^{\text{eq}} \neq 0} = 1 \quad \text{et} \quad -\frac{(q^{\text{eq}})^2}{2h_L^2 h_R^2} [h^2] + g[h + z] = 0.$$

En utilisant les deux résultats ci-dessus ainsi que l'égalité  $(F_{LR}^2)^{\text{eq}} = 2(q^{\text{eq}})^2 / (g h_L h_R (h_L + h_R))$  donnée par le Lemme 3.1.4.1-*iii*), il suit

$$\begin{aligned} (\Delta x \hat{s}_{LR}^{\text{FWB}})^{\text{eq}}|_{q^{\text{eq}} \neq 0} &= -g \frac{h_L + h_R}{2} [z] + \frac{g}{8} [h]^3 \frac{h_L + h_R}{h_L h_R} F_{LR}^2|_{q^{\text{eq}} \neq 0}, \\ &= -g \frac{h_L + h_R}{2} [z] + \frac{[h]^2}{4} \frac{[h^2] (q^{\text{eq}})^2}{h_L^2 h_R^2 (h_L + h_R)}, \\ &= -g \frac{h_L + h_R}{2} [z] + \frac{g[h]^2}{2} \frac{[h + z]}{h_L + h_R}. \end{aligned}$$

En insérant l'équation ci-dessus dans la définition de  $h^{\text{HLL}}\hat{u}^{\text{HLL}}$  donnée par (3.24a), on en déduit

$$\begin{aligned}
& (h^{\text{HLL}}\hat{u}^{\text{HLL}})^{\text{eq}}\Big|_{q^{\text{eq}}\neq 0} \\
&= \left(h^{\text{HLL}}u^{\text{HLL}} + \frac{\Delta x \hat{s}_{LR}^{\text{FWB}}}{2\lambda}\right)^{\text{eq}}\Big|_{q^{\text{eq}}\neq 0}, \\
&= q^{\text{eq}} - \frac{1}{2(h_L + h_R)\lambda} \left( -\frac{(q^{\text{eq}})^2[h^2]}{h_R h_L} + \frac{g}{2}(h_L + h_R)^2[h] \right) + \frac{g}{4\lambda} \frac{-(h_L + h_R)^2[z] + [h]^2[h + z]}{h_L + h_R}, \\
&= q^{\text{eq}} - \frac{g[h + z]}{4\lambda(h_L + h_R)} \left( -4h_L h_R + (h_L + h_R)^2 - [h]^2 \right), \\
&= q^{\text{eq}}.
\end{aligned}$$

L'égalité ci-dessus termine la preuve.  $\square$

Le second résultat intermédiaire nécessaire à la preuve de la Proposition 3.1.4.1 repose sur l'obtention d'une inégalité large qui conditionne l'existence de réels  $h_L^*$  et  $h_R^*$  satisfaisant le système d'équations (3.26), (3.55).

**Lemme 3.1.4.3.** *Soient  $\hat{w}_L, \hat{w}_R$  deux états constants de l'intérieur de  $\hat{\Omega}$ . Il existe  $\lambda > 0$  suffisamment grand et  $\Delta t > 0$  suffisamment petit tels que si la condition CFL (3.21) est satisfaite et si  $\hat{s}_{LR} = \hat{s}_{LR}^{\text{FWB}}$  dans l'expression de  $h^{\text{HLL}}\hat{u}^{\text{HLL}}$  donnée par (3.24a) alors, en considérant les quantités  $\hat{w}^{\text{HLL}}, \hat{\eta}^{\text{HLL}}$  et  $F_{LR}^2$  respectivement définies dans (3.24b), (3.24c) et dans (3.54d), l'inégalité suivante est satisfaite :*

$$\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8}[z]^2 - \frac{g[z]^2}{8} \frac{F_{LR}^2}{1 + F_{LR}^2} \geq 0. \quad (3.58)$$

*Démonstration.* La preuve est conduite par disjonction de trois cas respectivement donnés par

- i)  $[z] = 0$ ,
- ii)  $[z] \neq 0$  et  $\zeta_{LR} = 0$ ,
- iii)  $[z] \neq 0$  et  $\zeta_{LR} \neq 0$ .

En considérant tout d'abord le cas *i*) défini par  $[z] = 0$  et en utilisant d'une part la définition (3.54b), on a  $q_{LR} = 0$  et en utilisant d'autre part l'équation (3.55b), on obtient  $\hat{s}_{LR}^{\text{FWB}} = 0$ . Par conséquent, dans le cas où  $z_L$  et  $z_R$  sont égaux, on en déduit que  $F_{LR}^2 = 0$  et que  $\hat{u}^{\text{HLL}} = u^{\text{HLL}}$ . En s'appuyant sur ces deux dernières égalités un calcul analogue à (3.36) donne

$$\begin{aligned}
& \left( \hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8}[z]^2 - \frac{g[z]^2}{8} \frac{F_{LR}^2}{1 + F_{LR}^2} \right)\Big|_{[z]=0} \\
&= \eta^{\text{HLL}} - \frac{(h^{\text{HLL}}\hat{u}^{\text{HLL}})^2}{2h^{\text{HLL}}}\Big|_{[z]=0} - \frac{g}{2}(h^{\text{HLL}})^2 + \left( \frac{g}{4}[z]([h] - \frac{h_L u_L + h_R u_R}{\lambda}) \right)\Big|_{[z]=0} \\
&= \eta^{\text{HLL}} - \eta(w^{\text{HLL}}).
\end{aligned}$$

Comme la condition CFL (3.21) est satisfaite, l'inégalité large  $\eta^{\text{HLL}} - \eta(w^{\text{HLL}}) \geq 0$  est assurée en vertu des propriétés du solveur HLL détaillées en Section 1.4.3.

Le cas *ii*) est caractérisé par  $[z] \neq 0$  et  $\zeta_{LR} = 0$ . Comme  $\zeta_{LR}$  est nul,  $\hat{w}_L$  et  $\hat{w}_R$  définissent un état d'équilibre mouvant local (3.53) et dans ce cas, un calcul direct montre que  $\hat{G}(\hat{w}_R) - \hat{G}(\hat{w}_L) = [\hat{B}]q^{\text{eq}} = 0$ . D'après le Lemme 3.1.4.2, l'égalité  $(h^{\text{HLL}}\hat{u}^{\text{HLL}})^{\text{eq}} = q^{\text{eq}}$  est également vérifiée. Par conséquent, il suit

$$\begin{aligned}
(\hat{\eta}^{\text{HLL}})^{\text{eq}} - \hat{\eta}(\hat{w}^{\text{HLL}})^{\text{eq}} &= \frac{\eta(w_R)^{\text{eq}} + \eta(w_L)^{\text{eq}}}{2} + \frac{g}{2}(h_L z_L + h_R z_R) - \frac{\hat{G}(\hat{w}_R)^{\text{eq}} - \hat{G}(\hat{w}_L)^{\text{eq}}}{2\lambda} \\
&\quad - \left( \frac{(h^{\text{HLL}}\hat{u}^{\text{HLL}})^2}{2h^{\text{HLL}}} \right)^{\text{eq}} - \left( \frac{g(h^{\text{HLL}})^2}{2} \right)^{\text{eq}} - g(h^{\text{HLL}})^{\text{eq}} \frac{z_L + z_R}{2},
\end{aligned}$$



$$\begin{aligned}
&= \frac{(q^{\text{eq}})^2}{4} \left( \frac{1}{h_L} + \frac{1}{h_R} \right) + \frac{g}{4} (h_L^2 + h_R^2) + \frac{g}{2} (h_L z_L + h_R z_R) \\
&\quad - \frac{(q^{\text{eq}})^2}{h_L + h_R} - \frac{g(h_L + h_R)^2}{8} - \frac{g}{4} (h_L + h_R) (z_L + z_R). \\
&= \frac{(q^{\text{eq}})^2 h_L h_R}{4(h_L + h_R)} \left[ \frac{1}{h} \right]^2 + \frac{g}{8} [h]^2 + \frac{g}{4} [z][h], \\
&= \frac{g}{8} (1 + (F_{LR}^2)^{\text{eq}}) [h]^2 + \frac{g}{4} [z][h].
\end{aligned}$$

En utilisant l'équation ci-dessus, on obtient

$$\begin{aligned}
\left( \hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8} [z]^2 - \frac{g[z]^2}{8} \frac{F_{LR}^2}{1 + F_{LR}^2} \right)^{\text{eq}} &= \frac{g}{8} (1 + (F_{LR}^2)^{\text{eq}}) [h]^2 + \frac{g}{4} [h][z] + \frac{g}{8} \frac{[z]^2}{1 + (F_{LR}^2)^{\text{eq}}}, \\
&= \frac{g}{8} \left( \sqrt{1 + (F_{LR}^2)^{\text{eq}}} [h] + \frac{[z]}{\sqrt{1 + (F_{LR}^2)^{\text{eq}}}} \right)^2 \geq 0.
\end{aligned}$$

Finalemnt, le dernier cas *iii*) est celui de  $[z]$  non nul et pour des couples  $\hat{w}_L$  et  $\hat{w}_R$  qui ne définissent pas un état d'équilibre mouvant local (3.53). Ce dernier cas s'écrit  $[z] \neq 0$  et  $\zeta_{LR} \neq 0$  et il a été établi dans le Lemme 3.1.2.2-*ii*) l'estimation suivante :

$$\begin{aligned}
\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8} [z]^2 &= \\
\frac{h_R h_L [u]^2}{4(h_L + h_R)} + \frac{g}{8} [h + z]^2 - \frac{[u]^2 h_R h_L (h_L u_R - h_R u_L)}{4\lambda (h_L + h_R)^2} - \frac{h_L u_L + h_R u_R}{4\lambda} \left( g[z] + \frac{\Delta x \hat{s}_{LR}}{h^{\text{HLL}}} \right) + \mathcal{O} \left( \frac{1}{\lambda^2} \right). & \quad (3.59)
\end{aligned}$$

Avec cette estimation et tant que  $[u]$  ou  $[h + z]$  est différent de zéro, il existe  $\lambda > 0$  suffisamment grand tel que l'inégalité stricte suivante est satisfaite :

$$\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g[z]^2}{8} > 0. \quad (3.60)$$

Dans les cas particuliers où  $[u]$  et  $[h + z]$  sont nuls, les équations suivantes sont vérifiées :

$$\beta_{LR} \Big|_{\substack{[u]=0 \\ [h+z]=0}} = 0, \quad \hat{s}_{LR}^{\text{FWB}} \Big|_{\beta_{LR}=0} = \hat{s}_{LR}^{\text{WBAR}}.$$

Ainsi, lorsque  $[u] = 0$  et  $[h + z] = 0$ , la définition de  $\beta_{LR}$  donnée par (3.54c) fait basculer  $\hat{s}_{LR}^{\text{FWB}}$  vers  $\hat{s}_{LR}^{\text{WBAR}}$ . Par conséquent, l'existence de l'inégalité stricte (3.60) pour ces cas particuliers relève de la démonstration du Lemme 3.1.3.1 qui concerne l'estimation (3.59) avec  $\hat{s}_{LR} = \hat{s}_{LR}^{\text{WBAR}}$ . Or, d'après les calculs (3.45) et (3.46) détaillés dans ce lemme, l'inégalité stricte (3.60) peut être maintenue sauf dans le cas où  $\hat{w}_L$  et  $\hat{w}_R$  sont à l'équilibre local pour le lac au repos (3.10). Comme le lac au repos s'inscrit dans les équilibres mouvants, il a déjà été traité dans le cas *ii*). Par conséquent, l'inégalité stricte (3.60) peut toujours être assurée pour ce cas *iii*). De plus, d'après les définitions (3.54b), (3.54d) et puisque  $\zeta_{LR} \neq 0$ ,  $F_{LR}^2$  est proportionnel à  $\Delta t^4$ . Ainsi, si l'inégalité stricte (3.60) est vérifiée, l'inégalité (3.58) s'obtient en choisissant  $\Delta t > 0$  tel que

$$\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8} [z]^2 \geq \frac{g[z]^2}{8} \frac{F_{LR}^2}{1 + F_{LR}^2}.$$

On en déduit ainsi le résultat attendu pour ce troisième cas d'étude ce qui termine la démonstration.  $\square$

Avec les deux lemmes ci-dessus, il est maintenant possible de procéder à la démonstration de la Proposition 3.1.4.1.

*Démonstration de la Proposition 3.1.4.1.* Il est d'abord nécessaire de montrer que les expressions (3.56)-(3.57) sont solutions du système (3.26), (3.55). En insérant l'équation (3.55a) dans l'équation (3.26c) et en utilisant la définition de  $F_{LR}^2$  donnée dans (3.54d), on en déduit que  $h_L^*$  et  $h_R^*$  vérifient les deux équations

$$\frac{h_L^* + h_R^*}{2} = h^{\text{HLL}}, \quad (3.61a)$$

$$\frac{g}{8}(1 + F_{LR}^2)[h^*]^2 + \frac{g}{4}[z][h^*] = \hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}). \quad (3.61b)$$

Pour déterminer  $h_L^*$  et  $h_R^*$ , il faut d'abord montrer que l'équation quadratique ci-dessus est bien posée. En divisant cette équation quadratique (3.61b) de part et d'autre par  $g/8(1+F_{LR}^2)$  puis en ajoutant  $[z]^2/(1+F_{LR}^2)^2$ , on obtient

$$\begin{aligned} \left([h^*] + \frac{[z]}{1 + F_{LR}^2}\right)^2 &= \frac{8\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}})}{g(1 + F_{LR}^2)} + \frac{[z]^2}{(1 + F_{LR}^2)^2}, \\ &= \frac{8}{g(1 + F_{LR}^2)} \left(\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8}[z]^2 - \frac{g[z]^2}{8} \frac{F_{LR}^2}{1 + F_{LR}^2}\right). \end{aligned}$$

D'après le Lemme 3.1.4.3, si  $\lambda$  (*resp.*  $\Delta t$ ) est suffisamment grand (*resp.* petit) alors le terme de droite de l'équation ci-dessus est positif. Par conséquent, si  $\lambda$  et  $\Delta t$  sont convenablement choisis, l'équation précédente est bien posée. On peut donc appliquer la racine carrée de part et d'autre pour en déduire l'expression de  $[h^*]$  donnée par (3.56c). Par suite, en couplant  $[h^*] = h_R^* - h_L^*$  à l'équation (3.61a), on en déduit les expressions de  $h_L^*$  et de  $h_R^*$  respectivement données dans (3.56a) et dans (3.56b).

Comme les quantités  $h_R^*$  et  $h_L^*$  sont désormais connues et comme elles sont supposées strictement positives, l'équation (3.55a) se réécrit

$$[u^*] = \pm \frac{q_{LR}}{\sqrt{h_L h_R}} \frac{[h^*]}{\sqrt{h_L^* h_R^*}}. \quad (3.62)$$

En utilisant l'équation ci-dessus ainsi que l'équation (3.61a) dans l'équation  $(h_L^* u_L^* + h_R^* u_R^*)/2 = h^{\text{HLL}} \hat{u}^{\text{HLL}}$ , il suit

$$\begin{aligned} h^{\text{HLL}} \hat{u}^{\text{HLL}} &= \frac{h_L^* u_L^* + h_R^* u_R^*}{2}, \\ &= \frac{h_L^* + h_R^*}{2} \frac{u_L^* + u_R^*}{2} + \frac{[h^*][u^*]}{4}, \\ &= h^{\text{HLL}} \frac{u_L^* + u_R^*}{2} \pm \frac{q_{LR}}{4\sqrt{h_L h_R}} \frac{[h^*]^2}{\sqrt{h_L^* h_R^*}}. \end{aligned} \quad (3.63)$$

L'équation ci-dessus associée aux égalités (3.62) et (3.61a) donne

$$\begin{aligned} u_R^* &= \hat{u}^{\text{HLL}} \pm \frac{q_{LR}[h^*]}{2\sqrt{h_L h_R h_L^* h_R^*}} \mp \frac{q_{LR}[h^*]^2}{4h^{\text{HLL}}\sqrt{h_L h_R}\sqrt{h_R^* h_L^*}}, \\ &= \hat{u}^{\text{HLL}} \pm \frac{q_{LR}[h^*]}{2\sqrt{h_L h_R h_L^* h_R^*}} \left(1 - \frac{[h^*]}{2h^{\text{HLL}}}\right), \\ &= \hat{u}^{\text{HLL}} \pm \frac{q_{LR}[h^*]}{2\sqrt{h_L h_R h_L^* h_R^*}} \frac{h_L^*}{h^{\text{HLL}}}. \end{aligned}$$

Cette dernière égalité permet de déduire le résultat (3.57b) quant à  $u_R^*$ . Puisque l'expression de  $u_L^*$  (3.57a) se déduit d'un calcul analogue ceci termine de montrer les expressions (3.56)-(3.57).

Concernant les assertions *i*) et *ii*), puisque  $\mathfrak{s}_{LR}^{\text{FWB}}$  est consistant et qu'un calcul montre que  $\hat{w}(\cdot, \hat{w}, \hat{w}) = w$ , on en déduit que le solveur de Riemann satisfait la Définition 1.4.4.2-*iii*). De plus, comme  $w_L^*$  et  $w_R^*$  sont définis par le système (3.26), le Lemme 3.1.2.1 assure que le solveur de Riemann vérifie la relation de consistance intégrale (3.12) et l'inégalité (3.13). Ainsi, d'après la Proposition

3.1.1.1-*i*)-*ii*), on en déduit le résultat annoncé quant à la consistance *i*) et quant à la stabilité entropique *ii*). Ce dernier argument conclut la preuve.  $\square$

L'introduction de l'équation (3.55a) permet de décomposer les équations qui définissent  $w_L^*$  et  $w_R^*$  en deux sous-systèmes d'équations respectivement donnés par (3.61) et par (3.55a), (3.63). Chacun de ces sous systèmes est constitué d'une équation linéaire et d'une équation quadratique.

Si la condition implicite (3.58) qui restreint le couple  $(\lambda, \Delta t)$  est satisfaite alors la première équation quadratique (3.61b) est bien posée. Le caractère bien posé de la seconde équation quadratique (3.55a) repose sur l'hypothèse que les quantités  $h_L^*$  et  $h_R^*$  données par (3.56) sont strictement positives. Si cette hypothèse n'est pas vérifiée, les inégalités  $h_L^* > 0$  et  $h_R^* > 0$  peuvent être imposées avec une technique de limitation additionnelle. Par conséquent, le système d'équations (3.26), (3.55) admet toujours des solutions.

D'après les expressions (3.56)-(3.57), ces solutions ne sont pas uniques et elles dépendent des choix retenus pour les symboles  $\pm_1$  et  $\pm_2$ . La sélection d'une solution s'opère selon la propriété *well-balanced*.

Pour satisfaire cette propriété, il est nécessaire de choisir  $\pm_2$  négatif (*ipso facto*  $\mp_2$  positif) et un calcul direct montre que le choix de  $\pm_1$  doit se faire selon une condition obtenue lorsque  $\hat{w}_L$  et  $\hat{w}_R$  forment un équilibre local (3.53). Cette condition s'écrit formellement

$$\pm_1^{\text{eq}} = \text{sign}((1 + (F_{LR}^2)^{\text{eq}})[h] + [z]).$$

Puisque l'égalité formelle ci-dessus doit être satisfaite à l'équilibre local (3.53), une expression simple de  $\pm_1$  est donnée par

$$\pm_1 = \text{sign}((1 + (F_{LR}^2))[h] + [z]).$$

Ce choix est retenu dans l'égalité (3.65) mais d'autres expressions demeurent possibles.

En considérant la propriété *well-balanced* ainsi énoncée, le théorème qui suit donne des expressions de  $(h_\alpha^*, u_\alpha^*)_{\alpha \in \{L, R\}}$  complétées par une procédure de limitation.

**Théorème 3.1.4.1** (Schéma de type Godunov robuste, entropique, et *well-balanced* pour tous les équilibres réguliers). *Soient  $\hat{w}_L, \hat{w}_R$  deux états constants de  $\hat{\Omega}$  et  $\tilde{w}(\cdot, \hat{w}_L, \hat{w}_R) : \mathbb{R} \rightarrow \mathbb{R}^2$  un solveur de Riemann de la forme (3.22). Supposons que  $\lambda > 0$  et  $\Delta t > 0$  sont tels que la condition CFL (3.21) est satisfaite et tel que le système d'équations (3.26), (3.55) admet des solutions réelles. Considérons les quantités  $\hat{w}^{\text{HLL}}, \hat{\eta}^{\text{HLL}}, q_{LR}$  et  $F_{LR}^2$  respectivement définies dans (3.24b), (3.24c) (3.54b) et (3.54d) ainsi que les quantités  $(\tilde{h}_L^*, \tilde{u}_L^*, \tilde{h}_R^*, \tilde{u}_R^*)$  dans  $\mathbb{R}^4$  telle que*

$$\tilde{h}_L^* = h^{\text{HLL}} - \frac{[h^*]}{2}, \quad (3.64a)$$

$$\tilde{u}_L^* = \hat{u}^{\text{HLL}} + \frac{q_{LR}[h^*]}{2h^{\text{HLL}}\sqrt{h_L h_R}} \sqrt{\frac{h_R^*}{h_L^*}}, \quad (3.64b)$$

$$\tilde{h}_R^* = h^{\text{HLL}} + \frac{[h^*]}{2}, \quad (3.64c)$$

$$\tilde{u}_R^* = \hat{u}^{\text{HLL}} - \frac{q_{LR}[h^*]}{2h^{\text{HLL}}\sqrt{h_L h_R}} \sqrt{\frac{h_L^*}{h_R^*}}, \quad (3.64d)$$

avec

$$[h^*] = -\frac{(z_R - z_L)}{1 + F_{LR}^2} + \text{sign}((1 + F_{LR}^2)[h] + [z]) \sqrt{\frac{8\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}})}{g} + \frac{(z_R - z_L)^2}{(1 + F_{LR}^2)^2}}. \quad (3.65)$$

Considérons ensuite  $\hat{s}_{LR}$  sous la forme

$$\Delta x \hat{s}_{LR} = \begin{cases} 0, & \text{si } h_L = 0 \text{ et } h_R = 0, \\ gh_R^2/2, & \text{si } h_R u_R = 0 \text{ et } h_L = 0 \text{ et } h_R + z_R \leq z_L, \\ -gh_L^2/2, & \text{si } h_L u_L = 0 \text{ et } h_R = 0 \text{ et } h_L + z_L \leq z_R, \\ -g(h_L + h_R)(z_R - z_L)/2, & \text{si } h_L = 0 \text{ ou } h_R = 0, \\ \Delta x \hat{s}_{LR}^{\text{FWB}}, & \text{sinon,} \end{cases} \quad (3.66)$$

avec  $\hat{s}_{LR}^{\text{FWB}}$  défini par (3.55b). Pour  $\varepsilon > 0$ , considérons enfin que les états intermédiaires du solveur notés  $w_L^*$  et  $w_R^*$  s'écrivent

$$\begin{pmatrix} h_L^* \\ u_L^* \\ h_R^* \\ u_R^* \end{pmatrix} = \begin{cases} (0, 0, 0, 0)^T, & \text{si } h_L = 0 \text{ et } h_R = 0, \\ (0, 0, 2h^{\text{HLL}}, \hat{u}^{\text{HLL}})^T, & \text{si } h_L = 0 \text{ et } h_R > 0, \\ (2h^{\text{HLL}}, \hat{u}^{\text{HLL}}, 0, 0)^T, & \text{si } h_L > 0 \text{ et } h_R = 0, \\ \left( \begin{array}{c} \min \left( \max(\tilde{h}_L^*, \varepsilon), 2h^{\text{HLL}} - \varepsilon \right) \\ \tilde{u}_L^* \\ \min \left( \max(\tilde{h}_R^*, \varepsilon), 2h^{\text{HLL}} - \varepsilon \right) \\ \tilde{u}_R^* \end{array} \right), & \text{sinon.} \end{cases} \quad (3.67)$$

Un schéma de type Godunov (3.6) défini par la donnée d'un tel solveur de Riemann

- i) est consistant avec les équations de Saint Venant (3.3),
- ii) préserve l'ensemble convexe  $\hat{\Omega}$ ,
- iii) est robuste pour les transitions en zones sèches-mouillées,
- iv) est *well-balanced* pour toutes les solutions stationnaires régulières de la forme (1.13).

Si de plus,  $\tilde{h}_L^* > 0$  et  $\tilde{h}_R^* > 0$  alors un schéma de type Godunov (3.6) associé à un tel solveur de Riemann satisfait une inégalité d'entropie discrète locale (3.14) pour le couple  $(\hat{\eta}, \hat{G})$  donné par (3.4). Le flux numérique d'entropie  $\hat{G}$  de cette inégalité est donné par (3.15).

*Démonstration.* En ce qui concerne *i*), il est seulement nécessaire de montrer la consistance en dehors des transitions en zones sèches-mouillées. Si  $\tilde{h}_L^*$  et  $\tilde{h}_R^*$  respectivement donnés par (3.64a), (3.64c) sont tous deux strictement positifs alors les expressions (3.67) coïncident avec les expressions (3.64). Dans ce cas, la consistance est montrée dans la Proposition 3.1.4.1-*i*). Si  $\tilde{h}_L^* \leq 0$  ou  $\tilde{h}_R^* \leq 0$ , alors les procédures de limitation  $\min(\max(\cdot, \cdot), \cdot)$  s'activent et un calcul analogue à celui de la preuve du Lemme 2.3.3-*i*) montre que ces procédures préservent la relation de consistance intégrale pour tout  $\varepsilon > 0$  suffisamment petit. Par conséquent, on en déduit la consistance annoncée dans l'assertion *i*).

Pour montrer l'assertion *ii*) relative à la préservation du domaine  $\hat{\Omega}$ , il est suffisant de montrer que  $h_L^* \geq 0$  et  $h_R^* \geq 0$  or d'après les expressions (3.67) ces inégalités sont immédiatement assurées pour  $\varepsilon > 0$  suffisamment petit.

Concernant *iii*) et d'après [164, Section 3.1.2.4], les disjonctions proposées dans (3.66) pour  $\hat{s}_{LR}$  et dans (3.67) pour les états  $w_L^*$ ,  $w_R^*$  assurent la robustesse du schéma dans le cas d'une transition en zones sèches-mouillées.

Pour montrer la propriété *well-balanced* de l'assertion *iv*), il est suffisant, d'après la Proposition 3.1.1.1, de montrer que si  $\hat{w}_L$  et  $\hat{w}_R$  forment un équilibre local (3.53) alors  $w_L^* = w_L$  et  $w_R^* = w_R$ . Dans le cas d'un tel équilibre local,  $\hat{s}_{LR}$ ,  $w_L^*$  et  $w_R^*$  sont donnés par (3.55b), (3.64) et la preuve du Lemme 3.1.4.2 a été établie que

$$(h^{\text{HLL}})^{\text{eq}} (\hat{u}^{\text{HLL}})^{\text{eq}} = \frac{h_L + h_R}{2} (\hat{u}^{\text{HLL}})^{\text{eq}} = q^{\text{eq}}.$$

Avec la notation  $[\cdot] = \cdot_R - \cdot_L$  et comme cela est détaillé dans la démonstration du Lemme 3.1.4.3-ii), l'équation ci dessus permet de déduire

$$\begin{aligned} \frac{g}{8} \frac{1}{1 + (F_{LR}^2)^{\text{eq}}} \left(1 + (F_{LR}^2)^{\text{eq}}[h] + [z]\right)^2 &= \left(\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8}[z]^2 - \frac{g[z]^2}{8} \frac{F_{LR}^2}{1 + F_{LR}^2}\right)^{\text{eq}}, \\ &= \frac{(\hat{\eta}^{\text{HLL}})^{\text{eq}} - \hat{\eta}(\hat{w}^{\text{HLL}})^{\text{eq}}}{1 + (F_{LR}^2)^{\text{eq}}} + \frac{g}{8} \frac{[z]}{1 + (F_{LR}^2)^{\text{eq}}}. \end{aligned}$$

En introduisant cette donnée dans l'expression (3.65), on en déduit que  $[h^*]^{\text{eq}}$  vérifie

$$\begin{aligned} [h^*]^{\text{eq}} &= \\ &= -\frac{[z]}{1 + (F_{LR}^2)^{\text{eq}}} + \left(\text{sign}((1 + (F_{LR}^2)^{\text{eq}})[h] + [z])\right)^{\text{eq}} \sqrt{\frac{8}{g} \frac{(\hat{\eta}^{\text{HLL}})^{\text{eq}} - \hat{\eta}(\hat{w}^{\text{HLL}})^{\text{eq}}}{1 + (F_{LR}^2)^{\text{eq}}} + \frac{[z]^2}{(1 + (F_{LR}^2)^{\text{eq}})^2}}, \\ &= -\frac{[z]}{1 + (F_{LR}^2)^{\text{eq}}} + \text{sign}((1 + (F_{LR}^2)^{\text{eq}})[h] + [z]) \frac{|((1 + (F_{LR}^2)^{\text{eq}})[h] + [z])|}{|1 + (F_{LR}^2)^{\text{eq}}|}, \\ &= \frac{-[z] + (1 + (F_{LR}^2)^{\text{eq}})[h] + [z]}{1 + (F_{LR}^2)^{\text{eq}}}, \\ &= [h]. \end{aligned} \tag{3.68}$$

En associant l'égalité ci-dessus aux définitions (3.64a), (3.64c), on déduit que  $(h_\alpha^*)^{\text{eq}} = h_\alpha$  pour tout  $\alpha$  dans  $\{L, R\}$ .

Il faut maintenant montrer les égalités

$$q^{\text{eq}} = (h_R^*)^{\text{eq}}(u_R^*)^{\text{eq}} = (h_L^*)^{\text{eq}}(u_L^*)^{\text{eq}}.$$

Comme le résultat  $(h_\alpha^*)^{\text{eq}} = h_\alpha$  vient d'être montré, les égalités ci dessus sont équivalentes à  $(u_\alpha^*)^{\text{eq}} = q^{\text{eq}}/h_\alpha$  pour tout  $\alpha$  dans  $\{L, R\}$ . D'après le Lemme 3.1.4.1,  $q_{LR}^{\text{eq}} = q^{\text{eq}}$  or en associant ce résultat à l'égalité entre  $(h_\alpha^*)^{\text{eq}}$  et  $h_\alpha$  précédemment montrée, il suit

$$\begin{aligned} (u_L^*)^{\text{eq}} &= (\hat{w}^{\text{HLL}})^{\text{eq}} + \frac{q_{LR}^{\text{eq}}[h^*]^{\text{eq}}}{2(h^{\text{HLL}})^{\text{eq}}\sqrt{h_L h_R}} \sqrt{\frac{(h_R^*)^{\text{eq}}}{(h_L^*)^{\text{eq}}}}, \\ &= \frac{2q^{\text{eq}}}{h_L + h_R} + \frac{q^{\text{eq}}[h]}{(h_L + h_R)h_L}, \\ &= q^{\text{eq}} \frac{2h_L + [h]}{(h_L + h_R)h_L}, \\ &= \frac{q^{\text{eq}}}{h_L} = u_L. \end{aligned}$$

Comme un calcul analogue montre que  $(u_R^*)^{\text{eq}} = u_R$ , on en déduit la propriété *well-balanced* annoncée.

Enfin, concernant l'inégalité d'entropie discrète, si  $\tilde{h}_L^*$  et  $\tilde{h}_R^*$  respectivement donnés par (3.64a) et par (3.64c) sont tous deux strictement positifs alors les expressions (3.67) coïncident avec les expressions (3.64). Dans ce cas, l'inégalité d'entropie discrète est une conséquence directe de la définition du solveur de Riemann et cette inégalité est démontrée dans la Proposition 3.1.4.1 ce qui conclut la preuve.  $\square$

Les transitions en zones sèches-mouillées dans le théorème ci-dessus s'opèrent à l'aide des disjonctions de cas dans les expressions (3.66) et (3.67). Ces disjonctions de cas sont robustes (voir [164, Section 3.1.2.4]) mais leurs expressions ne sont pas continues avec le schéma initialement défini dans la Proposition 3.1.4.1.

La procédure de limitation  $\min(\max(\cdot, \cdot), \cdot)$  utilisée dans (3.67) est une version paramétrée par  $\varepsilon > 0$  des procédures de limitations déployées dans les sections 2.3 et 3.1.3. L'introduction de ce *petit*

paramètre  $\varepsilon$  permet de garantir les inégalités strictes  $h_L^* > 0$  et  $h_R^* > 0$ . Ces inégalités sont essentielles sur le plan théorique puisqu'elles sont aux fondements de la résolution du système d'équations (3.26), (3.55) exposée dans la Proposition 3.1.4.1. De plus, d'après les définitions (3.64b) et (3.64d),  $u_L^*$  et  $u_R^*$  ne sont pas définis lorsque  $h_L^* = 0$  ou  $h_R^* = 0$ .

Il est néanmoins d'usage de considérer en pratique le cas limite  $\varepsilon = 0$  et le lecteur intéressé trouvera des exemples dans [165, 166]. Ce cas limite peut se traiter sans difficulté pour le schéma numérique du Théorème 3.1.4.1 car que la programmation du solveur de Riemann (3.22) ne nécessite que la connaissance de  $h_L^*$ ,  $h_R^*$  et des produits  $h_L^* u_L^*$ ,  $h_R^* u_R^*$ . Or, en admettant  $h_L^* > 0$  et  $h_R^* > 0$  dans les définitions (3.67), il suit

$$h_L^* u_L^* = h_L^* \hat{u}^{\text{HLL}} + \frac{q_{LR}[h^*]}{2h^{\text{HLL}}} \sqrt{\frac{h_L^* h_R^*}{h_L h_R}}, \quad h_R^* u_R^* = h_R^* \hat{u}^{\text{HLL}} - \frac{q_{LR}[h^*]}{2h^{\text{HLL}}} \sqrt{\frac{h_L^* h_R^*}{h_L h_R}}. \quad (3.69)$$

Puisque les deux expressions ci-dessus sont continues lorsque  $h_L^*$  ou  $h_R^*$  tendent vers zéro, il est suffisant en pratique de programmer les quantités  $h_L^*$  et  $h_R^*$  en passant formellement à la limite  $\varepsilon = 0$  dans (3.67) puis de programmer les produits  $h_L^* u_L^*$  et  $h_R^* u_R^*$  données par (3.69). Enfin, lorsque les procédures de limitations sont actives, le système d'équations (3.26) n'est plus nécessairement vérifié et par conséquent la stabilité entropique pourrait être localement perdue dans ce cas.

Le Théorème 3.1.4.1 montre l'existence d'un schéma numérique pour les équations de Saint Venant (3.3) à stabilité entropique locale et *well-balanced* pour toutes les solutions stationnaires régulières. Ce schéma est une généralisation du schéma entropique *well-balanced* pour le lac au repos proposé dans le Théorème 3.1.3.1 car un calcul direct montre que si  $q_{LR}$  est nul alors les formulations (3.67) dégénèrent vers (3.51). Par conséquent, cette généralisation unifie les schémas numériques entropiques présentés dans le Théorème 3.1.3.1 (*well-balanced* pour le lac au repos) ainsi que dans le Lemme 2.3.3 (cas d'une fonction  $z$  constante).

La section suivante présente un schéma numérique entropique et *well-balanced* pour les équations de Ripa construit à partir d'une proposition analogue à la Proposition 3.1.1.1.

## 3.2 Schéma entropique et *well-balanced* pour les équations de Ripa

### 3.2.1 Condition suffisante de stabilité entropique locale

Pour la donnée d'une fonction  $z : \mathbb{R} \rightarrow \mathbb{R}$  bornée, régulière et indépendante du temps, le système de Ripa en une dimension d'espace est maintenant considéré. Ce système d'équations s'écrit

$$\partial_t \begin{pmatrix} h \\ hu \\ h\theta \end{pmatrix} + \partial_x \begin{pmatrix} hu \\ hu^2 + gh^2\varphi(\theta)/2 \\ h\theta u \end{pmatrix} = \begin{pmatrix} 0 \\ -gh\varphi(\theta) \\ 0 \end{pmatrix} \partial_x z. \quad (3.70)$$

L'inconnue  $w$  du système ci-dessus est  $(h, hu, h\theta)^T$  dans  $\Omega = \{(h, hu, h\theta) \in \mathbb{R}^3 \mid h > 0, hu \in \mathbb{R}, h\theta > 0\}$  et  $\varphi : \mathbb{R}_*^+ \rightarrow \mathbb{R}_*^+$  est une fonction régulière bijective donnée. La fonction de flux  $f : \Omega \rightarrow \mathbb{R}^3$  et le terme source  $S : \Omega \rightarrow \mathbb{R}^3$  du système de Ripa (3.70) sont respectivement donné par

$$f(w) = (hu, hu^2 + gh^2\varphi(\theta)/2, h\theta u)^T, \\ S(w) = (0, -gh\varphi(\theta), 0)^T.$$

Puisque la fonction  $z$  donnée est bornée, régulière et que le système d'équations (3.70) est invariant par translation en  $z$ , on peut supposer, sans perdre de généralité, que  $z \geq 0$ . Par conséquent, et en notant  $\hat{w} = (h, hu, h\theta, z)^T$ ,  $\hat{w}$  prend ses valeurs dans un espace convexe  $\hat{\Omega}$  défini par

$$\hat{\Omega} = \{(h, hu, h\theta, z) \in \mathbb{R}^4 \mid h > 0, hu \in \mathbb{R}, h\theta > 0, z \geq 0\}.$$

De plus, les solutions du système de Ripa (3.70) vérifient une inégalité d'entropie (3.2) dans laquelle le couple entropie, flux d'entropie  $(\hat{\eta}, \hat{G})$  est donné par

$$\hat{\eta}(\hat{w}) = \frac{(hu)^2}{2h} + \frac{g\varphi((h\theta)/h)h}{2}(h + 2z), \quad \hat{G}(\hat{w}) = \frac{(hu)^3}{2h^2} + g(hu)\varphi((h\theta)/h)(h + z), \quad \forall \hat{w} \in \hat{\Omega}. \quad (3.71)$$

Pour caractériser le couple  $(\hat{\eta}, \hat{G})$  défini ci-dessus, on introduit les couples  $(\eta, G)$  et  $(\psi, H)$  respectivement donné par

$$\eta(w) = \frac{(hu)^2}{2h} + \frac{g\varphi((h\theta)/h)h^2}{2}, \quad G(w) = \frac{(hu)^3}{2h^2} + g(hu)\varphi((h\theta)/h)h, \quad (3.72a)$$

$$\psi(w) = gh\varphi((h\theta)/h), \quad H(w) = g(hu)\varphi((h\theta)/h). \quad (3.72b)$$

Les fonctions  $(\eta, G)$  définies ci-dessus sont un couple entropie, flux d'entropie du système de Ripa (3.70) dans le cas particulier où la fonction  $z$  est constante. Les fonctions  $(\psi, H)$  forment également un couple entropie, flux d'entropie du système de Ripa (3.70). Ces fonctions sont introduites dans ce mémoire de thèse car elles permettent une réécriture synthétique de  $\hat{\eta}$  et  $\hat{G}$  sous la forme

$$\hat{\eta}(\hat{w}) = \eta(w) + \psi(w)z, \quad \hat{G}(\hat{w}) = G(w) + H(w)z, \quad \forall \hat{w} \in \hat{\Omega}. \quad (3.73)$$

Avec ces définitions et en ajoutant des contraintes sur la fonction  $\theta \mapsto \varphi(\theta)$ , il est possible d'énoncer des propriétés pour le couple  $(\hat{\eta}, \hat{G})$  donnée par (3.71).

**Lemme 3.2.1.1** (Propriétés du couple  $(\hat{\eta}, \hat{G})$  (3.71)). *Soit  $\varphi : \mathbb{R}_*^+ \rightarrow \mathbb{R}_*^+$  une fonction régulière bijective. Considérons les couples de fonctions  $(\eta, G)$  et  $(\psi, H)$  à valeurs dans  $\mathbb{R}$  et définis par (3.72). Les assertions suivantes sont vérifiées*

i) *Si  $\theta \mapsto \varphi(\theta)$  satisfait les inégalités strictes*

$$\varphi''(\theta)\varphi(\theta) - \frac{\varphi'(\theta)^2}{2} > 0, \quad \varphi(\theta) - \theta\varphi'(\theta) + \frac{\theta^2}{2}\varphi''(\theta) > 0, \quad \forall \theta \in \mathbb{R}_*^+, \quad (3.74)$$

*alors  $w \mapsto \eta(w)$  est strictement convexe.*

ii) *Si  $\theta \mapsto \varphi(\theta)$  est strictement convexe alors  $w \mapsto \psi(w)$  est convexe.*

iii) *L'ensemble des fonctions régulières  $\varphi$  définies sur  $\mathbb{R}_*^+$ , à valeurs dans  $\mathbb{R}_*^+$ , strictement convexes, bijectives et qui vérifient les inégalités (3.74) est non vide.*

*Démonstration.* Comme l'assertion i) est montrée dans [78], il est seulement nécessaire de montrer ii) et iii).

En ce qui concerne ii), en considérant  $(w_1, z_1)$ ,  $(w_2, z_2)$  dans  $\hat{\Omega}^2$  et  $s$  un réel dans  $[0, 1]$ , la convexité de  $\theta \mapsto \varphi(\theta)$  permet d'écrire

$$\begin{aligned} \psi(sw_1 + (1-s)w_2) &= g(sh_1 + (1-s)h_2)\varphi\left(\theta_1 \frac{sh_1}{sh_1 + (1-s)h_2} + \theta_2 \frac{(1-s)h_2}{sh_1 + (1-s)h_2}\right), \\ &\leq g(sh_1 + (1-s)h_2)\left(\frac{sh_1}{sh_1 + (1-s)h_2}\varphi(\theta_1) + \frac{(1-s)h_2}{sh_1 + (1-s)h_2}\varphi(\theta_2)\right), \\ &\leq s\psi(w_1) + (1-s)\psi(w_2). \end{aligned}$$

Puisque l'expression ci-dessus est la définition d'une fonction convexe, ce calcul termine la preuve de l'assertion ii).

Pour montrer l'assertion iii), un calcul direct montre que la fonction  $\varphi : \theta \mapsto \theta^3$  satisfait les conditions énoncées ce qui conclut la preuve.  $\square$

Le lemme ci-dessus présente une écriture de l'entropie  $\hat{\eta}$  avec les fonctions  $w \mapsto \eta(w)$  et  $w \mapsto \psi(w)$ . Quel que soit la fonction  $\varphi$  choisie, le critère de Sylvester montre que l'entropie  $\hat{w} \mapsto \hat{\eta}(\hat{w})$  ne peut pas être convexe. Cependant, les auteurs de [78] introduisent les contraintes (3.74) pour forcer la convexité de  $w \mapsto \eta(w)$ .

Dans ce mémoire de thèse, la convexité stricte de la fonction  $\theta \mapsto \varphi(\theta)$  est proposée comme hypothèse additionnelle. Cette hypothèse garantit la convexité de  $w \mapsto \psi(w)$  et d'après le Lemme 3.2.1.1-iii), elle est également compatible avec les contraintes existantes. Par conséquent, il est possible de formuler  $\hat{\eta}$  à l'aide des deux fonctions convexes  $\eta$  et  $\psi$ .

Concernant l'approximation numérique, la fonction régulière donnée  $z : \mathbb{R} \rightarrow \mathbb{R}^+$  est tout d'abord discrétisée comme suit :

$$z_i = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} z(x) dx, \quad \forall i \in \mathbb{Z}. \quad (3.75)$$

Puisque la fonction  $z$  est positive, la séquence  $(z_i)_{i \in \mathbb{Z}}$  définie ci-dessus est également positive.

Ensuite, les solutions du système de Ripa (3.70) sont approchées par un schéma de type Godunov. D'après la section 1.4.4, un tel schéma est défini par la donnée d'un solveur de Riemann  $\tilde{w} : \mathbb{R} \times \hat{\Omega} \times \hat{\Omega} \rightarrow \Omega$  qui satisfait la relation de consistance intégrale (1.63). En complétant cette donnée par la notation  $\hat{w}_i^n = (h_i^n, h_i^n u_i^n, h_i^n \theta_i^n, z_i)^T$  dans  $\hat{\Omega}$  et par

$$\xi_{i+\frac{1}{2}}^n : (x, t) \mapsto \frac{x - x_{i+\frac{1}{2}}}{t^n + t}, \quad \forall (i, n, x, t) \in \mathbb{Z} \times \mathbb{N} \times [x_i, x_{i+1}] \times [0, \Delta t],$$

un schéma numérique de type Godunov pour le système de Ripa (3.70) s'écrit pour tout  $i$  dans  $\mathbb{Z}$

$$w_i^{n+1} = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_i} \tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), \hat{w}_{i-1}^n, \hat{w}_i^n) dx + \frac{1}{\Delta x} \int_{x_i}^{x_{i+\frac{1}{2}}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n) dx. \quad (3.76)$$

Enfin, pour  $(\hat{S}_{i+1/2}^n)_{i \in \mathbb{Z}}$  une séquence de  $\mathbb{R}^3$  qui vérifie la condition de consistance (1.62), la Définition 1.4.4.1 donne l'existence d'une fonction  $\hat{\sigma} : (\hat{\Omega})^2 \rightarrow \mathbb{R}$  telle que

$$\hat{S}_{i+\frac{1}{2}}^n = \begin{pmatrix} 0 \\ \hat{\sigma}(\hat{w}_i^n, \hat{w}_{i+1}^n) \\ 0 \end{pmatrix} \frac{z_{i+1} - z_i}{\Delta x}, \quad \hat{\sigma}(\hat{w}, \hat{w}) = -gh\varphi((h\theta)/h), \quad \forall (i, \hat{w}) \in \mathbb{Z} \times \hat{\Omega}. \quad (3.77)$$

La fonction  $\hat{\sigma}$  est un paramètre libre à choisir au moins en vertu de la relation de consistance  $\hat{\sigma}(\hat{w}, \hat{w}) = -gh$ . D'après l'expression (3.77), la donnée de la fonction  $\hat{\sigma}$  définit la séquence  $(\hat{S}_{i+1/2}^n)_{i \in \mathbb{Z}}$  et puisqu'il n'y a pas d'ambiguïté, la notation suivante est désormais adoptée :

$$\hat{s}_{i+\frac{1}{2}} = \hat{\sigma}(\hat{w}_i^n, \hat{w}_{i+1}^n) \frac{z_{i+1} - z_i}{\Delta x}, \quad \forall i \in \mathbb{Z}.$$

En utilisant la séquence ci-dessus, le Lemme 1.4.4.1 autorise une réécriture équivalente du schéma de type Godunov (3.76) sous la forme

$$w_i^{n+1} = w_i^n - \frac{\Delta t}{\Delta x} (\hat{\mathcal{F}}(\hat{w}_i^n, \hat{w}_{i+1}^n) - \hat{\mathcal{F}}(\hat{w}_{i-1}^n, \hat{w}_i^n)) + \frac{\Delta t}{2} (\hat{S}_{i+\frac{1}{2}}^n + \hat{S}_{i-\frac{1}{2}}^n), \quad \forall i \in \mathbb{Z},$$

avec,

$$\begin{aligned} \hat{\mathcal{F}}(\hat{w}_i^n, \hat{w}_{i+1}^n) &= \frac{f(w_{i+1}^n) + f(w_i^n)}{2} - \frac{\Delta x}{4\Delta t} (w_{i+1}^n - w_i^n) + \frac{1}{2\Delta t} \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n) dx \\ &\quad - \frac{1}{2\Delta t} \int_{x_i}^{x_{i+\frac{1}{2}}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n) dx. \end{aligned}$$

L'objectif de cette section est d'écrire la relation de consistance intégrale, des conditions suffisantes de stabilité entropique ainsi que la propriété *well-balanced* pour un solveur de Riemann dédié aux équations de Ripa (3.70).

Tout d'abord, et comme cela est indiqué dans la section 1.4.4, la relation de consistance intégrale est valide pour tous les systèmes d'équations avec un terme source de la forme  $S(w)\partial_x z$ . Par conséquent, la relation de consistance intégrale (3.12) se lit *mutatis mutandis* pour le cas du système de Ripa (3.70).

Ensuite, la propriété *well-balanced* nécessite une définition des équilibres locaux pour les solutions stationnaires du système de Ripa (3.70). Plusieurs définitions sont possibles [78, 95] mais l'extension directe des équilibres locaux introduits en Section 3.1.1 est donnée par la définition suivante :



**Définition 3.2.1.1** (Équilibre local et schéma *well-balanced* pour les équations de Ripa (3.70)). *Considérons les équations de Ripa (3.70) complétées par la notation*

$$\hat{B}(\hat{w}) = \frac{u^2}{2} + g\varphi(\theta)(h+z), \quad \forall \hat{w} \in \hat{\Omega}.$$

À la date  $t^n$ , considérons  $\hat{w}_i^n$  et  $\hat{w}_{i+1}^n$  deux états constants dans  $\hat{\Omega}$  respectivement à gauche et à droite de l'interface  $x_{i+\frac{1}{2}}$ .

i) Les deux états  $\hat{w}_i^n$  et  $\hat{w}_{i+1}^n$  sont dits à l'équilibre local pour les équilibres mouvants si

$$h_i^n u_i^n = h_{i+1}^n u_{i+1}^n, \quad \hat{B}(\hat{w}_i^n) = \hat{B}(\hat{w}_{i+1}^n). \quad (3.78)$$

ii) Les deux états  $\hat{w}_i^n$  et  $\hat{w}_{i+1}^n$  sont dits à l'équilibre local pour le lac au repos si

$$u_i^n = u_{i+1}^n = 0, \quad \theta_i^n = \theta_{i+1}^n, \quad h_i^n + z_i = h_{i+1}^n + z_{i+1}. \quad (3.79)$$

iii) Un schéma numérique de type Godunov (3.76) est dit *well-balanced* pour les équilibres mouvants s'il préserve exactement une séquence  $(\hat{w}_i^n)_{i \in \mathbb{Z}}$  qui satisfait à chaque interface du maillage un équilibre local (3.78). Cette propriété s'écrit

$$\text{si } \forall i \in \mathbb{Z} \left\{ \begin{array}{l} h_i^n u_i^n = h_{i+1}^n u_{i+1}^n, \\ \hat{B}(\hat{w}_i^n) = \hat{B}(\hat{w}_{i+1}^n), \end{array} \right. \text{ alors } w_i^{n+1} = w_i^n, \quad \forall i \in \mathbb{Z}.$$

Un schéma numérique *well-balanced* pour le lac au repos s'énonce mutadis mutandis en considérant les équilibres locaux (3.79). Cette propriété s'écrit dans ce cas

$$\text{si } \forall i \in \mathbb{Z} \left\{ \begin{array}{l} u_i^n = u_{i+1}^n = 0, \\ \theta_i^n = \theta_{i+1}^n, \\ h_i^n + z_i = h_{i+1}^n + z_{i+1}, \end{array} \right. \text{ alors } w_i^{n+1} = w_i^n, \quad \forall i \in \mathbb{Z}.$$

Avec la définition ci-dessus, la condition suffisante à la propriété *well-balanced* donnée par (3.16) reste valable pour le cas du système de Ripa (3.70).

En revanche, la construction d'une inégalité de la forme (3.13) pour le système de Ripa (3.70) n'est pas simple et la difficulté provient une nouvelle fois de la perte de convexité de l'entropie  $\hat{w} \mapsto \hat{\eta}(\hat{w})$ . L'idée proposée dans la suite repose sur la réécriture du couple  $(\hat{\eta}, \hat{G})$  donnée par (3.73) puis sur l'utilisation des convexités de  $w \mapsto \eta(w)$  et  $w \mapsto \psi(w)$  obtenues avec les contraintes énoncées dans le Lemme 3.2.1.1.

Avec ces éléments, une extension de la Proposition 3.1.1.1 est maintenant accessible.

**Proposition 3.2.1.1** (Solveur de Riemann entropique et *well-balanced* pour le système de Ripa). *Avec la donnée d'une fonction  $\varphi : \mathbb{R}_*^+ \rightarrow \mathbb{R}_*^+$ , régulière, strictement convexe, bijective et satisfaisant les inégalités (3.74), considérons les couples  $(\eta, G)$  et  $(\psi, H)$  définis par (3.72).*

*Soient  $(\hat{S}_{i+\frac{1}{2}}^n)_{i \in \mathbb{Z}}$  une suite de  $\mathbb{R}^3$  et un solveur de Riemann  $\tilde{w} : \mathbb{R} \times \hat{\Omega} \times \hat{\Omega} \rightarrow \Omega$  satisfaisant respectivement les conditions de consistances des Définitions 1.4.4.1 et 1.4.4.2. Supposons qu'une condition CFL de non interaction est satisfaite. Considérons  $(\mathcal{H}_{i+\frac{1}{2}})_{i \in \mathbb{Z}}$  une séquence de réels telle que*

$$\begin{aligned} \mathcal{H}_{i+\frac{1}{2}} &= \frac{H(w_i^n) + H(w_{i+1}^n)}{2} - \frac{\Delta x}{4\Delta t} (\psi(w_{i+1}^n) - \psi(w_i^n)) + \frac{1}{2\Delta t} \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \psi(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n)) \, dx \\ &\quad - \frac{1}{2\Delta t} \int_{x_i}^{x_{i+\frac{1}{2}}} \psi(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n)) \, dx. \end{aligned} \quad (3.80)$$

Les assertions suivantes sont vérifiées.

i) Si pour tout  $i$  dans  $\mathbb{Z}$ , le solveur de Riemann vérifie la relation de consistance intégrale

$$\frac{1}{\Delta x} \int_{x_i}^{x_{i+1}} \tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n) dx = \frac{w_i^n + w_{i+1}^n}{2} - \frac{\Delta t}{\Delta x} (f(w_{i+1}^n) - f(w_i^n)) + \Delta t \hat{S}_{i+\frac{1}{2}}^n, \quad (3.81)$$

alors le schéma de type Godunov (3.76) est consistant avec les équations de Ripa (3.70).

ii) Si le solveur de Riemann satisfait les inégalités suivantes pour tout  $i$  dans  $\mathbb{Z}$

$$\begin{aligned} \frac{1}{\Delta x} \int_{x_i}^{x_{i+1}} \psi(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n)) dx \leq \\ \frac{\psi(w_i^n) + \psi(w_{i+1}^n)}{2} - \frac{\Delta t}{\Delta x} (H(w_{i+1}^n) - H(w_i^n)), \end{aligned} \quad (3.82a)$$

$$\begin{aligned} \frac{1}{\Delta x} \int_{x_i}^{x_{i+1}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), w_i^n, w_{i+1}^n)) dx \leq \\ \frac{\eta(w_i^n) + \eta(w_{i+1}^n)}{2} - \frac{\Delta t}{\Delta x} (G(w_{i+1}^n) - G(w_i^n)) - \frac{\Delta t}{\Delta x} (z_{i+1} - z_i) \mathcal{H}_{i+\frac{1}{2}}, \end{aligned} \quad (3.82b)$$

alors le schéma de type Godunov (3.76) vérifie une inégalité d'entropie discrète locale pour le couple  $(\hat{\eta}, \hat{G})$  donné par (3.71). Cette inégalité s'écrit

$$\frac{\hat{\eta}(\hat{w}_i^{n+1}) - \hat{\eta}(\hat{w}_i^n)}{\Delta t} + \frac{\hat{G}_{i+\frac{1}{2}} - \hat{G}_{i-\frac{1}{2}}}{\Delta x} \leq 0, \quad \forall i \in \mathbb{Z}, \quad (3.83)$$

avec,

$$\begin{aligned} \hat{G}_{i+\frac{1}{2}} &= \frac{G(w_{i+1}^n) + G(w_i^n)}{2} + \mathcal{H}_{i+\frac{1}{2}} \frac{z_{i+1} + z_i}{2} - \frac{\Delta x}{4\Delta t} (\eta(w_{i+1}^n) - \eta(w_i^n)) \\ &+ \frac{1}{2\Delta t} \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n)) dx \\ &- \frac{1}{2\Delta t} \int_{x_i}^{x_{i+\frac{1}{2}}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n)) dx. \end{aligned} \quad (3.84)$$

iii) Si le solveur de Riemann satisfait

$$\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n) = \begin{cases} w_i^n & \text{si } x - x_{i+\frac{1}{2}} < 0, \\ w_{i+1}^n & \text{sinon,} \end{cases} \quad \forall i \in \mathbb{Z}, \quad (3.85)$$

dès que la séquence  $(\hat{w}_i^n)_{i \in \mathbb{Z}}$  vérifie, à chaque interface du maillage, un équilibre local pour les équilibres mouvants (3.78) (resp. le lac au repos (3.79)) alors le schéma de type Godunov (3.76) est well-balanced pour les équilibres mouvants (resp. le lac au repos)

*Démonstration.* Pour l'assertion i), et comme  $z$  est régulière, la quantité  $\hat{S}_{i+\frac{1}{2}}^n$  donnée par (3.77) est consistante avec  $(0, -gh\varphi(\theta)\partial_x z, 0)^T$  et elle vérifie de plus

$$\hat{S}(\hat{w}, \hat{w}) = 0, \quad \forall \hat{w} \in \hat{\Omega}. \quad (3.86)$$

Par conséquent, un solveur de Riemann qui vérifie la relation de consistance intégrale (3.81) donne un schéma de type Godunov (3.76) consistant.

Concernant l'assertion ii), et comme  $\varphi$  satisfait les conditions (3.74), la fonction  $w \mapsto \eta(w)$  est strictement convexe. Ainsi, l'inégalité de Jensen appliquée au schéma de type Godunov (3.76) donne

$$\begin{aligned} \eta(w_i^{n+1}) \leq \\ \frac{1}{2\Delta x} \int_{x_i}^{x_{i+\frac{1}{2}}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n)) dx + \frac{1}{2\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_i} \eta(\tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), \hat{w}_{i-1}^n, \hat{w}_i^n)) dx \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{2\Delta x} \int_{x_{i-1}}^{x_i} \eta(\tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), \hat{w}_{i-1}^n, \hat{w}_i^n)) \, dx + \frac{1}{2\Delta x} \int_{x_i}^{x_{i+1}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n)) \, dx \\
& - \frac{1}{2\Delta x} \int_{x_{i-1}}^{x_{i-\frac{1}{2}}} \eta(\tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), \hat{w}_{i-1}^n, \hat{w}_i^n)) \, dx - \frac{1}{2\Delta x} \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n)) \, dx.
\end{aligned}$$

Avec l'introduction de l'inégalité (3.82b) dans le résultat ci-dessus, il suit

$$\begin{aligned}
\eta(w_i^{n+1}) & \leq \\
& \frac{1}{2\Delta x} \int_{x_i}^{x_{i+\frac{1}{2}}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n)) \, dx + \frac{1}{2\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_i} \eta(\tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), \hat{w}_{i-1}^n, \hat{w}_i^n)) \, dx \\
& + \frac{1}{2} \left( \frac{\eta(w_{i-1}^n) + \eta(w_i^n)}{2} - \frac{\Delta t}{\Delta x} (G(w_i^n) - G(w_{i-1}^n)) - \frac{\Delta t}{\Delta x} (z_i - z_{i-1}) \mathcal{H}_{i-\frac{1}{2}} \right) \\
& + \frac{1}{2} \left( \frac{\eta(w_i^n) + \eta(w_{i+1}^n)}{2} - \frac{\Delta t}{\Delta x} (G(w_{i+1}^n) - G(w_i^n)) - \frac{\Delta t}{\Delta x} (z_{i+1} - z_i) \mathcal{H}_{i+\frac{1}{2}} \right) \\
& - \frac{1}{2\Delta x} \int_{x_{i-1}}^{x_{i-\frac{1}{2}}} \eta(\tilde{w}(\xi_{i-\frac{1}{2}}^n(x, \Delta t), \hat{w}_{i-1}^n, \hat{w}_i^n)) \, dx - \frac{1}{2\Delta x} \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \eta(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}_i^n, \hat{w}_{i+1}^n)) \, dx.
\end{aligned} \tag{3.87}$$

Par ailleurs, l'inégalité (3.82a) coïncide avec l'inégalité (2.4) pour le couple  $(\psi, H)$  défini dans (3.72b). Comme  $\theta \mapsto \varphi(\theta)$  est strictement convexe, la fonction  $w \mapsto \psi(w)$  est convexe et par application de la Proposition 2.1.1, il suit immédiatement

$$\psi(w_i^{n+1}) \leq \psi(w_i^n) - \frac{\Delta t}{\Delta x} (\mathcal{H}_{i+\frac{1}{2}} - \mathcal{H}_{i-\frac{1}{2}}). \tag{3.88}$$

Comme la fonction  $z$  est positive, la quantité  $z_i$  donnée par (3.75) est également positive pour tout  $i$  dans  $\mathbb{Z}$  et par conséquent en multipliant l'inégalité ci-dessus par  $z_i$ , on en déduit

$$z_i \psi(w_i^{n+1}) \leq z_i \psi(w_i^n) - \frac{\Delta t}{\Delta x} z_i (\mathcal{H}_{i+\frac{1}{2}} - \mathcal{H}_{i-\frac{1}{2}}).$$

La sommation membre à membre de l'inégalité ci-dessus à l'inégalité (3.87) permet d'écrire une inégalité de la forme

$$\eta(w_i^{n+1}) + \psi(w_i^{n+1}) z_i \leq \eta(w_i^n) + \psi(w_i^n) z_i - \frac{\Delta t}{\Delta x} (\hat{\mathcal{G}}_{i+\frac{1}{2}} - \hat{\mathcal{G}}_{i-\frac{1}{2}}), \quad \forall i \in \mathbb{Z},$$

avec  $\mathcal{G}_{i+\frac{1}{2}}$  donné par (3.84). En utilisant la réécriture de  $\hat{\eta}(\hat{w})$  donnée par (3.73), il vient  $\hat{\eta}(\hat{w}) = \eta(w) + z\psi(w)$  et comme la fonction  $z$  est indépendante du temps, l'inégalité précédente se réécrit sous la forme de l'inégalité (3.83).

Avant de terminer la preuve de l'assertion *ii*), il faut vérifier que la quantité  $\hat{\mathcal{G}}_{i+\frac{1}{2}}$  donnée par (3.84) est consistante avec  $\hat{G}$ . En utilisant l'égalité de consistance du solveur  $\tilde{w}(\xi, \hat{w}, \hat{w}) = w$  dans la définition de  $\mathcal{H}$  donnée par (3.80), il suit

$$\begin{aligned}
\mathcal{H}(\hat{w}, \hat{w}) & = \frac{H(w) + H(w)}{2} - \frac{\Delta x}{4\Delta t} (\psi(w) - \psi(w)) \\
& + \frac{1}{2\Delta t} \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \psi(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}, \hat{w})) \, dx - \frac{1}{2\Delta t} \int_{x_i}^{x_{i+\frac{1}{2}}} \psi(\tilde{w}(\xi_{i+\frac{1}{2}}^n(x, \Delta t), \hat{w}, \hat{w})) \, dx, \\
& = H(w) + \frac{1}{2\Delta t} \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \psi(w) \, dx - \frac{1}{2\Delta t} \int_{x_i}^{x_{i+\frac{1}{2}}} \psi(w) \, dx, \\
& = H(w).
\end{aligned}$$

Avec l'égalité ci-dessus et un calcul analogue, on montre alors  $\hat{\mathcal{G}}(\hat{w}, \hat{w}) = \hat{G}(w) + zH(w)$ . Puisque l'expression (3.73) donne  $\hat{G}(\hat{w}) = G(w) + zH(w)$ , cette dernière égalité termine de montrer l'assertion *ii*).

Pour l'assertion *iii*), la condition suffisante de propriété *well-balanced* donnée par (3.85) est identique à (3.16) et par conséquent sa démonstration s'identifie *mutatis mutandis* à celle menée dans la Proposition 3.1.1.1-*iii*). Ce dernier argument conclut la preuve.  $\square$

La proposition ci-dessus fait apparaître une condition suffisante de stabilité entropique discrète sous la forme de deux inégalités respectivement données par (3.82a)-(3.82b). La première de ces inégalités est un support pour la démonstration mais elle produit aussi l'inégalité (3.88) qui s'identifie à une stabilité entropique pour le couple  $(\psi, H)$  défini dans (3.72b). Pour une fonction  $\varphi$  donnée, un schéma numérique défini à partir de la Proposition 3.2.1.1 vérifie donc deux stabilités entropiques locales : une pour le couple  $(\psi, H)$  et l'autre pour le couple  $(\hat{\eta}, \hat{G})$ . La seconde de ces stabilités entropiques est obtenue à partir de l'inégalité (3.82b) qui est un analogue de (3.13). On détaille maintenant un solveur de Riemann constant par morceaux de la forme (3.22) qui satisfait la relation de consistance intégrale (3.81) et les inégalités (3.82).

### 3.2.2 Solveur de Riemann à deux états intermédiaires symétriques, consistant et entropique

Pour définir un solveur de Riemann dans le cas des équations de Ripa (3.70), les notations (3.20) demeurent mais elles sont complétées par  $\mathcal{H}_{i+\frac{1}{2}} = \mathcal{H}_{LR}$  ainsi que par un réel  $\lambda > 0$  et une condition CFL de non interaction donnés par

$$\lambda \geq \max_{\alpha \in \{L, R\}} (|u_\alpha \pm \sqrt{gh_\alpha \varphi(\theta_\alpha)}|, |u_\alpha|), \quad \frac{\lambda \Delta t}{\Delta x} \leq \frac{1}{2}. \quad (3.89)$$

En considérant les inégalités ci-dessus satisfaites, un solveur de Riemann à deux états intermédiaires symétriques noté  $\tilde{w}(\cdot, \hat{w}_L, \hat{w}_R) : \mathbb{R} \rightarrow \mathbb{R}^3$  s'écrit

$$\tilde{w}(x/\Delta t, \hat{w}_L, \hat{w}_R) = \begin{cases} w_L & \text{si } \frac{x}{\Delta t} \leq -\lambda, \\ w_L^* & \text{si } -\lambda < \frac{x}{\Delta t} \leq 0, \\ w_R^* & \text{si } 0 < \frac{x}{\Delta t} \leq \lambda, \\ w_R & \text{si } \lambda < \frac{x}{\Delta t}. \end{cases} \quad (3.90)$$

Dans le cas du système de Ripa (3.70), les quantités  $w^{\text{HLL}}$  et  $\eta^{\text{HLL}}$  respectivement dans  $\mathbb{R}^3$  et dans  $\mathbb{R}$  se lisent

$$w^{\text{HLL}} = (h^{\text{HLL}}, (hu)^{\text{HLL}}, (h\theta)^{\text{HLL}})^T = \frac{w_R + w_L}{2} - \frac{f(w_R) - f(w_L)}{2\lambda}, \quad (3.91a)$$

$$\eta^{\text{HLL}} = \frac{\eta(w_R) + \eta(w_L)}{2} - \frac{G(w_R) - G(w_L)}{2\lambda}. \quad (3.91b)$$

Ces quantités sont maintenant complétées par un analogue défini avec le couple  $(\psi, H)$  et pour lequel des propriétés peuvent être établies.

**Lemme 3.2.2.1.** *Soient  $\hat{w}_L$  et  $\hat{w}_R$  deux éléments constants de  $\hat{\Omega}$  et  $\varphi : \mathbb{R}_*^+ \rightarrow \mathbb{R}_*^+$  une fonction régulière, strictement convexe, bijective. Considérons le couple  $(\psi, H)$  défini dans (3.72b) et le réel  $\psi^{\text{HLL}}$  tel que*

$$\psi^{\text{HLL}} = \frac{\psi(w_R) + \psi(w_L)}{2} - \frac{H(w_R) - H(w_L)}{2\lambda}. \quad (3.92)$$

*Les deux assertions ci-dessous sont vérifiées.*

i) Si  $\lambda > 0$  satisfait l'inégalité (3.89) alors

$$h^{\text{HLL}} > 0, \quad \frac{(h\theta)^{\text{HLL}}}{h^{\text{HLL}}} = \theta^{\text{HLL}} > 0, \quad \psi^{\text{HLL}} \geq \psi(w^{\text{HLL}}). \quad (3.93)$$

ii) L'égalité  $\psi^{\text{HLL}} = \psi(w^{\text{HLL}})$  advient si et seulement si  $\theta_L = \theta_R = \theta$ . Dans ce cas, les égalités suivantes sont satisfaites :

$$\psi^{\text{HLL}} = \psi(w^{\text{HLL}}) = gh^{\text{HLL}}\varphi(\theta), \quad \theta^{\text{HLL}} = \theta. \quad (3.94)$$

*Démonstration.* L'inégalité stricte  $h^{\text{HLL}} > 0$  est démontrée dans le Lemme 2.3.1 en Section 2.3 et elle permet de définir  $\theta^{\text{HLL}} = (h\theta)^{\text{HLL}}/h^{\text{HLL}}$ . Pour montrer l'inégalité stricte  $\theta^{\text{HLL}} > 0$  et pour alléger les écritures, la notation  $[\cdot] = \cdot_R - \cdot_L$  est utilisée. Par définition,  $\theta^{\text{HLL}}$  s'écrit

$$\begin{aligned} \theta^{\text{HLL}} &= \frac{(h\theta)^{\text{HLL}}}{h^{\text{HLL}}} = \frac{(h\theta)_L + (h\theta)_R - [h\theta u/\lambda]}{2h^{\text{HLL}}}, \\ &= \theta_L \frac{h_L(1 + u_L/\lambda)}{2h^{\text{HLL}}} + \theta_R \frac{h_R(1 - u_R/\lambda)}{2h^{\text{HLL}}}. \end{aligned} \quad (3.95)$$

Comme  $\lambda$  répond à l'inégalité (3.89), il vérifie en particulier  $\lambda > \max(|u_L|, |u_R|)$ . Par conséquent, le membre de droite de l'égalité (3.95) est une combinaison convexe de  $\theta_L$  et de  $\theta_R$  qui sont tous deux strictement positifs, ce qui permet de déduire  $\theta^{\text{HLL}} > 0$ .

Concernant la preuve de l'inégalité  $\psi(w^{\text{HLL}}) \leq \psi^{\text{HLL}}$ , en utilisant l'expression de  $\theta^{\text{HLL}}$  donnée sous la forme d'une combinaison convexe dans (3.95) et en s'appuyant sur la convexité de  $\theta \mapsto \varphi(\theta)$ , il suit

$$\begin{aligned} \psi(w^{\text{HLL}}) &= h^{\text{HLL}}\varphi(\theta^{\text{HLL}}) = h^{\text{HLL}}\varphi\left(\theta_L \frac{h_L(1 + u_L/\lambda)}{2h^{\text{HLL}}} + \theta_R \frac{h_R(1 - u_R/\lambda)}{2h^{\text{HLL}}}\right), \\ &\leq \frac{h_L(1 + u_L/\lambda)}{2}\varphi(\theta_L) + \frac{h_R(1 - u_R/\lambda)}{2}\varphi(\theta_R). \end{aligned} \quad (3.96)$$

En développant et en réorganisant les termes de droite de l'inégalité ci-dessus, il vient

$$\begin{aligned} \frac{h_L(1 + u_L/\lambda)}{2}\varphi(\theta_L) + \frac{h_R(1 - u_R/\lambda)}{2}\varphi(\theta_R) &= \frac{h_L\varphi(\theta_L) + h_R\varphi(\theta_R)}{2} - \frac{h_R\varphi(\theta_R)u_R - h_L\varphi(\theta_L)u_L}{2\lambda} \\ &= \psi^{\text{HLL}}. \end{aligned} \quad (3.97)$$

Cette dernière expression associée à l'inégalité (3.96) permet de déduire l'inégalité  $\psi(w^{\text{HLL}}) \leq \psi^{\text{HLL}}$ .

Pour l'assertion *ii*), l'égalité (3.95) montre immédiatement que si  $\theta_L = \theta_R = \theta$  alors  $\theta^{\text{HLL}} = \theta$ . De plus, l'égalité (3.97) montre que la condition  $\theta_L = \theta_R$  est une condition suffisante à l'égalité  $\psi(w^{\text{HLL}}) = \psi^{\text{HLL}}$  et on montre maintenant qu'il s'agit également d'une condition nécessaire.

Puisque  $\varphi$  est régulière, un développement de Taylor avec reste intégrale permet d'écrire pour tout  $\alpha$  dans  $\{L, R\}$

$$\varphi(\theta_\alpha) = \varphi(\theta^{\text{HLL}}) + \varphi'(\theta^{\text{HLL}})(\theta_\alpha - \theta^{\text{HLL}}) + (\theta_\alpha - \theta^{\text{HLL}})^2 \int_0^1 (1-s)\varphi''(\theta^{\text{HLL}} + s(\theta_\alpha - \theta^{\text{HLL}})) ds.$$

Par conséquent, la quantité  $\psi^{\text{HLL}}$  définie dans (3.92) se lit également

$$\begin{aligned} \psi^{\text{HLL}} &= \frac{h_L(1 + u_L/\lambda)}{2} \left( \varphi(\theta^{\text{HLL}}) + \varphi'(\theta^{\text{HLL}})(\theta_L - \theta^{\text{HLL}}) \right) \\ &\quad + \frac{h_R(1 - u_R/\lambda)}{2} \left( \varphi(\theta^{\text{HLL}}) + \varphi'(\theta^{\text{HLL}})(\theta_R - \theta^{\text{HLL}}) \right) \\ &\quad + \frac{h_L(1 + u_L/\lambda)}{2} (\theta_L - \theta^{\text{HLL}})^2 \int_0^1 (1-s)\varphi''(\theta^{\text{HLL}} + s(\theta_L - \theta^{\text{HLL}})) ds \\ &\quad + \frac{h_R(1 - u_R/\lambda)}{2} (\theta_R - \theta^{\text{HLL}})^2 \int_0^1 (1-s)\varphi''(\theta^{\text{HLL}} + s(\theta_R - \theta^{\text{HLL}})) ds, \end{aligned}$$

$$\begin{aligned}
&= h^{\text{HLL}}\varphi(\theta^{\text{HLL}}) + \varphi'(\theta^{\text{HLL}})((h\theta)^{\text{HLL}} - h^{\text{HLL}}\theta^{\text{HLL}}) \\
&\quad + \frac{h_L(1+u_L/\lambda)}{2}(\theta_L - \theta^{\text{HLL}})^2 \int_0^1 (1-s)\varphi''(\theta^{\text{HLL}} + s(\theta_L - \theta^{\text{HLL}})) \, ds \\
&\quad + \frac{h_R(1-u_R/\lambda)}{2}(\theta_R - \theta^{\text{HLL}})^2 \int_0^1 (1-s)\varphi''(\theta^{\text{HLL}} + s(\theta_R - \theta^{\text{HLL}})) \, ds, \\
&= \psi(w^{\text{HLL}}) + \frac{h_L(1+u_L/\lambda)}{2}(\theta_L - \theta^{\text{HLL}})^2 \int_0^1 (1-s)\varphi''(\theta^{\text{HLL}} + s(\theta_L - \theta^{\text{HLL}})) \, ds \\
&\quad + \frac{h_R(1-u_R/\lambda)}{2}(\theta_R - \theta^{\text{HLL}})^2 \int_0^1 (1-s)\varphi''(\theta^{\text{HLL}} + s(\theta_R - \theta^{\text{HLL}})) \, ds. \tag{3.98}
\end{aligned}$$

Puisque  $\lambda$  vérifie en particulier  $\lambda > \max(|u_L|, |u_R|)$ , les quantités  $h_L(1+u_L/\lambda)$  et  $h_R(1-u_R/\lambda)$  sont strictement positives. Ainsi, si  $\psi^{\text{HLL}} = \psi(w^{\text{HLL}})$  alors l'expression (3.98) conduit nécessairement à

$$0 = (\theta_\alpha - \theta^{\text{HLL}})^2 \int_0^1 (1-s)\varphi''(\theta^{\text{HLL}} + s(\theta_\alpha - \theta^{\text{HLL}})) \, ds, \quad \alpha \in \{L, R\}.$$

En utilisant la convexité stricte de  $\theta \mapsto \varphi(\theta)$ , le résultat ci-dessus permet de déduire que  $\theta_L = \theta_R = \theta^{\text{HLL}}$  est une condition nécessaire à l'égalité  $\psi^{\text{HLL}} = \psi(w^{\text{HLL}})$  ce qui conclut la preuve.  $\square$

Avec un solveur de Riemann de la forme (3.90) donné et les propriétés du lemme ci-dessus, l'objectif est de déterminer les états intermédiaires  $w_L^*$  et  $w_R^*$  en vertu des équations (3.81) et des inégalités (3.82). Pour le cas des équations de Ripa (3.70), ces états intermédiaires se notent  $w_\alpha^* = (h_\alpha^*, h_\alpha^* u_\alpha^*, h_\alpha^* \theta_\alpha^*)^T$  pour  $\alpha$  dans  $\{L, R\}$ . Le lemme ci-dessous présente les équations et les inégalités qui les définissent.

**Lemme 3.2.2.2** (Application des (in-)équations (3.81)-(3.82) à un solveur de Riemann (3.90)). *Soient  $\hat{w}_L, \hat{w}_R$  deux états constants de  $\hat{\Omega}$  ainsi que  $\varphi : \mathbb{R}_*^+ \rightarrow \mathbb{R}_*^+$  une fonction régulière, strictement convexe, bijective et satisfaisant les inégalités (3.74). Considérons les couples les couples  $(\hat{\eta}, \hat{G}), (\eta, G), (\psi, H)$  respectivement défini par (3.71)-(3.72). Considérons ensuite  $\hat{s}_{LR}$  une discrétisation consistante de  $\frac{1}{\Delta x} \int_{-\Delta x/2}^{\Delta x/2} (-gh\varphi(\theta)\partial_x z) \, dx$  au sens de l'expression (3.77) et  $\tilde{w}(\cdot, \hat{w}_L, \hat{w}_R) : \mathbb{R} \rightarrow \mathbb{R}^3$  un solveur de Riemann de la forme (3.90). Supposons que  $\lambda > 0$  et  $\Delta t > 0$  satisfont les inégalités (3.89) puis considérons enfin les quantités  $w^{\text{HLL}}, \psi^{\text{HLL}}$  respectivement définies dans (3.91a)-(3.92) ainsi que les quantités réelles*

$$h^{\text{HLL}}\hat{u}^{\text{HLL}} = (hu)^{\text{HLL}} + \frac{\Delta x \hat{s}_{LR}}{2\lambda}, \tag{3.99a}$$

$$\hat{\eta}^{\text{HLL}} = \frac{\hat{\eta}(\hat{w}_R) + \hat{\eta}(\hat{w}_L)}{2} - \frac{\hat{G}(\hat{w}_R) - \hat{G}(\hat{w}_L)}{2\lambda}. \tag{3.99b}$$

Si le solveur de Riemann (3.90) vérifie la relation de consistance intégrale (3.81) ainsi que les inégalités (3.82), alors les états intermédiaires  $w_L^*$  et  $w_R^*$  du solveur de Riemann (3.90) satisfont

$$\frac{h_L^* + h_R^*}{2} = h^{\text{HLL}}, \tag{3.100a}$$

$$\frac{h_L^* u_L^* + h_R^* u_R^*}{2} = h^{\text{HLL}}\hat{u}^{\text{HLL}}, \tag{3.100b}$$

$$\frac{h_L^* \theta_L^* + h_R^* \theta_R^*}{2} = h^{\text{HLL}}\theta^{\text{HLL}}, \tag{3.100c}$$

$$\frac{\psi(w_L^*) + \psi(w_R^*)}{2} \leq \psi^{\text{HLL}}, \tag{3.100d}$$

$$\begin{aligned}
&\frac{h_L^* h_R^*}{8h^{\text{HLL}}}(u_R^* - u_L^*)^2 + \frac{h^{\text{HLL}}}{2} \frac{\psi(w_L^*) + \psi(w_R^*)}{2} + (h_R^* - h_L^* + 2(z_R - z_L)) \frac{\psi(w_R^*) - \psi(w_L^*)}{8} \leq \\
&\hat{\eta}^{\text{HLL}} - \frac{(h^{\text{HLL}}\hat{u}^{\text{HLL}})^2}{2h^{\text{HLL}}} - \psi^{\text{HLL}} \frac{z_L + z_R}{2}. \tag{3.100e}
\end{aligned}$$

Réciproquement, si les états intermédiaires  $w_L^*$  et  $w_R^*$  vérifient les équations et les inégalités ci-dessus alors le solveur de Riemann (3.90) satisfait (3.81)-(3.82).

Avant de démontrer le lemme énoncé ci-dessus, il est nécessaire de remarquer que la quantité  $h^{\text{HLL}}\hat{u}^{\text{HLL}}$  donnée par (3.99a) est paramétrée par  $\hat{s}_{LR}$ . Ce paramètre libre doit être choisi en vertu de la condition de consistante donnée dans (3.77) et une expression possible sera proposée dans la section suivante. Avec cette remarque, la démonstration du Lemme 3.2.2.2 est maintenant établie.

*Démonstration du Lemme 3.2.2.2.* Il faut d'abord montrer que (3.81)-(3.82) impliquent (3.100). Puisqu'une condition CFL de non interaction est vérifiée, une réécriture équivalente la relation de consistance intégrale est possible. En notant  $\nu = \lambda\Delta t/\Delta x$ , cette réécriture est donnée par

$$\begin{aligned} & \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \tilde{w}(x/\Delta t, \hat{w}_L, \hat{w}_R) dx \\ &= \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{-\lambda\Delta t} w_L dx + \frac{1}{\Delta x} \int_{-\lambda\Delta t}^0 w_L^* dx + \frac{1}{\Delta x} \int_0^{\lambda\Delta t} w_R^* dx + \frac{1}{\Delta x} \int_{\lambda\Delta t}^{\frac{\Delta x}{2}} w_R dx, \\ &= (-\nu + 1/2)w_L + \nu w_L^* + \nu w_R^* + (1/2 - \nu)w_R, \\ &= \frac{w_L + w_R}{2} - \nu \frac{f(w_R) - f(w_L)}{\lambda} + \nu (0, \Delta x \hat{s}_{LR}/\lambda, 0)^T. \end{aligned}$$

En écrivant composante par composante l'équation ci-dessus et en divisant de part et d'autre par  $2\nu$ , on en déduit les équations (3.100a)-(3.100b)-(3.100c).

Comme l'inégalité (3.100d) se montre par un calcul analogue, il est maintenant nécessaire de montrer l'inégalité (3.100e) à partir de (3.82b). En scindant l'intégrale figurant dans (3.82b) et en conservant la notation  $\nu = \lambda\Delta t/\Delta x$ , il suit

$$\begin{aligned} & \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \eta(\tilde{w}(x/\Delta t, \hat{w}_L, \hat{w}_R)) dx \\ &= \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{-\lambda\Delta t} \eta(w_L) dx + \frac{1}{\Delta x} \int_{-\lambda\Delta t}^0 \eta(w_L^*) dx + \frac{1}{\Delta x} \int_0^{\lambda\Delta t} \eta(w_R^*) dx + \frac{1}{\Delta x} \int_{\lambda\Delta t}^{\frac{\Delta x}{2}} \eta(w_R) dx, \\ &= (-\nu + 1/2)\eta(w_L) + \nu\eta(w_L^*) + \nu\eta(w_R^*) + (1/2 - \nu)\eta(w_R), \\ &\leq \frac{\eta(w_L) + \eta(w_R)}{2} - \nu \frac{G(w_R) - G(w_L)}{\lambda} - \mathcal{H}_{LR} \frac{\nu[z]}{\lambda}. \end{aligned}$$

En simplifiant par  $2\nu > 0$  l'expression ci-dessus puis en utilisant la quantité  $\eta^{\text{HLL}}$  définie dans (3.91b), on en déduit que l'inégalité (3.82b) est équivalente à

$$\frac{\eta(w_L^*) + \eta(w_R^*)}{2} \leq \eta^{\text{HLL}} - \mathcal{H}_{LR} \frac{[z]}{2\lambda}. \quad (3.101)$$

L'étape de simplification suivante de l'inégalité (3.101) consiste à donner une reformulation équivalente de  $-\mathcal{H}_{LR}[z]/(2\lambda)$ . Pour clarifier les écritures, la notation  $[\cdot] = \cdot_R - \cdot_L$  est utilisée et pour  $\Phi$  une fonction quelconque, cette notation est complétée par la convention  $[\Phi] = \Phi(\cdot_R) - \Phi(\cdot_L)$ . En associant la définition de  $\mathcal{H}_{LR}$  énoncée dans (3.80) à la définition du solveur de Riemann constant par morceaux (3.90), un calcul direct donne

$$\mathcal{H}_{LR} = \frac{H(w_L) + H(w_R)}{2} - \frac{\lambda}{2}[\psi] + \frac{\lambda}{2}[\psi^*].$$

Par conséquent, et en utilisant la quantité  $\hat{\eta}^{\text{HLL}}$  définie dans (3.99b), il vient

$$\begin{aligned}
& -\mathcal{H}_{LR} \frac{[z]}{2\lambda} \\
&= -\frac{[z][\psi^*]}{4} + \frac{[z]}{4} \left( [\psi] - \frac{H(w_L) + H(w_R)}{\lambda} \right), \\
&= -\frac{[z][\psi^*]}{4} + \frac{\psi(w_L)z_L + \psi(w_R)z_R}{2} - \frac{\psi(w_L) + \psi(w_R)}{2} \frac{z_L + z_R}{2} - \frac{1}{2\lambda} \frac{H(w_L) + H(w_R)}{2} [z], \\
&= -\frac{[z][\psi^*]}{4} + \frac{\psi(w_L)z_L + \psi(w_R)z_R}{2} - \frac{\psi(w_L) + \psi(w_R)}{2} \frac{z_L + z_R}{2} - \frac{1}{2\lambda} \left( [Hz] - [H] \frac{z_L + z_R}{2} \right), \\
&= -\frac{[z][\psi^*]}{4} + \hat{\eta}^{\text{HLL}} - \eta^{\text{HLL}} - \psi^{\text{HLL}} \frac{z_L + z_R}{2}.
\end{aligned}$$

En introduisant l'égalité ci-dessus dans l'expression (3.101), il est possible de déduire qu'une reformulation équivalente de l'inégalité (3.82b) est donnée par

$$\frac{\eta(w_L^*) + \eta(w_R^*)}{2} + \frac{[z][\psi^*]}{4} \leq \hat{\eta}^{\text{HLL}} - \psi^{\text{HLL}} \frac{z_L + z_R}{2}. \quad (3.102)$$

L'étape suivante consiste à récrire les quantités  $\eta(w_L^*)$  et  $\eta(w_R^*)$ . D'après la définition de  $w \mapsto \eta(w)$  exprimée dans (3.72a) et d'après un calcul analogue à (3.31), on a

$$\begin{aligned}
\frac{\eta(w_L^*) + \eta(w_R^*)}{2} &= \frac{h_L^*(u_L^*)^2 + h_R^*(u_R^*)^2}{4} + g \frac{(h_L^*)^2 \varphi(\theta_L^*) + (h_R^*)^2 \varphi(\theta_R^*)}{4}, \\
&= \frac{(h^{\text{HLL}} \hat{u}^{\text{HLL}})^2}{2h^{\text{HLL}}} + \frac{h_L^* h_R^* [u^*]^2}{8h^{\text{HLL}}} + \frac{h_L^* \psi(w_L^*) + h_R^* \psi(w_R^*)}{4}, \\
&= \frac{(h^{\text{HLL}} \hat{u}^{\text{HLL}})^2}{2h^{\text{HLL}}} + \frac{h_L^* h_R^* [u^*]^2}{8h^{\text{HLL}}} + \frac{1}{2} \frac{h_L^* + h_R^*}{2} \frac{\psi(w_L^*) + \psi(w_R^*)}{2} + \frac{[h^*][\psi^*]}{8}.
\end{aligned}$$

En utilisant l'équation (3.100a), il est possible de substituer  $h_L^* + h_R^*$  par  $2h^{\text{HLL}}$  dans l'égalité ci-dessus puis en insérant le résultat dans l'inéquation (3.102), on obtient une nouvelle forme équivalente de (3.82b) qui s'écrit

$$\frac{h_L^* h_R^* [u^*]^2}{8h^{\text{HLL}}} + \frac{[h^* + 2z][\psi^*]}{8} + \frac{h^{\text{HLL}}}{2} \frac{\psi(w_L^*) + \psi(w_R^*)}{2} \leq \hat{\eta}^{\text{HLL}} - \frac{(h^{\text{HLL}} \hat{u}^{\text{HLL}})^2}{2h^{\text{HLL}}} - \psi^{\text{HLL}} \frac{z_L + z_R}{2}.$$

Cette dernière expression correspond à la forme (3.100e) souhaitée et comme la preuve de l'implication réciproque est immédiate, elle termine également la démonstration.  $\square$

D'après la Proposition 3.2.1.1-*i-ii*), le lemme ci-dessus donne des conditions suffisantes de consistance à l'aide des équations (3.100a), (3.100b) et (3.100c) et il donne de plus des conditions suffisantes de stabilité entropique avec les inégalités (3.100d) et (3.100e). Un système d'équations non linéaires qui permet de satisfaire l'ensemble de ces conditions est maintenant proposé.

### 3.2.3 Schéma entropique et *well-balanced* pour le lac au repos

#### Définition des états $w_L^*$ et $w_R^*$ du solveur de Riemann

Dans l'objectif de satisfaire les équations et les inégalités (3.100), cette section propose un système d'équations définissant les états  $w_L^*$  et  $w_R^*$  d'un solveur de Riemann de la forme (3.90).

Pour  $\lambda > 0$  et  $\Delta t > 0$  donnés, la quantité  $w^{\text{HLL}}$  définie par (3.91a) est considérée ainsi que

$$\tilde{\Delta t}^2 = \frac{g\Delta t^2}{h^{\text{HLL}}} \left( \frac{\theta_R - \theta_L}{\theta^{\text{HLL}}} \right)^2, \quad \phi(\lambda) = e^{\frac{\lambda}{\sqrt{gh^{\text{HLL}}}} \left( \frac{\theta_R - \theta_L}{\theta^{\text{HLL}}} \right)^2}. \quad (3.103)$$

D'après les définitions ci-dessus, si  $\Delta t > 0$  alors trivialement  $\tilde{\Delta t} = 0$  si et seulement si  $\theta_L = \theta_R$  et dans ce cas,  $\phi(\lambda) = 1$  pour tout  $\lambda$ . Si  $\theta_L \neq \theta_R$  alors  $\tilde{\Delta t}^2$  est proportionnel à  $\Delta t^2$  et  $\phi$  tend vers l'infini



lorsque  $\lambda$  tend vers l'infini. Avec ces notations et en considérant la fonction  $w \mapsto \psi(w)$  et la quantité  $h^{\text{HLL}}\hat{u}^{\text{HLL}}$  respectivement définies dans (3.72b) et dans (3.99a), on propose de définir les états  $w_L^*$  et  $w_R^*$  d'un solveur de Riemann (3.90) par les équations suivantes :

$$\frac{h_L^* + h_R^*}{2} = h^{\text{HLL}}, \quad (3.104a)$$

$$\frac{h_L^* u_L^* + h_R^* u_R^*}{2} = h^{\text{HLL}} \hat{u}^{\text{HLL}}, \quad (3.104b)$$

$$\frac{h_L^* \theta_L^* + h_R^* \theta_R^*}{2} = h^{\text{HLL}} \theta^{\text{HLL}}, \quad (3.104c)$$

$$u_L^* = u_R^*, \quad (3.104d)$$

$$\frac{\psi(w_L^*) + \psi(w_R^*)}{2} = \psi(w^{\text{HLL}})(1 + \tilde{\Delta}t^2), \quad (3.104e)$$

$$h_R^* - h_L^* + \frac{h_L^* h_R^*}{h^{\text{HLL}}} \frac{\varphi(\theta_R^*) - \varphi(\theta_L^*)}{\varphi(\theta^{\text{HLL}})} = -\phi(\lambda)(h_R^* - h_L^* + 2z_R - 2z_L). \quad (3.104f)$$

Les équations (3.104a)-(3.104b)-(3.104c) sont exactement les trois premières équations de (3.100) et elles s'identifient à la relation de consistance intégrale. Les deux équations (3.104e)-(3.104f) sont paramétrées par  $\lambda > 0$  et  $\Delta t > 0$  et sont destinées à assurer les inégalités (3.100d)-(3.100e). Avant de démontrer ces résultats, il est d'abord nécessaire d'étudier le système d'équations (3.104).

Pour conduire l'étude du système d'équations (3.104), il est utile de définir les ensembles

$$\mathcal{U} = \{(\hat{w}_L, \hat{w}_R) \in \hat{\Omega}^2 \mid \theta_L \neq \theta_R\},$$

$$\mathcal{U}^c = \{(\hat{w}_L, \hat{w}_R) \in \hat{\Omega}^2 \mid \theta_L = \theta_R\},$$

$$\mathcal{V} = \{(\sigma^*, \zeta^*) \in \mathbb{R}^2 \mid \sigma^* \in \mathbb{R}, \theta^{\text{HLL}} - (1 + \sigma^*)\zeta^* > 0, \theta^{\text{HLL}} + (1 - \sigma^*)\zeta^* > 0\}.$$

Le premier lemme traite de deux cas particuliers pour lesquels une solution explicite du système (3.104) peut être exhibée.

**Lemme 3.2.3.1** (Solution explicite partielle de (3.104)). *Soient  $\hat{w}_L, \hat{w}_R$  deux états constants de  $\hat{\Omega}$ . Considérons les quantités  $w^{\text{HLL}}$  et  $h^{\text{HLL}}\hat{u}^{\text{HLL}}$  respectivement définies par (3.91a) et par (3.99a). Supposons que la condition CFL de non interaction (3.89) est satisfaite. Les assertions suivantes sont vérifiées.*

i) *Les quantités  $u_L^*$  et  $u_R^*$  du système d'équations (3.104) sont toujours explicitement connues et leurs expressions sont données par*

$$u_L^* = u_R^* = \hat{u}^{\text{HLL}}. \quad (3.105)$$

ii) *Si  $(\hat{w}_L, \hat{w}_R)$  est dans  $\mathcal{U}^c$  alors une solution du système d'équations (3.104) est donnée par*

$$\theta_L^* = \theta_R^* = \theta^{\text{HLL}}, \quad (3.106a)$$

$$h_L^* = h^{\text{HLL}} + \frac{z_R - z_L}{2}, \quad (3.106b)$$

$$h_R^* = h^{\text{HLL}} - \frac{z_R - z_L}{2}. \quad (3.106c)$$

iii) *Si  $\Delta t$  est nul alors une solution du système d'équations (3.104) est donnée par*

$$\theta_L^* = \theta_R^* = \theta^{\text{HLL}}, \quad (3.107a)$$

$$h_L^* = h^{\text{HLL}} + \phi(\lambda) \frac{z_R - z_L}{1 + \phi(\lambda)}, \quad (3.107b)$$

$$h_R^* = h^{\text{HLL}} - \phi(\lambda) \frac{z_R - z_L}{1 + \phi(\lambda)}. \quad (3.107c)$$

*Démonstration.* Pour l'assertion *i*), puisque  $u_L^*$  et  $u_R^*$  sont égaux d'après (3.104d), notons  $u^* = u_L^* = u_R^*$ . Dans ce cas, les équations (3.104a) et (3.104b) permettent d'écrire

$$\begin{aligned} h^{\text{HLL}} \hat{u}^{\text{HLL}} &= \frac{h_L^* u_L^* + h_R^* u_R^*}{2}, \\ &= \frac{h_L^* + h_R^*}{2} u^*, \\ &= h^{\text{HLL}} u^*. \end{aligned} \tag{3.108}$$

Puisque que la condition CFL (3.89) est satisfaite, l'inégalité stricte  $h^{\text{HLL}} > 0$  est vérifiée. Par conséquent, une multiplication de l'équation (3.108) par  $1/h^{\text{HLL}}$  produit le résultat attendu quant à  $u_L^*$  et  $u_R^*$ .

Pour montrer l'assertion *ii*), il est suffisant de montrer que les expressions proposées dans (3.106) satisfont les quatre équations (3.104a)-(3.104c)-(3.104e)-(3.104f) lorsque  $\theta_L = \theta_R$ .

En sommant les expressions de  $h_L^*$  et de  $h_R^*$  respectivement données dans (3.106b) et dans (3.106c), on en déduit que l'équation (3.104a) est vérifiée. Les expressions  $\theta_L^* = \theta_R^* = \theta^{\text{HLL}}$  complétées par l'équation (3.104a) permettent de maintenir l'équation (3.104c). D'après la quantité  $\hat{\Delta}t$  définie par (3.103), l'équation (3.104e) s'écrit dans le cas  $\theta_L = \theta_R$  de la manière suivante :

$$\frac{\psi(w_L^*) + \psi(w_R^*)}{2} = \psi(w^{\text{HLL}}). \tag{3.109}$$

En utilisant la définition de  $w \mapsto \psi(w)$  donnée dans (3.72b) et les égalités  $\theta_L^* = \theta_R^* = \theta^{\text{HLL}}$ , le membre de gauche de l'équation ci dessus s'écrit

$$\begin{aligned} \frac{\psi(w_L^*) + \psi(w_R^*)}{2} &= g \frac{h_L^* \varphi(\theta_L^*) + h_R^* \varphi(\theta_R^*)}{2}, \\ &= g \frac{h_L^* + h_R^*}{2} \varphi(\theta^{\text{HLL}}). \end{aligned} \tag{3.110}$$

L'égalité ci-dessus associée à l'équation (3.104a) permet de déduire que l'équation (3.104e) est vérifiée sous la forme de (3.109). Enfin, et puisque  $\theta_L = \theta_R = \theta^{\text{HLL}} = \theta_L^* = \theta_R^*$ , l'équation (3.104f) se lit

$$h_R^* - h_L^* = -(h_R^* - h_L^*) - 2(z_R - z_L).$$

La différence des quantités  $h_R^*$  et  $h_L^*$  données dans (3.106c)-(3.106b) satisfait l'équation ci-dessus ce qui termine de montrer l'assertion *ii*).

Concernant l'assertion *iii*), il suffit de montrer que les expressions proposées dans (3.107) permettent de satisfaire les quatre équations (3.104a)-(3.104c)-(3.104e)-(3.104f) dans le cas  $\Delta t = 0$ .

La somme des expressions de  $h_L^*$  et de  $h_R^*$  respectivement données dans (3.107b) et dans (3.107c) montre immédiatement que l'équation (3.104a) est satisfaite. Les égalités  $\theta_L^* = \theta_R^* = \theta^{\text{HLL}}$  associées à l'équation (3.104a) permettent de satisfaire l'équation (3.104c). Dans le cas  $\Delta t = 0$ , l'équation (3.104e) est une nouvelle fois équivalente à (3.109). Par conséquent, le calcul (3.110) est de nouveau valable et il permet de déduire que l'équation (3.104e) est vérifiée. Finalement, et comme  $\theta_L^* = \theta_R^* = \theta^{\text{HLL}}$ , l'équation (3.104f) s'écrit

$$h_R^* - h_L^* = -\phi(\lambda)(h_R^* - h_L^* + 2(z_R - z_L)).$$

En faisant la différence des quantités  $h_R^*$  et  $h_L^*$  données dans (3.107c)-(3.107b) on en déduit que l'équation ci-dessus est satisfaite ce qui termine la preuve. □

Le lemme ci-dessus montre que les quantités  $u_L^*$  et  $u_R^*$  qui satisfont le système d'équations (3.104) sont toujours explicitement connues et leurs expressions sont données par (3.105).

Par conséquent, la difficulté dans la résolution du système d'équations (3.104) réside dans la détermination d'un quadruplet  $h_L^*, h_R^*, \theta_L^*, \theta_R^*$  satisfaisant les quatre équations (3.104a)-(3.104c)-(3.104e)-(3.104f). Le lemme suivant établit une reformulation équivalente de ces quatre équations.

**Lemme 3.2.3.2** (Reformulation équivalente des équations (3.104a)-(3.104c)-(3.104e)-(3.104f)). Soient  $\hat{w}_L, \hat{w}_R$  deux états constants de  $\hat{\Omega}$ . Considérons la quantité  $w^{\text{HLL}}$  définie dans (3.91a) ainsi que le couple  $(\sigma^*, \zeta^*)$  dans  $\mathbb{R}^2$  tel que

$$\sigma^* = \frac{h_R^* - h_L^*}{2h^{\text{HLL}}}, \quad \zeta^* = \frac{\theta_R^* - \theta_L^*}{2}. \quad (3.111)$$

Il y a équivalence entre les quatre équations (3.104a)-(3.104c)-(3.104e)-(3.104f) et le système suivant :

$$h^{\text{HLL}} = \frac{h_L^* + h_R^*}{2}, \quad (3.112a)$$

$$\theta^{\text{HLL}} = \frac{\theta_L^* + \theta_R^*}{2} + \frac{\sigma^* \zeta^*}{2}, \quad (3.112b)$$

$$2\varphi(\theta^{\text{HLL}})(1 + \tilde{\Delta}t^2) = \varphi(\theta^{\text{HLL}} - (1 + \sigma^*)\zeta^*)(1 - \sigma^*) + \varphi(\theta^{\text{HLL}} + (1 - \sigma^*)\zeta^*)(1 + \sigma^*) \quad (3.112c)$$

$$\begin{aligned} -\phi(\lambda) \frac{z_R - z_L}{h^{\text{HLL}}} &= \sigma^*(1 + \phi(\lambda)) \\ &+ \frac{1 - (\sigma^*)^2}{2\varphi(\theta^{\text{HLL}})} (\varphi(\theta^{\text{HLL}} + (1 - \sigma^*)\zeta^*) - \varphi(\theta^{\text{HLL}} - (1 + \sigma^*)\zeta^*)). \end{aligned} \quad (3.112d)$$

*Démonstration.* Puisque les équations (3.104a) et (3.112a) sont identiques, il faut seulement montrer les trois équations (3.112b)-(3.112c)-(3.112d). Pour ce faire, la notation  $[\cdot] = \cdot_R - \cdot_L$  est adoptée.

Pour l'équation (3.112b), en divisant l'équation (3.104c) par  $h^{\text{HLL}} > 0$  puis en introduisant l'équation (3.104a) dans le résultat, il suit

$$\begin{aligned} \theta^{\text{HLL}} &= \frac{h_L^* \theta_L^* + h_R^* \theta_R^*}{2h^{\text{HLL}}}, \\ &= \frac{1}{h^{\text{HLL}}} \left( \frac{h_L^* + h_R^*}{2} \frac{\theta_L^* + \theta_R^*}{2} + \frac{[h^*][\theta^*]}{4} \right), \\ &= \frac{\theta_L^* + \theta_R^*}{2} + \frac{[h^*][\theta^*]}{4h^{\text{HLL}}}. \end{aligned} \quad (3.113)$$

L'utilisation des définitions de  $\sigma^*$  et de  $\zeta^*$  données par (3.111) dans l'égalité ci-dessus permet de déduire l'équation (3.112b).

En ce qui concerne l'équation (3.112c), en utilisant un calcul analogue à (3.113) et l'expression de  $w \mapsto \psi(w)$  donnée dans (3.72b), l'équation (3.104e) s'écrit

$$\begin{aligned} 2\varphi(\theta^{\text{HLL}})(1 + \tilde{\Delta}t^2) &= \varphi(\theta_L^*) + \varphi(\theta_R^*) + \frac{[h^*](\varphi(\theta_R^*) - \varphi(\theta_L^*))}{2h^{\text{HLL}}}, \\ &=: (1 + \sigma^*)\varphi(\theta_R^*) + (1 - \sigma^*)\varphi(\theta_L^*). \end{aligned} \quad (3.114)$$

Par ailleurs, un calcul direct utilisant l'équation (3.112b) permet d'obtenir

$$\begin{aligned} \theta_L^* &= \frac{\theta_L^* + \theta_R^*}{2} - \frac{[\theta^*]}{2}, \\ &= \theta^{\text{HLL}} - \left(1 + \frac{[h^*]}{2h^{\text{HLL}}}\right) \frac{[\theta^*]}{2}, \\ &= \theta^{\text{HLL}} - (1 + \sigma^*)\zeta^*. \end{aligned} \quad (3.115)$$

De la même manière, il suit

$$\theta_R^* = \theta^{\text{HLL}} + (1 - \sigma^*)\zeta^*. \quad (3.116)$$

L'introduction des deux équations (3.115)-(3.116) dans l'expression (3.114) permet d'obtenir l'équation (3.112c).

Pour l'équation (3.112d), en divisant (3.104f) par  $2h^{\text{HLL}} > 0$  et en utilisant l'équation (3.112a), il suit

$$\begin{aligned}
-\phi(\lambda) \frac{[z]}{h^{\text{HLL}}} &= (1 + \phi(\lambda)) \frac{[h^*]}{2h^{\text{HLL}}} + \frac{h_L^* h_R^*}{(h^{\text{HLL}})^2} \frac{\varphi(\theta_R^*) - \varphi(\theta_L^*)}{2\varphi(\theta^{\text{HLL}})}, \\
&= (1 + \phi(\lambda)) \sigma^* + \frac{(h_L^* + h_R^*)^2 - [h^*]^2}{(2h^{\text{HLL}})^2} \frac{\varphi(\theta_R^*) - \varphi(\theta_L^*)}{2\varphi(\theta^{\text{HLL}})}, \\
&= (1 + \phi(\lambda)) \sigma^* + (1 - (\sigma^*)^2) \frac{\varphi(\theta_R^*) - \varphi(\theta_L^*)}{2\varphi(\theta^{\text{HLL}})}.
\end{aligned}$$

En insérant les égalités (3.115)-(3.116) dans l'équation ci-dessus on en déduit l'expression (3.112d) ce qui termine de montrer que (3.104a)-(3.104c)-(3.104e)-(3.104f) impliquent le système d'équations (3.112). Comme l'implication réciproque est immédiate ceci termine la preuve du lemme.  $\square$

D'après le lemme ci-dessus, il est clair que si  $\zeta^*$  et  $\sigma^*$  définis dans (3.111) sont connus alors les quantités  $h_L^*$ ,  $h_R^*$ ,  $\theta_L^*$  et  $\theta_R^*$  s'obtiennent facilement avec les équations (3.112a)-(3.112b). Par conséquent, la difficulté initiale qui reposait sur la détermination du quadruplet  $h_L^*$ ,  $h_R^*$ ,  $\theta_L^*$ ,  $\theta_R^*$  se réduit avec le Lemme 3.2.3.2, à la détermination de  $\sigma^*$  et de  $\zeta^*$  vérifiant les équations (3.112c)-(3.112d).

Si un tel doublet est explicitement exhibée dans le Lemme 3.2.3.1 pour le cas  $\theta_L = \theta_R$ , les recherches menées n'ont pas permis d'établir l'existence de ce doublet dans le cas  $\theta_L \neq \theta_R$ . Comme la solution explicite du Lemme 3.2.3.1-ii) est singulière, il est malheureusement impossible d'appliquer le Théorème des Fonctions Implicites. Néanmoins, les très nombreuses résolutions numériques du système (3.112c)-(3.112d) réalisées dans le cas  $\theta_L \neq \theta_R$  et avec une méthode de Newton ou de point fixe n'ont jamais montré de défaut de convergence. Par conséquent, il semble légitime de supposer l'existence de  $(\sigma^*, \zeta^*)$  dans  $\mathbb{R}^2$  tel que les équations (3.112c)-(3.112d) sont satisfaites. Dans la suite de ce mémoire de thèse, cette existence ainsi que le caractère bien posé des équations (3.112c)-(3.112d) sont donc admis et on suppose de plus que  $\sigma^*$  est tel que  $h_L^* > 0$  et  $h_R^* > 0$ . Ces hypothèses s'écrivent sous la forme de la proposition suivante :

i)  $\forall (\hat{w}_L, \hat{w}_R) \in \hat{\Omega}^2, \exists (\sigma^*, \zeta^*, \lambda, \Delta t) \in ]-1, 1[ \times \mathbb{R} \times (\mathbb{R}_*^+)^2, \text{ tel que}$

$$\left\{ \begin{array}{l} 2\varphi(\theta^{\text{HLL}})(1 + \tilde{\Delta}t^2) = \\ \quad \varphi(\theta^{\text{HLL}} - (1 + \sigma^*)\zeta^*)(1 - \sigma^*) + \varphi(\theta^{\text{HLL}} + (1 - \sigma^*)\zeta^*)(1 + \sigma^*), \\ -\phi(\lambda) \frac{z_R - z_L}{h^{\text{HLL}}} = \\ \quad \sigma^*(1 + \phi(\lambda)) + \frac{1 - (\sigma^*)^2}{2\varphi(\theta^{\text{HLL}})} (\varphi(\theta^{\text{HLL}} + (1 - \sigma^*)\zeta^*) - \varphi(\theta^{\text{HLL}} - (1 + \sigma^*)\zeta^*)), \end{array} \right. \quad (3.117a)$$

ii)  $(\hat{w}_L, \hat{w}_R, \lambda, \Delta t) \mapsto (\sigma^*, \zeta^*) \in C^0(\hat{\Omega}^2 \times (\mathbb{R}_*^+)^2, \mathbb{R}^2)$ . (3.117b)

Avec les hypothèses précédentes, il est possible de caractériser les solutions du système d'équations (3.104). Cette caractérisation est donnée dans le lemme ci-dessous et elle sera nécessaire dans la section suivante qui concerne la stabilité entropique.

**Lemme 3.2.3.3** (Propriétés des solutions du système d'équations (3.104)). *Soient  $(\hat{w}_L, \hat{w}_R)$  dans  $\mathcal{U}$  i.e  $\theta_L \neq \theta_R$  et  $\tilde{w}(\cdot, \hat{w}_L, \hat{w}_R) : \mathbb{R} \rightarrow \mathbb{R}^3$  un solveur de Riemann de la forme (3.90). Soit  $\varphi : \mathbb{R}_*^+ \rightarrow \mathbb{R}_*^+$  une fonction régulière, strictement convexe, bijective et satisfaisant les inégalités (3.74). Supposons que les hypothèses (3.117) sont vérifiées et notons  $w_L^*$ ,  $w_R^*$  une solution du système d'équations (3.104). Les assertions suivantes sont vérifiées.*

i) Si  $\Delta t \rightarrow 0$  alors  $w_L^*$  et  $w_R^*$  converge simplement vers les expressions (3.107).

ii) Si  $z_L$  et  $z_R$  sont différents alors  $h_R^* - h_L^* + 2(z_R - z_L)$  est différent de zero.

*Démonstration.* Concernant l'assertion *i*), comme les expressions de  $u_L^*$  et de  $u_R^*$  sont connues d'après le Lemme 3.2.3.1-*i*), il est seulement nécessaire de montrer

$$(\theta_L^*, \theta_R^*, h_L^*, h_R^*)^T \xrightarrow{\Delta t \rightarrow 0} \left( \theta^{\text{HLL}}, \theta^{\text{HLL}}, h^{\text{HLL}} + \phi(\lambda) \frac{z_R - z_L}{1 + \phi(\lambda)}, h^{\text{HLL}} - \phi(\lambda) \frac{z_R - z_L}{1 + \phi(\lambda)} \right)^T.$$

Pour le couple  $(\theta_L^*, \theta_R^*)$ , et puisque les hypothèses (3.117) sont vérifiées, le système d'équations (3.104) admet des solutions réelles. Ces solutions dépendent continûment des données et elles sont telles que  $h_L^* > 0$  et  $h_R^* > 0$ . En s'appuyant sur la régularité de  $\varphi$ , il est donc possible d'utiliser un développement de Taylor avec reste intégral pour récrire certaines équations du système (3.104). En associant un tel développement à l'équation  $h_L^* + h_R^* = 2h^{\text{HLL}}$ , une solution du système (3.104) vérifie la réécriture de l'équation (3.104e) qui suit :

$$\begin{aligned} \frac{\psi(w_L^*) + \psi(w_R^*)}{2} &= g \frac{h_L^* \varphi(\theta_L^*) + h_R^* \varphi(\theta_R^*)}{2}, \\ &= g \frac{h_L^* (\varphi(\theta_L^*) - \varphi(\theta^{\text{HLL}})) + h_R^* (\varphi(\theta_R^*) - \varphi(\theta^{\text{HLL}}))}{2} + g \frac{h_L^* + h_R^*}{2} \varphi(\theta^{\text{HLL}}), \\ &= \frac{gh_L^*}{2} \left( \varphi'(\theta^{\text{HLL}})(\theta_L^* - \theta^{\text{HLL}}) + (\theta_L^* - \theta^{\text{HLL}})^2 \int_0^1 (1-s) \varphi''(\theta^{\text{HLL}} + s(\theta_L^* - \theta^{\text{HLL}})) ds \right) \\ &\quad + \frac{gh_R^*}{2} \left( \varphi'(\theta^{\text{HLL}})(\theta_R^* - \theta^{\text{HLL}}) + (\theta_R^* - \theta^{\text{HLL}})^2 \int_0^1 (1-s) \varphi''(\theta^{\text{HLL}} + s(\theta_R^* - \theta^{\text{HLL}})) ds \right) \\ &\quad + \psi(w^{\text{HLL}}), \\ &= \psi(w^{\text{HLL}}) + g \varphi'(\theta^{\text{HLL}}) \left( \frac{h_R \theta_R^* + h_L^* \theta_L^*}{2} - h^{\text{HLL}} \theta^{\text{HLL}} \right) \\ &\quad + \frac{gh_R^*}{2} (\theta_R^* - \theta^{\text{HLL}})^2 \int_0^1 (1-s) \varphi''(\theta^{\text{HLL}} + s(\theta_R^* - \theta^{\text{HLL}})) ds \\ &\quad + \frac{gh_L^*}{2} (\theta_L^* - \theta^{\text{HLL}})^2 \int_0^1 (1-s) \varphi''(\theta^{\text{HLL}} + s(\theta_L^* - \theta^{\text{HLL}})) ds. \end{aligned}$$

En insérant l'équation (3.104c) dans le résultat ci-dessus, l'équation (3.104e) se récrit de manière équivalente sous la forme

$$\sum_{\alpha \in \{L, R\}} h_\alpha^* (\theta_\alpha^* - \theta^{\text{HLL}})^2 \int_0^1 (1-s) \varphi''(\theta^{\text{HLL}} + s(\theta_\alpha^* - \theta^{\text{HLL}})) ds = \psi(w^{\text{HLL}}) \tilde{\Delta t}^2. \quad (3.118)$$

Comme  $(\hat{w}_L, \hat{w}_R)$  est dans  $\mathcal{U}$ ,  $\tilde{\Delta t}^2$  défini dans (3.103) est proportionnel à  $\Delta t^2$  et par conséquent

$$\sum_{\alpha \in \{L, R\}} h_\alpha^* (\theta_\alpha^* - \theta^{\text{HLL}})^2 \int_0^1 (1-s) \varphi''(\theta^{\text{HLL}} + s(\theta_\alpha^* - \theta^{\text{HLL}})) ds \xrightarrow{\Delta t \rightarrow 0} 0.$$

Puisque les quantités  $h_L^*$  et  $h_R^*$  sont par hypothèse strictement positives et que  $\theta \mapsto \varphi(\theta)$  est strictement convexe, le membre de gauche de l'expression ci-dessus est constitué d'une somme de deux termes positifs. Par conséquent, il suit nécessairement

$$h_\alpha^* (\theta_\alpha^* - \theta^{\text{HLL}})^2 \int_0^1 (1-s) \varphi''(\theta^{\text{HLL}} + s(\theta_\alpha^* - \theta^{\text{HLL}})) ds \xrightarrow{\Delta t \rightarrow 0} 0, \quad \forall \alpha \in \{L, R\}.$$

En arguant de nouveau  $h_L^* > 0$  et  $h_R^* > 0$  et comme que les intégrales ci dessus ne peuvent pas être nulles en vertu de la convexité stricte de  $\theta \mapsto \varphi(\theta)$ , on en déduit

$$\theta_\alpha^* - \theta^{\text{HLL}} \xrightarrow{\Delta t \rightarrow 0} 0, \quad \forall \alpha \in \{L, R\}.$$

Ainsi,  $\theta_L^*$  et  $\theta_R^*$  tendent tous deux vers la même limite  $\theta^{\text{HLL}}$  ce qui correspond à l'expression (3.107a).

Pour les expressions limites de  $h_L^*$  et de  $h_R^*$  données par (3.107b)-(3.107c), en utilisant le résultat ci-dessus et la régularité de  $\varphi$ , l'équation limite de (3.104f) lorsque  $\Delta t \rightarrow 0$  s'écrit

$$h_R^* - h_L^* = -2\phi(\lambda) \frac{z_R - z_L}{1 + \phi(\lambda)}.$$

L'égalité précédente associée à l'équation (3.104a) permet de conclure la preuve de l'assertion *i*).

Pour montrer l'assertion *ii*), raisonnons par l'absurde et supposons que  $z_R \neq z_L$  et  $h_R^* - h_L^* + 2(z_R - z_L) = 0$ . Dans ce cas, l'équation (3.104f) se lit

$$h_R^* - h_L^* + \frac{h_L^* h_R^* \varphi(\theta_R^*) - \varphi(\theta_L^*)}{h^{\text{HLL}} \varphi(\theta^{\text{HLL}})} = 0. \quad (3.119)$$

Comme  $\theta_L \neq \theta_R$ ,  $\tilde{\Delta t}^2$  est proportionnel à  $\Delta t^2$  or l'équation ci-dessus est en particulier valable lorsque  $\Delta t \rightarrow 0$  et *ipso facto*  $\tilde{\Delta t} \rightarrow 0$ . Dans ce cas, l'assertion *i*) montre que  $\theta_L^*, \theta_R^* \rightarrow \theta^{\text{HLL}}$ . Par conséquent, et comme  $\varphi$  est régulière, lorsque  $\Delta t \rightarrow 0$  l'équation (3.119) tend vers

$$h_R^* - h_L^* = 0.$$

Le résultat ci-dessus introduit dans  $h_R^* - h_L^* + 2(z_R - z_L) = 0$  conduit à conclure que  $z_L$  tend nécessairement vers  $z_R$  ce qui est absurde.  $\square$

Avec les propriétés décrites dans le lemme ci-dessus, il est maintenant possible de détailler un schéma de type Godunov entropique pour le couple  $(\hat{\eta}, \hat{G})$  défini dans (3.71) et *well-balanced* pour le lac au repos.

### Un schéma de type Godunov entropique et *well-balanced* pour le lac au repos

Le schéma de type Godunov proposé est défini par un solveur de Riemann de la forme (3.90) avec des états intermédiaires  $w_L^*$  et  $w_R^*$  qui vérifient le système (3.104). Ce système contient les équations de consistance et il est paramétré par les réels  $\lambda > 0$ ,  $\Delta t > 0$  et par la quantité  $\hat{s}_{LR}$ .

Cette dernière quantité est un paramètre libre qui intervient dans la définition de  $\hat{u}^{\text{HLL}}$  donnée par (3.99a) et qui pour rappel s'écrit

$$h^{\text{HLL}} \hat{u}^{\text{HLL}} = (hu)^{\text{HLL}} + \frac{\Delta x \hat{s}_{LR}}{2\lambda}.$$

D'après la section 3.2.1, le choix retenu pour  $\hat{s}_{LR}$  doit être consistant avec  $\frac{1}{\Delta x} \int_{-\Delta x/2}^{\Delta x/2} (-gh\varphi(\theta)\partial_x z) dx$  au sens de l'expression (3.77). Le lemme qui suit propose une expression consistante de  $\hat{s}_{LR}$  qui, accompagnée d'un choix convenable de  $\lambda > 0$ ,  $\Delta t > 0$ , permet aux équations (3.104e)-(3.104f) de satisfaire les inégalités (3.100d)-(3.100e).

**Lemme 3.2.3.4** (Satisfaction des inégalités (3.100d)-(3.100e)). *Soient  $\hat{w}_L, \hat{w}_R$  deux états constants de  $\hat{\Omega}$  ainsi que  $\varphi : \mathbb{R}_*^+ \rightarrow \mathbb{R}_*^+$  une fonction régulière, strictement convexe, bijective et satisfaisant les inégalités (3.74). Considérons le couple  $(\psi, H)$  défini dans (3.72b). Il existe  $\lambda > 0$  suffisamment grand et  $\Delta t > 0$  suffisamment petit tels que si les hypothèses (3.117) sont vérifiées, si la condition CFL (3.89) est satisfaite et si*

$$\hat{s}_{LR} = -gh^{\text{HLL}} \varphi(\theta^{\text{HLL}}) \frac{z_R - z_L}{\Delta x} - \frac{H(w_L) + H(w_R)}{\sqrt{g}(h^{\text{HLL}})^{3/2}} \frac{(z_R - z_L)^2}{\Delta x}, \quad (3.120)$$

alors, les inégalités suivantes sont assurées :

$$\frac{\psi(w_L^*) + \psi(w_R^*)}{2} \leq \psi^{\text{HLL}}, \quad (3.121a)$$

$$\begin{aligned} \frac{h_L^* h_R^*}{8h^{\text{HLL}}} (u_R^* - u_L^*)^2 + \frac{h^{\text{HLL}}}{2} \frac{\psi(w_L^*) + \psi(w_R^*)}{2} + (h_R^* - h_L^* + 2(z_R - z_L)) \frac{\psi(w_R^*) - \psi(w_L^*)}{8} \leq \\ \hat{\eta}^{\text{HLL}} - \frac{(h^{\text{HLL}} \hat{u}^{\text{HLL}})^2}{2h^{\text{HLL}}} - \psi^{\text{HLL}} \frac{z_L + z_R}{2}. \end{aligned} \quad (3.121b)$$

*Démonstration.* Le première inégalité (3.121a) se montre par une disjonction des cas  $\theta_L = \theta_R$  puis  $\theta_L \neq \theta_R$ . Dans le cas où  $\theta_L$  et  $\theta_R$  sont égaux, les Lemmes 3.2.2.1-ii) et 3.2.3.1-i) donnent respectivement  $\psi(w^{\text{HLL}}) = \psi^{\text{HLL}}$  et  $\theta_L^* = \theta_R^* = \theta^{\text{HLL}}$ . Par conséquent, et d'après la définition de  $w \mapsto \psi(w)$  énoncée dans (3.72b), il suit

$$\begin{aligned} \frac{\psi(w_L^*) + \psi(w_R^*)}{2} &= g \frac{h_L^* + h_R^*}{2} \varphi(\theta), \\ &= gh^{\text{HLL}} \varphi(\theta^{\text{HLL}}), \\ &= \psi(w^{\text{HLL}}) = \psi^{\text{HLL}}. \end{aligned}$$

Ainsi, d'après l'égalité ci-dessus, lorsque  $\theta_L = \theta_R$ , l'inégalité (3.121a) est satisfaite au sens de l'égalité. Dans le cas  $\theta_L \neq \theta_R$ ,  $\tilde{\Delta}t$  défini par (3.103) est proportionnel à  $\Delta t > 0$  et l'équation (3.104e) se lit

$$\frac{\psi(w_L^*) + \psi(w_R^*)}{2} = \psi(w^{\text{HLL}})(1 + \tilde{\Delta}t^2). \quad (3.122)$$

De plus, et d'après le Lemme 3.2.2.1-ii), la convexité stricte de  $\theta \mapsto \varphi(\theta)$  impose l'inégalité stricte  $\psi(w^{\text{HLL}}) < \psi^{\text{HLL}}$ . Par conséquent, et puisque  $0 < \psi(w^{\text{HLL}})$ , il existe  $\Delta t > 0$  assez petit tel que  $\tilde{\Delta}t$  vérifie

$$0 < \psi(w^{\text{HLL}})\tilde{\Delta}t^2 \leq \psi^{\text{HLL}} - \psi(w^{\text{HLL}}).$$

L'association de cette dernière inégalité à l'équation (3.122) donne

$$\frac{\psi(w_L^*) + \psi(w_R^*)}{2} = \psi(w^{\text{HLL}})(1 + \tilde{\Delta}t^2) \leq \psi^{\text{HLL}}.$$

Le résultat ci-dessus correspond à l'inégalité (3.121a).

En ce qui concerne l'inégalité (3.121b), le premier objectif est d'obtenir une reformulation de cette inégalité à l'aide du système d'équations (3.104). Pour effectuer cette reformulation, la notation  $[\cdot] = \cdot_R - \cdot_L$  est adoptée et elle est complétée par  $[\varphi^*] = \varphi(\theta_R^*) - \varphi(\theta_L^*)$ .

En utilisant l'équation (3.104a) ainsi que le calcul (3.114) qui donne une expression de  $\varphi(\theta_L^*) + \varphi(\theta_R^*)$ , il suit

$$\begin{aligned} \frac{\psi(w_R^*) - \psi(w_L^*)}{g} &= [h^*] \frac{\varphi(\theta_L^*) + \varphi(\theta_R^*)}{2} + \frac{h_L^* + h_R^*}{2} [\varphi^*], \\ &= [h^*] \left( \varphi(\theta^{\text{HLL}}) + \varphi(\theta^{\text{HLL}})\tilde{\Delta}t^2 - \frac{[h^*][\varphi^*]}{4h^{\text{HLL}}} \right) + h^{\text{HLL}}[\varphi^*], \\ &= \varphi(\theta^{\text{HLL}}) \left( [h^*] + h^{\text{HLL}} \left( 1 - \left( \frac{[h^*]}{2h^{\text{HLL}}} \right)^2 \right) \frac{[\varphi^*]}{\varphi(\theta^{\text{HLL}})} \right) + [h^*] \varphi(\theta^{\text{HLL}})\tilde{\Delta}t^2, \\ &= \varphi(\theta^{\text{HLL}}) \left( [h^*] + \frac{h_R^* h_L^*}{h^{\text{HLL}}} \frac{[\varphi^*]}{\varphi(\theta^{\text{HLL}})} \right) + [h^*] \varphi(\theta^{\text{HLL}})\tilde{\Delta}t^2. \end{aligned}$$

En utilisant les équations (3.104d)-(3.104e)-(3.104f) et le résultat ci-dessus, on en déduit que le membre de gauche de l'inégalité (3.121b) s'écrit

$$\begin{aligned} &\frac{h_L^* h_R^* [u^*]^2}{8h^{\text{HLL}}} + \frac{h^{\text{HLL}}}{2} \frac{\psi(w_L^*) + \psi(w_R^*)}{2} + [h^* + 2z] \frac{\psi(w_R^*) - \psi(w_L^*)}{8} \\ &= \frac{h^{\text{HLL}}}{2} \left( \psi(w^{\text{HLL}})(1 + \tilde{\Delta}t^2) \right) + \frac{g}{8} \varphi(\theta^{\text{HLL}}) [h^* + 2z] \left( [h^*] + \frac{h_R^* h_L^*}{h^{\text{HLL}}} \frac{[\varphi^*]}{\varphi(\theta^{\text{HLL}})} + [h^*] \tilde{\Delta}t^2 \right), \\ &= \frac{h^{\text{HLL}}}{2} \psi(w^{\text{HLL}}) - \frac{g}{8} \varphi(\theta^{\text{HLL}}) \phi(\lambda) [h^* + 2z]^2 + \frac{g}{8} \varphi(\theta^{\text{HLL}}) \tilde{\Delta}t^2 (4(h^{\text{HLL}})^2 + [h^* + 2z][h^*]), \\ &= \frac{h^{\text{HLL}}}{2} \psi(w^{\text{HLL}}) - \frac{g}{8} \varphi(\theta^{\text{HLL}}) \phi(\lambda) [h^* + 2z]^2 + \frac{g}{4} \varphi(\theta^{\text{HLL}}) \tilde{\Delta}t^2 ((h_L^*)^2 + (h_R^*)^2 + [z][h^*]). \end{aligned}$$

Par conséquent, l'inégalité (3.121b) est équivalente à

$$-\frac{g}{8}\varphi(\theta^{\text{HLL}})\phi(\lambda)[h^* + 2z]^2 + \frac{g}{4}\varphi(\theta^{\text{HLL}})\tilde{\Delta}t^2((h_L^*)^2 + (h_R^*)^2 + [z][h^*]) \leq \hat{\eta}^{\text{HLL}} - \frac{(h^{\text{HLL}}\hat{u}^{\text{HLL}})^2}{2h^{\text{HLL}}} - \frac{h^{\text{HLL}}}{2}\psi(w^{\text{HLL}}) - \psi^{\text{HLL}}\frac{z_L + z_R}{2}. \quad (3.123)$$

Cette écriture de l'inégalité (3.121b) acquise, considérons tout d'abord le cas  $\theta_L = \theta_R$ . Dans ce cas,  $\psi(w^{\text{HLL}}) = \psi^{\text{HLL}}$  et en utilisant respectivement les définitions de  $\tilde{\Delta}t$ ,  $\phi(\lambda)$  données dans (3.103) et les expressions de  $(h_L^*, h_R^*)$  énoncées dans le Lemme 3.2.3.1-ii), on obtient

$$\phi(\lambda) = 1, \quad \tilde{\Delta}t^2 = 0, \quad [h^* + z] = 0.$$

Ainsi, en notant  $\hat{w}^{\text{HLL}} = (h^{\text{HLL}}, h^{\text{HLL}}\hat{u}^{\text{HLL}}, h^{\text{HLL}}\theta^{\text{HLL}}, (z_L + z_R)/2)^{\text{T}}$  dans  $\mathbb{R}^4$ , l'inégalité (3.121b) (ou sa forme équivalente (3.123)) devient

$$\begin{aligned} 0 &\leq \hat{\eta}^{\text{HLL}} - \frac{(h^{\text{HLL}}\hat{u}^{\text{HLL}})^2}{2h^{\text{HLL}}} - \frac{g}{2}(h^{\text{HLL}})^2\varphi(\theta^{\text{HLL}}) - \psi^{\text{HLL}}\frac{z_L + z_R}{2} + \frac{g}{8}\varphi(\theta^{\text{HLL}})[z]^2, \\ &\leq \hat{\eta}^{\text{HLL}} - \frac{(h^{\text{HLL}}\hat{u}^{\text{HLL}})^2}{2h^{\text{HLL}}} - \frac{g}{2}(h^{\text{HLL}})^2\varphi(\theta^{\text{HLL}}) - \psi(w^{\text{HLL}})\frac{z_L + z_R}{2} + \frac{g}{8}\varphi(\theta^{\text{HLL}})[z]^2, \\ &\leq \hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g}{8}\varphi(\theta^{\text{HLL}})[z]^2. \end{aligned} \quad (3.124)$$

Lorsque  $\theta_L = \theta_R$ , et en considérant la quantité  $\hat{s}_{LR}^{\text{WBAR}}$  définie dans (3.40b),  $\hat{s}_{LR}$  donné par (3.120) satisfait

$$\hat{s}_{LR} = \varphi(\theta^{\text{HLL}})\hat{s}_{LR}^{\text{WBAR}}.$$

Ainsi, en s'appuyant sur les définitions de  $\hat{w} \mapsto \hat{\eta}(\hat{w})$  et de  $\hat{\eta}^{\text{HLL}}$  respectivement explicitées dans (3.71) et dans (3.99), la quantité  $\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + g\varphi(\theta^{\text{HLL}})[z]^2/8$  du membre de droite de (3.124) peut également être calculée en substituant  $g$  par  $g\varphi(\theta^{\text{HLL}})$  dans la preuve du Lemme 3.1.3.1-i). Cette remarque permet de déduire que le membre de droite de l'inégalité (3.124) satisfait l'estimation

$$\begin{aligned} \hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g\varphi(\theta^{\text{HLL}})[z]^2}{8} &= \\ \frac{h_R h_L [u]^2}{4(h_L + h_R)} + \frac{g\varphi(\theta^{\text{HLL}})}{8}[h + z]^2 &- \\ - \frac{[u]^2 h_R h_L (h_L u_R - h_R u_L)}{4\lambda(h_L + h_R)^2} - \varphi(\theta^{\text{HLL}}) \frac{h_L u_L + h_R u_R}{4\lambda} \left( g[z] + \frac{\Delta x \hat{s}_{LR}^{\text{WBAR}}}{h^{\text{HLL}}} \right) &+ \mathcal{O}\left(\frac{1}{\lambda^2}\right). \end{aligned}$$

Puisque  $\varphi$  est positive, on peut donc directement utiliser le Lemme 3.1.3.1-i) pour déduire l'existence de  $\lambda > 0$  suffisamment grand tel que l'inégalité (3.121b) est satisfaite dans le cas  $\theta_L = \theta_R$ .

Pour le second cas formé par  $\theta_L \neq \theta_R$ , les définitions (3.103) donnent

$$\tilde{\Delta}t^2 \xrightarrow{\Delta t \rightarrow 0} 0, \quad \phi(\lambda) \xrightarrow{\lambda \rightarrow +\infty} +\infty.$$

Si  $[z] \neq 0$  alors le Lemme 3.2.3.3-ii) assure que  $[h^* + 2z] \neq 0$  et par conséquent, il existe alors  $\lambda > 0$  suffisamment grand et  $\Delta t > 0$  suffisamment petit tels que l'inégalité (3.123) (équivalente à (3.121b)) est satisfaite. Si  $[z] = 0$  et comme  $\hat{s}_{LR}$  donné par (3.120) est consistant, il suit immédiatement  $\hat{u}^{\text{HLL}}|_{[z]=0} = (hu)^{\text{HLL}}/h^{\text{HLL}}$ . Ainsi, en notant  $z_0 = z_L = z_R$ , le membre de droite de l'inégalité (3.123) s'écrit

$$\begin{aligned} &\left( \hat{\eta}^{\text{HLL}} - \frac{(h^{\text{HLL}}\hat{u}^{\text{HLL}})^2}{2h^{\text{HLL}}} - \frac{h^{\text{HLL}}}{2}\psi(w^{\text{HLL}}) - \psi^{\text{HLL}}\frac{z_L + z_R}{2} \right) \Big|_{[z]=0}^{\theta_L \neq \theta_R} \\ &= \eta^{\text{HLL}} + z_0\psi^{\text{HLL}} - \frac{((hu)^{\text{HLL}})^2}{2h^{\text{HLL}}} - \frac{h^{\text{HLL}}}{2}\psi(w^{\text{HLL}}) - z_0\psi^{\text{HLL}}, \\ &= \eta^{\text{HLL}} - \eta(w^{\text{HLL}}). \end{aligned}$$



Par conséquent, lorsque  $\theta_L \neq \theta_R$  et  $[z] = 0$ , l'inégalité (3.123) se lit

$$-\frac{g}{8}\varphi(\theta^{\text{HLL}})\phi(\lambda)[h^*]^2 + \frac{g}{4}\varphi(\theta^{\text{HLL}})\tilde{\Delta}t^2((h_L^*)^2 + (h_R^*)^2) \leq \eta^{\text{HLL}} - \eta(w^{\text{HLL}}). \quad (3.125)$$

Puisque  $\theta \mapsto \varphi(\theta)$  vérifie les conditions (3.74), la fonction  $w \mapsto \eta(w)$  est strictement convexe. Comme  $\theta_L \neq \theta_R$ , cette convexité stricte associée à la définition du schéma HLL rappelée en Section 1.4.3 assurent que le terme de droite de l'inégalité ci-dessus est strictement positif. Ainsi, même dans le cas le plus défavorable caractérisé par  $[h^*] = 0$ , il est toujours possible de sélectionner  $\Delta t > 0$  pour que l'inégalité (3.125) soit assurée. Ce dernier argument termine de montrer l'inégalité (3.121b) ainsi que la preuve du lemme.  $\square$

Le Lemme ci-dessus montre que les équations (3.104e)-(3.104f) permettent de satisfaire les inégalités (3.100d)-(3.100e) (également données par (3.121a)-(3.121b)) or ces inégalités sont des conditions suffisantes à la stabilité entropique pour le couple  $(\hat{\eta}, \hat{G})$  donné par (3.71). Par conséquent, un schéma de type Godunov entropique et *well-balanced* peut maintenant être détaillé.

**Théorème 3.2.3.1** (Schéma de type Godunov entropique et *well-balanced* pour le lac au repos). *Soient  $\hat{w}_L, \hat{w}_R$  deux états constants de  $\hat{\Omega}$ ,  $\tilde{w}(\cdot, \hat{w}_L, \hat{w}_R) : \mathbb{R} \rightarrow \mathbb{R}^3$  un solveur de Riemann de la forme (3.90) ainsi que  $\varphi : \mathbb{R}_*^+ \rightarrow \mathbb{R}_*^+$  une fonction régulière, strictement convexe, bijective et satisfaisant les inégalités (3.74).*

*En considérant la quantité  $\hat{s}_{LR}$  donnée par (3.120), il existe  $\lambda > 0$  suffisamment grand et  $\Delta t > 0$  suffisamment petit tels que si les hypothèses (3.117) sont vérifiées, si la condition CFL (3.89) et les inégalités (3.100d)-(3.100e) sont satisfaites et si les états intermédiaires du solveur de Riemann sont donnés par une solution du système (3.104) alors un schéma de type Godunov (3.76) défini par la donnée d'un tel solveur de Riemann*

- i) est consistant avec les équations de Ripa (3.70),*
- ii) satisfait une inégalité d'entropie discrète locale (3.83) pour le couple  $(\hat{\eta}, \hat{G})$  donné dans (3.71) et dans laquelle le flux numérique d'entropie  $\hat{G}$  est donné par (3.84),*
- iii) est well-balanced pour le lac au repos.*

*Démonstration.* Concernant l'assertion *i)*, lorsque  $\hat{w}_L = \hat{w}_R = \hat{w}$ , une solution triviale du système d'équations (3.104) est donnée par  $w_L^* = w_R^* = w$ . Par conséquent, le solveur de Riemann vérifie  $\tilde{w}(\cdot, \hat{w}, \hat{w}) = w$ . De plus, d'après le Lemme 3.2.2.2, les équations (3.104a)-(3.104b)-(3.104c) assurent que le solveur de Riemann satisfait la relation de consistance intégrale (3.81). Ainsi, le solveur de Riemann vérifie la notion de consistance donnée dans la Définition 1.4.4.2. Par suite, et comme  $\hat{s}_{LR}$  défini par (3.120) est consistant, la Proposition 3.2.1.1-*i)* permet de déduire la consistance du schéma associé.

Pour l'inégalité d'entropie *ii)* et puisque les inégalités (3.100d)-(3.100e) sont satisfaites, le Lemme 3.2.2.2 permet de déduire que le solveur de Riemann vérifie les inégalités (3.82). Par conséquent, l'inégalité d'entropie discrète suit immédiatement par application de la Proposition 3.2.1.1-*ii)*.

Pour montrer *iii)*, il est suffisant, d'après la Proposition 3.2.1.1-*iii)*, de montrer que si  $\hat{w}_L$  et  $\hat{w}_R$  forment un équilibre local pour le lac au repos (3.79) alors  $w_L^* = w_L$  et  $w_R^* = w_R$ . Comme un équilibre local (3.79) impose  $\theta_L = \theta_R = \theta$ , les solutions du système d'équations (3.104) sont dans ce cas explicites et elles sont données par le Lemme 3.2.3.1-*i)*-*ii)*. Ainsi, dans le cas d'un équilibre local du lac au repos (3.79) et en notant  $[\cdot] = \cdot_R - \cdot_L$ , il suit

$$\begin{aligned} h_R^* \Big|_{\substack{u_L=u_R=0, \\ \theta_L=\theta_R=\theta, \\ [h+z]=0}} &= \frac{h_L + h_R - [z]}{2} = \frac{h_L + h_R + [h]}{2} = h_R, \\ h_L^* \Big|_{\substack{u_L=u_R=0, \\ \theta_L=\theta_R=\theta, \\ [h+z]=0}} &= \frac{h_L + h_R + [z]}{2} = \frac{h_L + h_R - [h]}{2} = h_L. \end{aligned}$$

Il faut maintenant montrer que  $u_L^*$  et  $u_R^*$  sont nuls. Dans le cas d'un équilibre local pour le lac au repos (3.79), la quantité  $\hat{s}_{LR}$  définie dans (3.120) s'écrit

$$\hat{s}_{LR} = -g\varphi(\theta) \frac{h_L + h_R}{2} \frac{z_R - z_L}{\Delta x}.$$

Par conséquent, un calcul immédiat montre que  $h^{\text{HLL}}\hat{u}^{\text{HLL}} = (hu)^{\text{HLL}} + \Delta x s_{LR}/(2\lambda) = 0$  et comme  $h^{\text{HLL}} > 0$ , on en déduit

$$\hat{u}^{\text{HLL}} \Big|_{\substack{u_L=u_R=0, \\ \theta_L=\theta_R=\theta, \\ [h+z]=0}} = 0$$

Étant donné que  $u_L^* = u_R^* = \hat{u}^{\text{HLL}}$ , cette dernière égalité termine de montrer la propriété *well-balanced* ainsi que la preuve du théorème.  $\square$

Puisque l'existence des solutions du système (3.104) n'est établi que pour le cas  $\theta_L = \theta_R$ , le théorème ci-dessus est un résultat très partiel d'existence d'un schéma numérique pour les équations de Ripa, entropique et *well-balanced* pour le lac au repos. Pour conclure à l'existence d'un tel schéma, il serait maintenant nécessaire de démontrer les hypothèses (3.117).

Néanmoins, les équations algébriques (3.117a) peuvent être *résolues* en pratique avec des méthodes numériques usuelles. Par conséquent, le Théorème 3.2.3.1 montre que la mise on œuvre d'un schéma entropique et *well-balanced* pour les équations de Ripa est possible.

La prochaine section concerne les expériences numériques pour le schéma du Théorème 3.2.3.1 ainsi que pour les schémas développées en Section 3.1.3 et 3.1.4.

## 3.3 Résultats de simulations numériques

### 3.3.1 Les équations de Saint Venant

#### 3.3.1.1 Schéma entropique et *well-balanced* pour le lac au repos du Théorème 3.1.3.1

Les cas test suivants illustrent le résultat du Théorème 3.1.3.1 qui présente deux schéma numériques entropique et *well-balanced* distincts. Lorsque la fonction  $z$  est constante, les schémas du Théorème 3.1.3.1 dégénèrent vers les schémas numériques définis par (2.27) or la section 2.5.1 montre que les schémas (2.27) produisent des résultats très similaires. Par conséquent, et pour chacun des cas tests suivants, l'un des deux schémas du Théorème 3.1.3.1 est sélectionné aléatoirement.

La principale difficulté à la mise en œuvre du Théorème 3.1.3.1 réside dans le choix du paramètre  $\lambda > 0$  qui doit permettre l'existence de solutions réelles au système d'équations (3.26), (3.40). Si le Théorème 3.1.3.1 assure l'existence d'un tel paramètre, il n'en donne pas de formule de calcul explicite. Pour choisir le paramètre  $\lambda$ , une procédure brute de force est adoptée. Cette procédure simple consiste à augmenter  $\lambda$  tant que le système (3.26), (3.40) n'admet pas de solution. Pour chaque interface du maillage ayant à sa gauche un état constant  $\hat{w}_L$  et à sa droite un autre état constant  $\hat{w}_R$ ,  $\lambda$  est d'abord initialisé avec

$$\lambda_0 = \max_{\alpha \in \{L, R\}} |u_\alpha \pm \sqrt{gh_\alpha}|, \quad \forall (L, R).$$

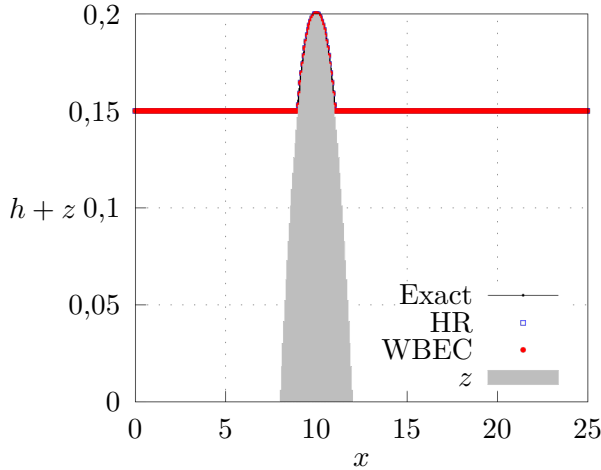
À partir de  $\lambda_0$  défini ci-dessus,  $\lambda$  est ensuite augmenté avec une progression arithmétique de raison  $\lambda_0$ . Après sélection du paramètre  $\lambda$ , le pas de temps  $\Delta t > 0$  est choisi en fixant le nombre CFL  $\lambda \Delta t / \Delta x$ . Par construction, la condition CFL  $\frac{\lambda \Delta t}{\Delta x} \leq \frac{1}{2}$  est satisfaite si et seulement si le nombre CFL prend ses valeurs dans  $]0, 1/2]$ .

Le premier cas test est celui de l'île immergée introduit dans [98] et qui illustre la robustesse du schéma dans les transitions en zones sèches. Le domaine spatial  $[0, 25]$  est discrétisé avec 400 cellules. Les conditions initiales et la fonction  $z$  sont données par

$$h_0(x) = \max(0, 15 - z(x), 0), \quad u_0(x) = 0, \quad z(x) = \max(0, 0, 2 - 0, 05(x - 10)^2). \quad (3.126)$$

Les conditions de bord sont périodiques et la solutions exacte se compose de deux états du lac au repos séparés par une zone où  $h$  est nul. La date finale de la simulation est  $t = 100$  et le nombre CFL est égal

à 0,45. La Figure 3.2 montre les résultats du schéma du Théorème 3.1.3.1 comparés à la reconstruction hydrostatique [13] appliquée au flux numérique de Rusanov [185].



Erreurs commises sur les variables $(h, hu)$ pour le lac au repos.			
	$h$		
	$L^1$	$L^2$	$L^\infty$
WBEC	1,69E-21	4,79E-21	2,29E-41
HR	6,55E-18	8,44E-18	1,25E-17
	$hu$		
	$L^1$	$L^2$	$L^\infty$
WBEC	9,88E-20	3,43E-20	1,14E-22
HR	3,17E-17	2,52E-17	3,07E-17

FIGURE 3.2 – Sur la gauche : résultats numériques à la date  $t = 100,0$  pour le problème de l’île immergée (3.126) avec la légende WBEC : schéma entropique et *well-balanced* pour le lac repos décrit dans le Théorème 3.1.3.1, HR : reconstruction hydrostatique [13] appliquée au flux numérique de Rusanov [185]. Sur la droite : erreurs numériques entre la solution exacte et la solution calculée.

La propriété *well-balanced* ainsi que la disjonction des cas présentes dans les définitions (3.50)-(3.51) permettent de préserver les états du lac au repos à la précision machine près.

Le second cas test concerne les trois problèmes de Goutal et Maurel [111] pour lesquels le domaine spatial  $[0, 25]$  est discrétisé avec 400 cellules et la fonction  $z$  est donnée par (3.126). Dans la suite, la notation avec des exposants  $GM_1$ ,  $GM_2$  and  $GM_3$  est employée pour désigner les quantités relatives à chacun des trois problèmes. En considérant  $(h^{GM_k})_{k \in \{1,2,3\}}$  et  $(q^{GM_k})_{k \in \{1,2,3\}}$  définis dans le tableau 3.1, les conditions initiales pour chaque problème sont

$$h_0^{GM_k}(x) = h^{GM_k}, \quad (hu)_0^{GM_k}(x) = q^{GM_k}, \quad \forall k \in \{1, 2, 3\}. \quad (3.127)$$

Sur la frontière gauche du domaine spatial,  $h$  satisfait une condition de Neumann homogène et  $hu$  est fixée avec les valeurs  $(q^{GM_k})_{k \in \{1,2,3\}}$ . Sur la frontière droite, si l’écoulement est sous-critique (*i.e*  $|u| < \sqrt{gh}$ ) alors  $h$  est fixée à la valeur  $h^{GM_k}$ , sinon  $h$  vérifie une condition de Neumann homogène. Sur cette frontière droite,  $hu$  satisfait une condition de Neumann homogène.

Paramètres des problèmes de Goutal et Maurel.			
	$GM_1$	$GM_2$	$GM_3$
Date finale	500	125	1000
$h^{GM_k}$	2	0,66	0,33
$q^{GM_k}$	4,42	1,53	0,18

TABLE 3.1 – Dates finales, valeurs initiales et conditions de bord pour les trois problèmes de Goutal et Maurel [111].

Lorsque le régime stationnaire est atteint, les solutions des problèmes de Goutal et Maurel sont caractérisés par des valeurs de  $hu$  et de  $u^2/2 + g(h + z)$  constantes. Pour les problèmes  $GM_1$  et  $GM_2$ , ces constantes sont continues et par conséquent, les solutions stationnaires s’identifient à des états d’équilibre mouvants. En revanche, pour le problème  $GM_3$ , ces valeurs sont seulement constantes par morceaux et la surface libre  $h + z$  se compose d’une onde de choc stationnaire. Pour les trois problèmes, des solutions stationnaires exactes sont calculées avec le logiciel SWASHES [71]. Les dates finales de simulation sont données dans le tableau 3.1 et le nombre CFL est égal à 0,45. La Figure 3.3 montre les résultats comparés.

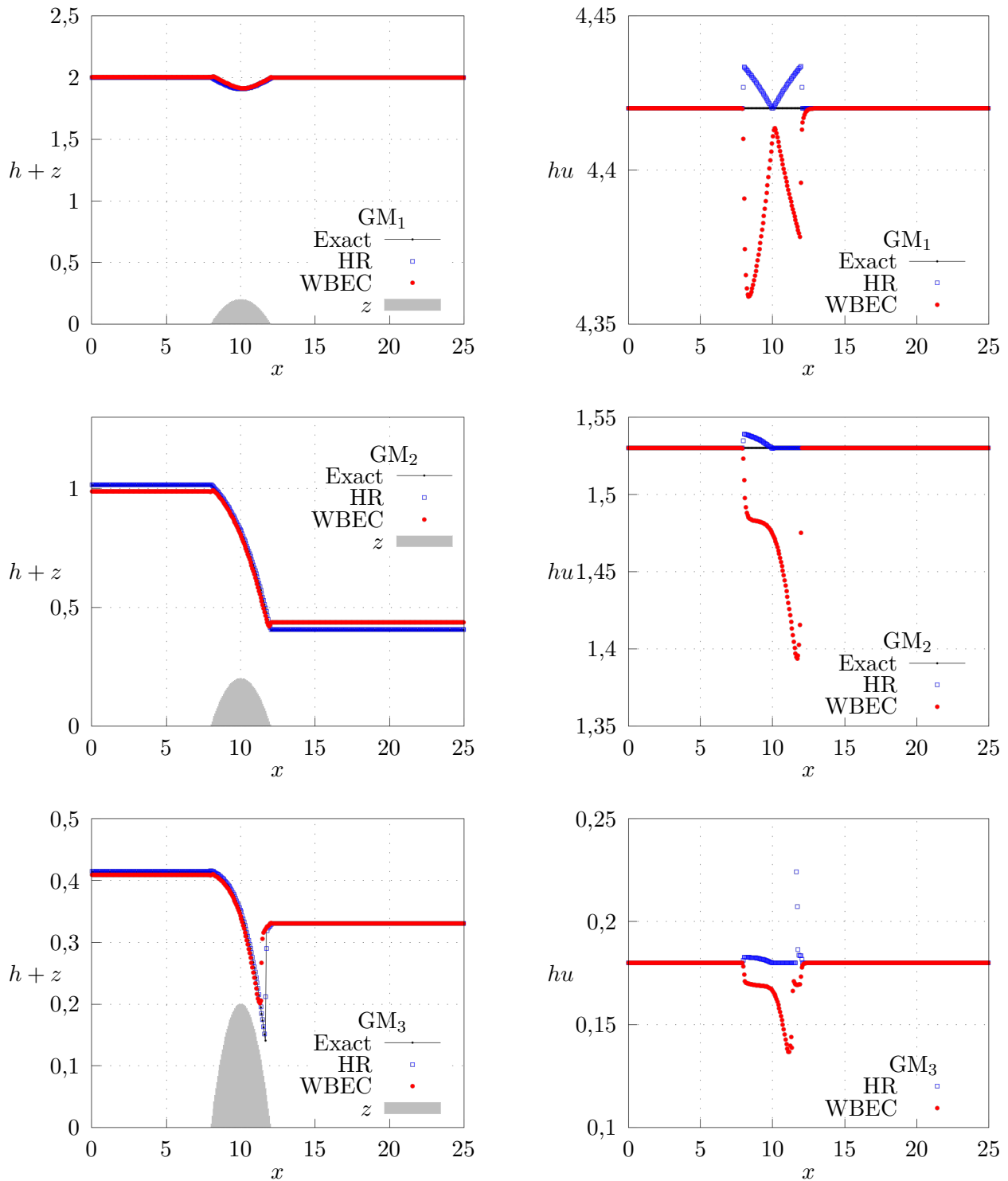


FIGURE 3.3 – Résultats numériques pour les trois problèmes de Goutal et Maurel (3.127) sur un maillage constitué de 400 mailles et avec la légende WBEC : schéma entropique et *well-balanced* pour le lac repos décrit dans le Théorème 3.1.3.1, HR : reconstruction hydrostatique [13] appliquée au flux numérique de Rusanov [185].

Les résultats sont globalement corrects mais la surface libre  $h+z$  est parfois mal placée. Ce défaut est particulièrement flagrant pour le problème GM<sub>2</sub>. La surface libre du problème GM<sub>3</sub> est sous évaluée avant le choc mais elle est correctement simulée après. Malgré les propriétés de stabilité entropique et *well-balanced* vérifiées par les schémas du Théorème 3.1.3.1, les solutions numériques peuvent converger vers des solutions qui ne sont pas celles de références. De tels écarts de convergence ont déjà été observés

avec d'autres schémas numériques [34, 42, 59]. De plus, les schémas du Théorème 3.1.3.1 produisent des oscillations parasites pour la variable  $hu$ . Ces oscillations existent également avec la reconstruction hydrostatique mais leurs amplitudes sont plus faibles. Pour ces trois problèmes, un schéma numérique qui utilise la reconstruction hydrostatique [13] couplée à un flux numérique de Rusanov [185] est plus robuste que les schémas du Théorème 3.1.3.1.

Le dernier cas test est celui de la rupture de barrage telle que décrite dans [166]. Le domaine spatial  $[-1, 1]$  est discrétisé avec 400 cellules. Les conditions initiales et la fonction  $z$  sont données par

$$h_0(x) + z(x) = \begin{cases} 3 & \text{si } x < 0, \\ 1 & \text{sinon,} \end{cases} \quad u_0(x) = 0, \quad z(x) = \frac{1}{2} \cos^2(\pi x). \quad (3.128)$$

Des conditions de Neumann homogènes sont imposées aux deux frontières du domaine spatial. Une solution de référence est calculée avec le schéma HLL [118] sur un maillage fin constitué de 50 000 cellules. La date finale de la simulation est  $t = 0,1$  et le nombre CFL vaut 0,45. La Figure 3.4 montre les résultats comparés.

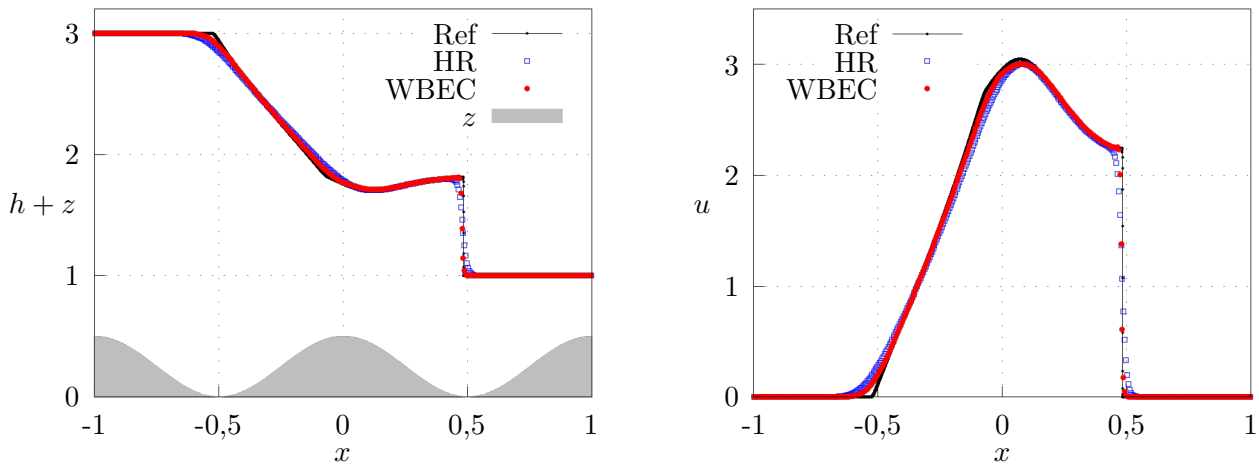


FIGURE 3.4 – Résultats numériques pour la rupture de barrage (3.128) à la date  $t = 0,1$ , sur un maillage constitué de 400 mailles et avec la légende WBEC : schéma entropique et *well-balanced* pour le lac repos décrit dans le Théorème 3.1.3.1, HR : reconstruction hydrostatique [13] appliquée au flux numérique de Rusanov [185].

Pour ce cas test, les solutions numériques calculées par les schémas du Théorème 3.1.3.1 épousent très bien la solution de référence. En particulier, la discontinuité localisée près de  $x = 0,5$  est restituée avec très peu de cellules.

### 3.3.1.2 Schéma entropique et *well-balanced* du Théorème 3.1.4.1

Cette section présente les résultats numériques du schéma du Théorème 3.1.4.1. Ce schéma vérifie une inégalité d'entropie discrète locale et il est *well-balanced* pour toutes les solutions stationnaires régulières.

Le schéma proposé est paramétré par  $\lambda > 0$  et  $\Delta t > 0$  et ces paramètres doivent être choisis afin d'assurer l'existence de solutions réelles au système d'équations (3.26), (3.55). Le Théorème 3.1.4.1 montre l'existence de tels paramètres mais il ne donne pas de formule de calcul explicite. La procédure de choix de  $\lambda$  et de  $\Delta t$  utilisée dans la suite repose sur le Lemme 3.1.4.3 qui donne une condition nécessaire et suffisante à l'existence de  $h_L^*$  et de  $h_R^*$  satisfaisants le système (3.26), (3.55). La preuve de ce lemme se décompose en trois études de cas et dans les deux premiers, il est suffisant de choisir  $\lambda$  et  $\Delta t$  selon les conditions

$$\lambda = \max_{\alpha \in \{L, R\}} |u_\alpha \pm \sqrt{gh_\alpha}|, \quad \frac{\lambda \Delta t}{\Delta x} \leq \frac{1}{2}, \quad \forall (L, R).$$

Pour le troisième cas de la preuve, le paramètre  $\lambda > 0$  est d'abord sélectionné par la méthode brute de force décrite dans la section 3.3.1.1 et il est augmenté tant que l'inégalité stricte  $\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + g[z]^2/s > 0$  n'est pas satisfaite. Ensuite, le pas de temps  $\Delta t > 0$  est initialisé avec  $\Delta x/2\lambda$  puis il est successivement diminué avec une progression géométrique de raison  $1/2$  et ce tant que l'inégalité large suivante n'est pas vérifiée :

$$\hat{\eta}^{\text{HLL}} - \hat{\eta}(\hat{w}^{\text{HLL}}) + \frac{g[z]^2}{8(1 + F_{LR}^2)} \geq 0.$$

Cette procédure de sélection, complétée si nécessaire avec les techniques de limitation exposées dans le Théorème 3.1.4.1, assure l'existence des solutions recherchées et satisfait également la condition CFL de non interaction. Enfin, en considérant les quantités  $\zeta_{LR}$  et  $\bar{h}$  définies dans (3.54), la fonction sign nécessaire à l'équation (3.65) est programmée comme suit :

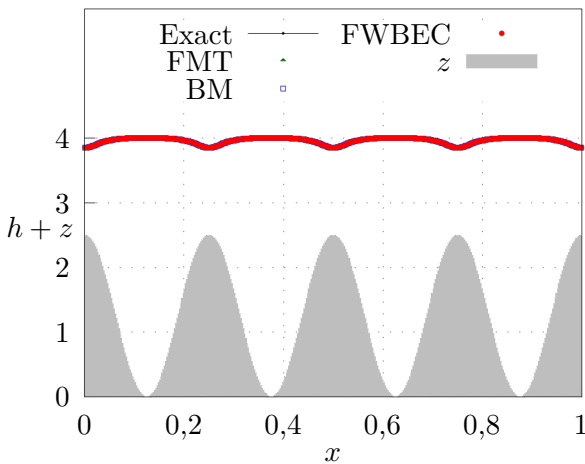
$$\text{sign}(r) \approx \frac{r}{|r| + \zeta_{LR}/(g\bar{h})}. \quad (3.129)$$

Les résultats numériques obtenus avec ces paramétrages sont comparés au résultats des schémas présentés dans [93, Section 3.2] et dans [35]. Ces deux schémas respectivement notés FMT et BM, sont retenus car ils sont *well-balanced* pour les équilibres mouvants et vérifient tous deux une inégalité d'entropie semi-discrète. Le schéma BM préserve également l'espace convexe  $\hat{\Omega}$  sans technique de limitation additionnelle. Pour ce comparatif, il est associé au flux numérique de Rusanov [185].

Le premier cas test est destiné à illustrer la propriété *well-balanced* pour les équilibres mouvants. Le domaine spatial  $[0, 1]$  est discrétisé avec 400 cellules. En notant  $q_0 = 5/2$  et  $B_0 = 25/98 + 4g$ , les conditions initiales et la fonction  $z$  sont données par

$$(hu)_0(x) = q_0, \quad \frac{u_0^2(x)}{2} + g(h_0 + z)(x) = B_0, \quad z(x) = \frac{5}{2} \cos^2(4\pi x). \quad (3.130)$$

À partir des équations ci-dessus, les valeurs initiales  $(h_0, u_0)(x)$  sont calculées avec une méthode de Newton. Des conditions de bord périodiques sont prescrites aux deux extrémités du domaine spatial. La solution exacte est un équilibre mouvant (1.13). Les résultats graphiques ainsi que les erreurs numériques à la date finale  $t = 1,0$  sont présentés dans la Figure 3.5.



Erreurs commises sur les variables $(h, hu)$ pour l'équilibre mouvant (3.130).			
	$h$		
	$L^1$	$L^2$	$L^\infty$
FWBEC	0,00E-00	0,00E-00	0,00E-00
BM	2,49E-17	4,41E-17	1,27E-17
FMT	1,84E-13	2,03E-13	3,28E-13
	$hu$		
	$L^1$	$L^2$	$L^\infty$
FWBEC	1,13E-20	8,74E-20	7,64E-39
BM	8,58E-17	1,29E-16	9,19E-17
FMT	2,87E-13	3,75E-13	4,10E-13

FIGURE 3.5 – Sur la gauche : résultats numériques à la date  $t = 1,0$  pour l'équilibre mouvant (3.130) avec la légende FWBEC : schéma entropique et *well-balanced* décrit dans le Théorème 3.1.4.1, BM : schéma décrit dans [35] associé au flux numérique de Rusanov [185], FMT : schéma proposé dans [93, Section 3.2]. Sur la droite : erreurs numériques entre la solution exacte et la solution calculée.

Grâce à la propriété *well-balanced*, l'équilibre mouvant (3.130) est préservé à la précision machine près.

Le second cas test est dédié au trois problèmes de Goutal et Maurel [111] dont les données ont été introduites en Section 3.3.1.1. La Figure 3.6 montre les résultats comparés et le tableau 3.2 rapporte les erreurs numériques entre la solution exacte et la solution calculée par les schémas numériques.

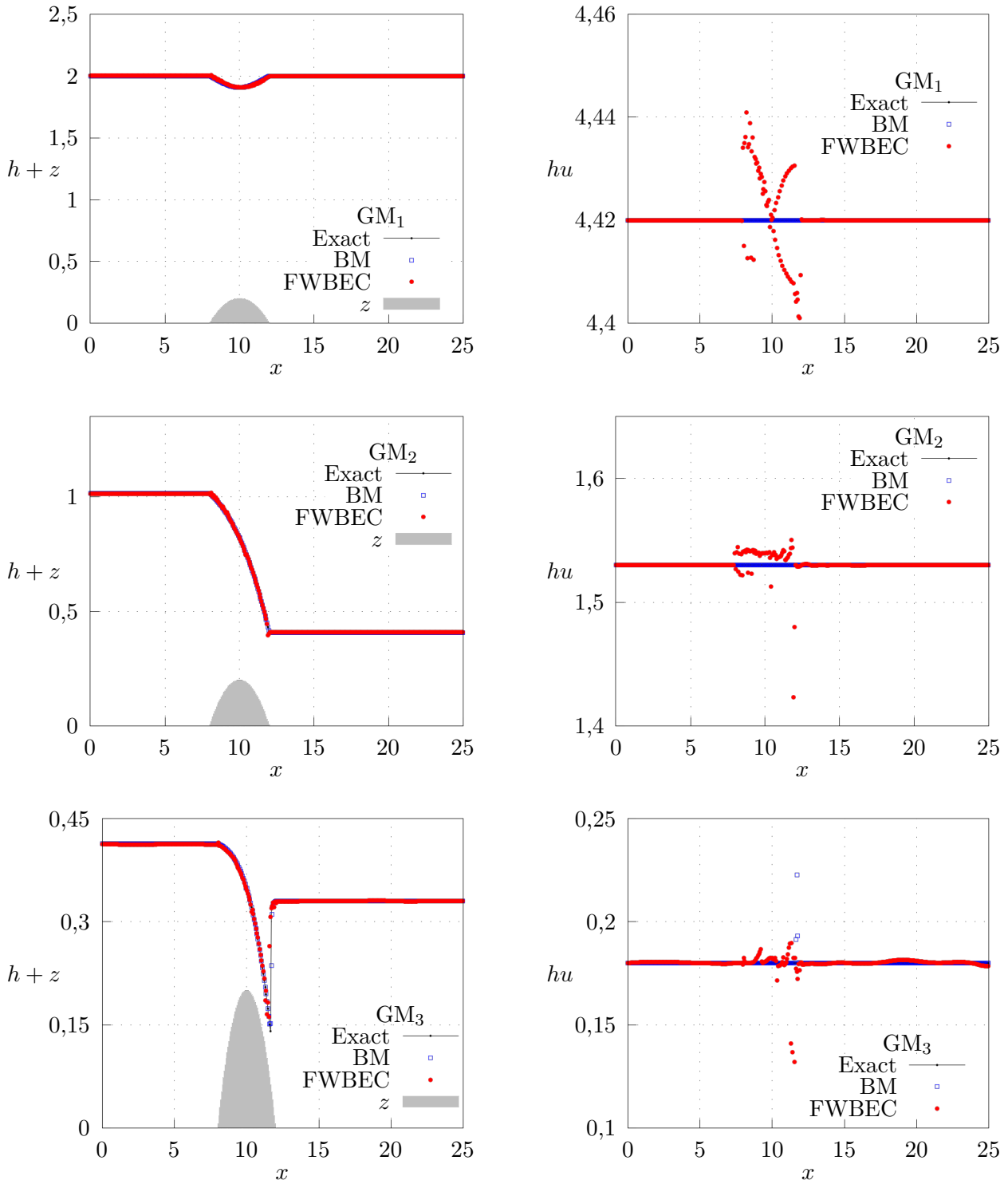


FIGURE 3.6 – Résultats numériques pour les trois problèmes de Goutal et Maurel (3.127) sur un maillage constitué de 400 mailles et avec la légende FWBEC : schéma entropique et *well-balanced* décrit dans le Théorème 3.1.4.1, BM : schéma décrit dans [35] associé au flux numérique de Rusanov [185].

Malgré les efforts de programmation consentis, des résultats de simulation pour les problèmes de Goutal et Maurel (3.127) n'ont pas pu être obtenus avec le schéma de [93, Section 3.2]. Ce schéma paraît très instable dans les phases transitoires et les auteurs eux-mêmes formulent cette remarque [93, Remarque 3.4]. Le schéma du Théorème 3.1.4.1 produit des oscillations parasites particulièrement visibles pour la variable  $hu$  mais il corrige nettement les écarts de convergence observés dans la section

3.3.1.1. Toutefois, un zoom sur la Figure 3.6 montre que la surface libre  $h + z$  demeure légèrement mal placée dans le cas du problème  $GM_1$ .

Erreurs commises sur les variables $(h, hu)$ pour le problème $GM_1$ .				Erreurs commises sur les variables $(h, hu)$ pour le problème $GM_2$ .			
	$h$				$h$		
	$L^1$	$L^2$	$L^\infty$		$L^1$	$L^2$	$L^\infty$
FWBEC	3,86E-02	1,41E-02	2,00E-04	FWBEC	9,35E-02	2,13E-02	4,46E-03
BM	1,91E-15	4,16E-16	6,11E-17	BM	9,06E-03	2,37E-03	5,12E-04
	$hu$				$hu$		
	$L^1$	$L^2$	$L^\infty$		$L^1$	$L^2$	$L^\infty$
FWBEC	3,77E-02	2,10E-02	4,439E-04	FWBEC	4,89E-02	3,53E-02	1,25E-03
BM	7,29E-15	1,49E-15	2,91E-16	BM	7,78E-05	3,38E-05	4,33E-08

Erreurs commises sur les variables $(h, hu)$ pour le problème $GM_3$ .			
	$h$		
	$L^1$	$L^2$	$L^\infty$
FWBEC	4,28E-02	5,11E-02	2,62E-03
BM	8,23E-03	2,20E-02	4,86E-04
	$hu$		
	$L^1$	$L^2$	$L^\infty$
FWBEC	2,57E-02	2,00E-02	1,59E-03
BM	4,21E-03	1,15E-02	1,32E-04

TABLE 3.2 – Erreurs entre les solutions calculées par les schémas numériques et les solutions exactes des problèmes de Goutal et Maurel (3.127) données par le logiciel SWASHES [71]. Les notations sont FWBEC : schéma entropique et *well-balanced* décrit dans le Théorème 3.1.4.1, BM : schéma décrit dans [35] associé au flux numérique de Rusanov [185].

Selon les tableaux 3.2, le schéma du Théorème 3.1.4.1 approche les solutions exactes avec une erreur excessive. En effet, comme ce schéma est *well-balanced* pour les équilibres mouvants réguliers, il devrait produire pour les problèmes  $GM_1$  et  $GM_2$ , une solution numérique dont l'erreur est au plus de l'ordre de la précision machine. L'origine exacte de ce défaut pratique est inconnue mais une hypothèse plausible est la programmation de la fonction sign nécessaire à l'équation (3.65). En effet, comme la fonction sign n'est pas continue en zéro, sa programmation demande une attention particulière. La définition (3.129) lève la difficulté de continuité mais d'autres choix sont possibles. Chacun de ces choix de programmation donnent un schéma numérique formellement *proche* du Théorème 3.1.4.1 mais qui demeure en pratique différent. Le schéma [35] couplé au flux numérique de Rusanov [185] préserve la solution exacte du problème  $GM_1$  et pour les trois problèmes, les erreurs commises par ce schéma sont inférieures à celles du schéma introduit dans le Théorème 3.1.4.1.

Le dernier cas test est celui de la rupture de barrage telle que décrite dans la section 3.3.1.1. La Figure 3.7 montre les résultats comparés.



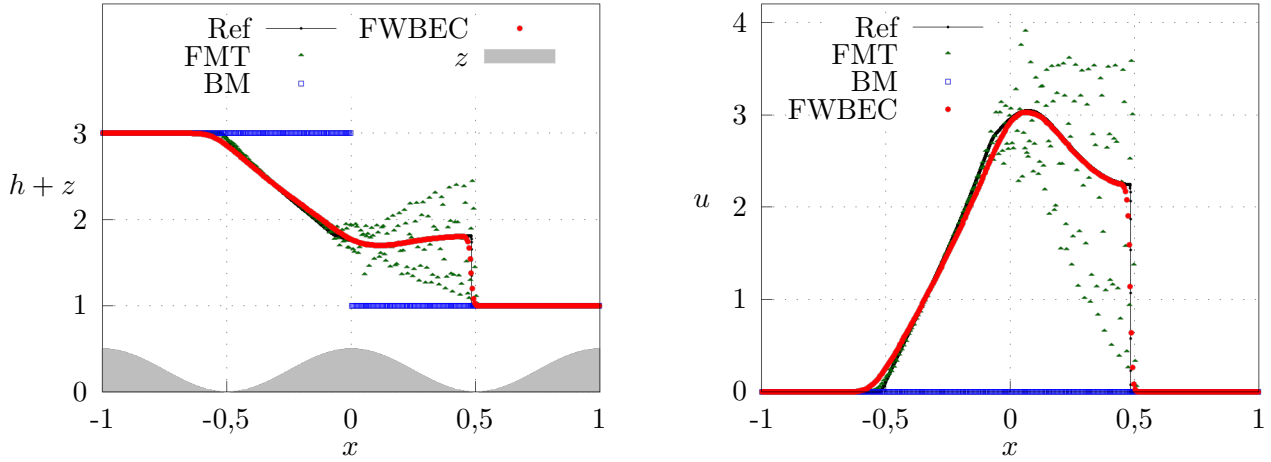


FIGURE 3.7 – Résultats numériques pour la rupture de barrage (3.128) à la date  $t = 0,1$ , sur un maillage constitué de 400 mailles et avec la légende FWBEC : schéma entropique et *well-balanced* décrit dans le Théorème 3.1.4.1, BM : schéma décrit dans [35] associé au flux numérique de Rusanov [185], FMT : schéma décrit dans [93, Section 3.2].

Le schéma [93, Section 3.2] génère de fortes oscillations à proximité de la discontinuité. Le schéma [35] associé au flux numérique de Rusanov [185] identifie ce cas test à deux états du lac au repos séparés par une discontinuité. Par conséquent, ce schéma préserve dans ce cas la donnée initiale (3.128). Le schéma du Théorème 3.1.4.1 restitue une solution numérique conforme à la solution de référence.

### 3.3.2 Les équations de Ripa

Cette section présente les résultats numériques pour les équations de Ripa (3.70) dans le cas  $\varphi = \exp$ . La fonction  $\exp : ]0, +\infty[ \rightarrow ]1, +\infty[$  est positive, bijective, strictement convexe et d'après [78], elle satisfait les inégalités strictes (3.74). Par conséquent, cette fonction est conforme aux conditions du Lemme 3.2.1.1 et elle peut donc être utilisée pour illustrer le schéma du Théorème 3.2.3.1.

Pour être mis en œuvre, ce théorème nécessite d'abord une méthode numérique de résolution des équations (3.104) puis un choix de  $\lambda > 0$  et de  $\Delta t > 0$  qui assure les inégalités (3.100d)-(3.100e) (également données par (3.121a)-(3.121b)).

D'abord, pour chaque interface du maillage ayant de part et d'autre un état gauche constant  $\hat{w}_L$  et un état droit constant  $\hat{w}_R$ , la résolution du système d'équations (3.104) est conduite par disjonction des cas. En considérant  $\tilde{\Delta t}$  défini dans (3.103), si  $|\tilde{\Delta t}| < 10^{-11}$  alors les solutions explicites du Lemme 3.2.3.1 sont utilisées, sinon, pour un couple  $(\lambda, \Delta t)$  donné, les solutions de (3.104) sont obtenue avec une méthode de Newton appliquée aux équations (3.117a). Il est admis que l'existence des solutions du système (3.104) est acquise si la procédure de Newton converge. Le critère de convergence choisi est l'erreur absolue entre deux itérations consécutives et la procédure est stoppée lorsque cette erreur est inférieure à  $10^{-13}$ .

Ensuite, les inégalités (3.100d)-(3.100e) sont obtenues par une méthode brute de force. Tout d'abord, le couple  $(\lambda, \Delta t)$  est initialisé avec

$$\lambda = \max_{\alpha \in \{L,R\}} (|u_\alpha \pm \sqrt{gh_\alpha \varphi(\theta_\alpha)}|, |u_\alpha|), \quad \Delta t = \frac{\Delta x}{2\lambda},$$

puis le système d'équations (3.104) est ensuite résolu tant que les inégalités (3.100d)-(3.100e) ne sont pas satisfaites. À chaque résolution, les réels  $\lambda$  et  $\Delta t$  sont respectivement augmentés et diminués par des progressions linéaires et géométriques de raison respectives  $\frac{23}{20}$ ,  $\frac{19}{20}$ . Par construction, la quantité  $\Delta t > 0$  obtenue satisfait la condition CFL de non interaction  $\frac{\lambda \Delta t}{\Delta x} \leq \frac{1}{2}$ .

Le cas test d'illustration choisi est une version modifiée de celui proposé par A. Chertock *et al* dans [58]. Dans ce cas test, la constante de gravitation  $g$  vaut 1, le domaine spatial est  $[-1, 1]$  et la fonction

$z$  est donnée par

$$z(x) = \begin{cases} 2 \cos(10\pi(x + 0,3)) + 2, & \text{si } -0,4 \leq x \leq -0,2, \\ 0,5 \cos(10\pi(x - 0,3)) + 0,5, & \text{si } 0,2 \leq x \leq 0,4, \\ 0, & \text{sinon.} \end{cases} \quad (3.131)$$

La donnée initiale est

$$(h, u, \varphi(\theta))^T(x, t = 0) = \begin{cases} (5 - z, 0, 1,5)^T, & \text{si } x \leq 0, \\ (1 - z, 0, 5)^T, & \text{sinon.} \end{cases} \quad (3.132)$$

Aux deux extrémités du domaine spatial, les conditions de bord sont de type Neumann homogènes. Une solutions de référence est calculée avec le schéma HLL standard [118] sur un maillage fin constitué de 50 000 cellules. La date finale de la simulation est  $t = 0,3$ . La Figure 3.8 montre les résultats du schéma numérique de type Godunov du Théorème 3.2.3.1.

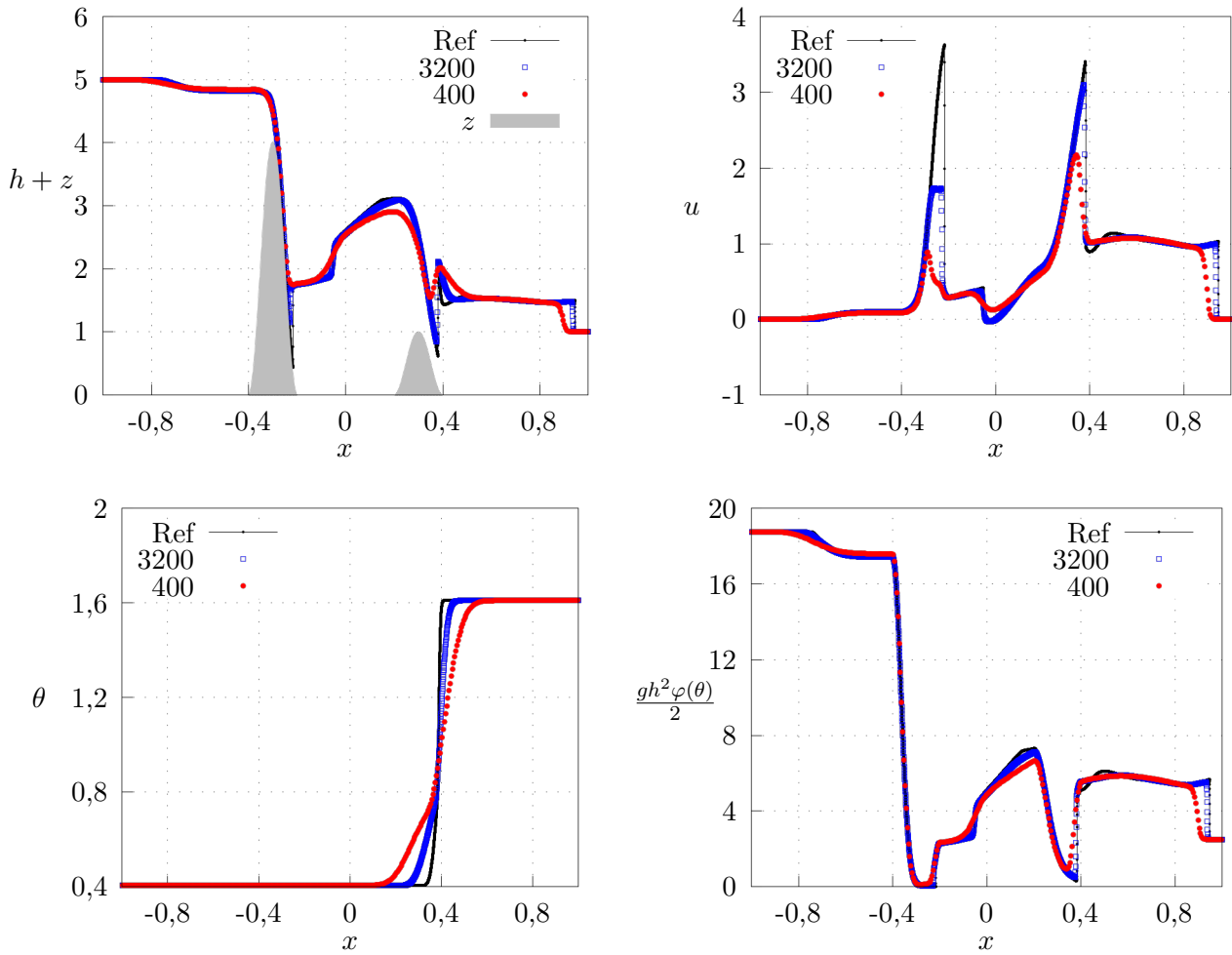


FIGURE 3.8 – Résultats à la date  $t = 0,3$  du schéma numérique de type Godunov du Théorème 3.2.3.1 pour le problème (3.131)-(3.132) et sur un maillage de 400 cellules en rouge et de 3 200 cellules en bleu.

La solution calculée épouse relativement bien la solution de référence mais elle demeure diffuse. Sur maillage grossier, une erreur de positionnement est observée sur le front du choc situé approximativement en  $x = 0,8$ . Cette erreur se corrige en raffinant le maillage.

### 3.4 Synthèse et conclusion

La première partie de ce chapitre donne une condition suffisante de stabilité entropique pour un schéma de type Godunov approchant les solutions des équations de Saint Venant. L'application de

cette condition à un solveur de Riemann à deux états intermédiaires constants produit un système d'équations non linéaires sous-déterminé.

Deux fermetures de ce système sont proposées et toutes deux conduisent à des équations quadratiques dont les solutions sont toujours explicites. En revanche, si l'existence des solutions réelles est assurée, elle dépend néanmoins de paramètres implicites pour lesquels il n'y a pas de formule de calcul établie.

Ces deux fermetures sont construites pour préserver la propriété *well-balanced* d'abord pour le lac au repos puis pour toutes les solutions stationnaires régulières. Pour les deux fermetures, si la séquence  $(\hat{w}_i^n)_{i \in \mathbb{Z}}$  n'est pas dans un *voisinage* immédiat du bord de  $\hat{\Omega}$ , les schémas numériques obtenus sont de plus entropiques. Dans le cas contraire, la stabilité entropique pourrait être localement perdue.

Les deux fermetures proposées peuvent d'interpréter comme des généralisations successives des schémas introduits en Section 2.1. En particulier, la seconde fermeture montre l'existence d'un schéma numérique entropique et *well-balanced* pour toutes les solutions stationnaires régulières des équations de Saint Venant.

Les deux schémas donnent des résultats numériques très satisfaisants en présence de solutions constituées d'onde de choc. Toutefois, la propriété *well-balanced* n'est pas restituée à la précision machine attendue dans le cas des équilibres mouvants atteints après régimes transitoires. L'origine de ce défaut est inconnue mais une hypothèse est avancée en Section 3.3.1.2. De plus, les deux schémas convergent parfois vers des solutions qui ne sont pas celles de référence. De tels écarts de convergence ont déjà été observés avec d'autres schémas numériques [34, 42, 59].

La seconde partie de ce chapitre établit d'abord des conditions suffisantes de stabilité entropique pour un schéma numérique dédié aux équations de Ripa.

Ensuite, le chapitre s'articule autour d'équations non linéaires qui définissent un solveur de Riemann constant par morceau à deux états intermédiaires. Si des solutions explicites sont parfois accessibles, l'existence des solutions aux équations proposées n'est malheureusement pas complètement établie. Néanmoins, en admettant cette existence, ces équations permettent de satisfaire les conditions de stabilité entropique proposées. Par conséquent, elles sont utilisées pour définir un schéma de type Godunov entropique et *well-balanced* pour le lac au repos. Ces deux propriétés sont démontrées mais l'existence théorique du schéma est de fait incomplète.

En pratique, et comme les équations proposées sont algébriques, ce schéma entropique et *well-balanced* peut être mis en œuvre avec une méthode de Newton ou de point fixe par exemple. Les résultats obtenus avec une méthode de Newton demeurent convenables.

## Chapitre 4

# Schémas d'ordre élevé, non limités et à stabilité entropique discrète globale

Ce chapitre concerne la stabilité entropique globale de schémas numériques d'ordre élevé. Ces schémas approchent les solutions d'un système hyperbolique homogène de  $m$  lois de conservation donné par

$$\begin{cases} \partial_t w + \partial_x f(w) = 0, & t > 0, \quad x \in \mathbb{R}, \\ w(x, t = 0) = w_0(x). \end{cases} \quad (4.1)$$

Les solutions  $w$  du système d'équations ci-dessus prennent leurs valeurs dans un ensemble convexe  $\Omega$  inclus dans  $\mathbb{R}^m$ . Le système (4.1) est muni d'une entropie strictement convexe  $\eta$  associée à un flux d'entropie  $G$ . Par conséquent, ce chapitre s'intéresse aux solutions de (4.1) qui vérifient une inégalité d'entropie de la forme

$$\partial_t \eta(w) + \partial_x G(w) \leq 0.$$

Pour des solutions telles que  $(x, t) \mapsto \eta(w(x, t))$  est à support compact ou de limite nulle en l'infini, l'intégration formelle de l'inégalité locale ci-dessus selon la variable  $x$  donne l'inégalité d'entropie globale suivante :

$$\frac{d}{dt} \int_{\mathbb{R}} \eta(w(x, t)) dx \leq 0, \quad \forall t > 0. \quad (4.2)$$

Les schémas d'ordre élevé utilisent des limiteurs de pente [55, 57, 94, 116, 182, 194, 196, 219] et ils vérifient souvent des stabilités entropiques semi-discrètes [7, 45, 51, 57, 92, 94, 137, 182]. L'objectif de ce chapitre est de définir des schémas numériques d'ordre élevé au sens de la Définition 1.4.1.1, sans limiteur de pente et dont la séquence mise à jour  $(w_i^{n+1})_{i \in \mathbb{Z}}$  vérifie un analogue discret de (4.2).

En faisant une hypothèse de régularité sur  $\eta$ , ce chapitre introduit la notion de schémas numériques dissipatifs puis montre comment cette notion permet d'obtenir une stabilité entropique discrète globale.

### 4.1 Schémas numériques dissipatifs

La notion de schéma numérique dissipatif repose sur une inégalité stricte globale qui contraint le bilan du flux numérique  $\mathcal{F}$  relativement à une entropie  $\eta$ . En supposant une régularité  $C^1$  de l'entropie, il est possible de proposer une définition de schéma numérique dissipatif.

**Définition 4.1.1** (Schéma numérique volumes finis dissipatif). *Soient  $l$  un entier et  $\mathcal{F} : (\mathbb{R}^m)^{2l+2} \rightarrow \mathbb{R}^m$  une fonction continue qui satisfait*

$$\mathcal{F}(w, \dots, w) = f(w), \quad \forall w \in \Omega.$$

*Pour approcher les solutions du système d'équations (4.1), considérons un schéma numérique volumes finis écrit sous sa forme conservative*

$$w_i^{n+1} = w_i^n - \frac{\Delta t}{\Delta x} (\mathcal{F}(w_{i-l}^n, \dots, w_{i+l+1}^n) - \mathcal{F}(w_{i-l-1}^n, \dots, w_{i+1}^n)), \quad \forall i \in \mathbb{Z}.$$

Supposons que l'entropie  $\eta$  associée au système (4.1) est une fonction de classe  $C^1(\Omega, \mathbb{R})$  et notons le produit scalaire usuel de  $\mathbb{R}^m$ . Un schéma numérique (4.3) (resp. flux numérique  $\mathcal{F}$ ) est dit dissipatif relativement à  $\eta$  si l'inégalité stricte suivante est satisfaite pour toute séquence  $(w_i^n)_{i \in \mathbb{Z}}$  non constante incluse dans  $\Omega$  :

$$\sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot \frac{\mathcal{F}(w_{i-l}^n, \dots, w_{i+l+1}^n) - \mathcal{F}(w_{i-l-1}^n, \dots, w_{i+1}^n)}{\Delta x} > 0.$$

L'inégalité ci-dessus est appelée inégalité de dissipation (relativement à  $\eta$ ).

À partir de la définition 4.1.1 et puisqu'il n'y a pas de confusion possible, la notation  $\mathcal{F}_{i+1/2} = \mathcal{F}(w_{i-l}^n, \dots, w_{i+l+1}^n)$  sera désormais employée. Par conséquent, un schéma numérique conservatif (4.3) et l'inégalité de dissipation relativement à  $\eta$  (4.4) se lisent respectivement

$$w_i^{n+1} = w_i^n - \frac{\Delta t}{\Delta x} (\mathcal{F}_{i+1/2} - \mathcal{F}_{i-1/2}), \quad \forall i \in \mathbb{Z}, \quad (4.3)$$

$$\sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot \frac{\mathcal{F}_{i+1/2} - \mathcal{F}_{i-1/2}}{\Delta x} > 0. \quad (4.4)$$

En supposant une régularité suffisante de l'entropie  $\eta$ , le membre de gauche de l'inégalité de dissipation donnée ci-dessus est exactement le terme d'ordre un du développement de Taylor de  $\sum_{i \in \mathbb{Z}} \eta(w_i^{n+1}) \Delta x$  où  $w_i^{n+1}$  est donnée par (4.3). Ainsi, l'inégalité de dissipation (4.4) peut entraîner une inégalité d'entropie discrète globale de la forme

$$\sum_{i \in \mathbb{Z}} \frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} \Delta x \leq 0.$$

Comme cette dernière inégalité repose sur l'évaluation de  $\sum_{i \in \mathbb{Z}} \eta(w_i^{n+1}) \Delta x$ , il faut d'abord s'assurer que la séquence mise à jour  $(w_i^{n+1})_{i \in \mathbb{Z}}$  demeure dans  $\Omega$  et que les sommes utilisées pour exprimer cette quantité sont toutes convergentes. Pour cela, deux lemmes techniques sont maintenant présentés et le premier repose sur un argument topologique et une restriction de  $\Delta t > 0$  qui permettent d'obtenir  $(w_i^{n+1})_{i \in \mathbb{Z}} \subset \Omega$ .

**Lemme 4.1.1** (Robustesse des schémas numériques). *Soit un schéma numérique conservatif (4.3) qui approche les solutions du système d'équations (4.1). Supposons que ces solutions prennent leurs valeurs dans un ensemble strictement convexe non vide  $\Omega$  inclus dans  $\mathbb{R}^m$  et soit  $K^n$  un ensemble compact inclus dans  $\Omega$ . Supposons que la séquence non constante  $(w_i^n)_{i \in \mathbb{Z}}$  est incluse dans  $K^n$ . Il existe une constante implicite  $c^n > 0$  telle que si  $0 < \frac{\Delta t}{\Delta x} \leq c^n$  alors la séquence  $(w_i^{n+1})_{i \in \mathbb{Z}}$  donnée par le schéma (4.3) est incluse dans  $\Omega$*

*Démonstration.* Soit  $\|\cdot\|$  une norme de  $\mathbb{R}^m$  et considérons la fonction de distance signée au bord de  $\Omega$  associée à  $\|\cdot\|$ . Pour tout  $w$  de  $\mathbb{R}^m$ , cette fonction de distance signée s'écrit

$$\text{dist}(w, \partial\Omega) = \begin{cases} \inf_{y \in \partial\Omega} \|w - y\|, & \text{si } w \in \overline{\Omega}, \\ - \inf_{y \in \partial\Omega} \|w - y\| & \text{si } w \in \mathbb{R}^m \setminus \overline{\Omega}. \end{cases}$$

D'après la définition du schéma conservatif (4.3) et comme la fonction ci-dessus est une fonction continue 1-Lipschitz de  $\mathbb{R}^m$ , il suit pour tout  $i$  dans  $\mathbb{Z}$

$$\left| \text{dist}(w_i^{n+1}, \partial\Omega) - \text{dist}(w_i^n, \partial\Omega) \right| \leq \|w_i^{n+1} - w_i^n\| = \frac{\Delta t}{\Delta x} \|\mathcal{F}_{i+1/2} - \mathcal{F}_{i-1/2}\|.$$

L'inégalité ci-dessus donne immédiatement

$$\text{dist}(w_i^{n+1}, \partial\Omega) \geq \text{dist}(w_i^n, \partial\Omega) - \frac{\Delta t}{\Delta x} \|\mathcal{F}_{i+1/2} - \mathcal{F}_{i-1/2}\|. \quad (4.5)$$

Comme le compact  $K^n$  est inclus dans l'ensemble ouvert  $\Omega$ , on a  $\text{dist}(K, \partial\Omega) > 0$  et en particulier

$$\text{dist}(w_i^n, \partial\Omega) \geq \inf_{w \in K^n} \text{dist}(w, \partial\Omega) = \min_{w \in K^n} \text{dist}(w, \partial\Omega) = \text{dist}(K, \partial\Omega) > 0. \quad (4.6)$$

Puisque  $(w_i^n)_{i \in \mathbb{Z}}$  n'est pas constante, la continuité du flux numérique  $\mathcal{F}$  sur le compact  $K^n$  permet de déduire l'existence d'une constante  $R^n > 0$  telle que pour tout  $i$  dans  $\mathbb{Z}$

$$\left\| \mathcal{F}_{i+\frac{1}{2}} - \mathcal{F}_{i-\frac{1}{2}} \right\| \leq R^n.$$

En introduisant l'inégalité ci-dessus ainsi que de l'inégalité (4.6) dans le résultat intermédiaire (4.5), on en déduit

$$\text{dist}(w_i^{n+1}, \partial\Omega) \geq \min_{w \in K^n} \text{dist}(w, \partial\Omega) - \frac{\Delta t}{\Delta x} R^n, \quad \forall i \in \mathbb{Z}.$$

Si  $0 < \frac{\Delta t}{\Delta x} \leq \frac{\min_{w \in K^n} \text{dist}(w, \partial\Omega)}{2R^n}$  alors  $\text{dist}(w_i^{n+1}, \partial\Omega) \geq \min_{w \in K^n} \text{dist}(w, \partial\Omega)/2 > 0$  pour tout  $i$  dans  $\mathbb{Z}$ . Comme cette dernière inégalité signifie que  $w_i^{n+1}$  est dans  $\Omega$ , ceci conclut la preuve.  $\square$

Le lemme de robustesse ci-dessus s'appuie sur des arguments topologiques qui conduisent à une restriction implicite sur  $\Delta t > 0$ . Par conséquent, cette restriction est difficile à utiliser en pratique. D'autres preuves de robustesse existent mais elles reposent sur des techniques de limitations [55, 57, 94, 116, 182, 194, 196, 219].

À l'aide du lemme de robustesse précédent, il est désormais possible de montrer qu'un schéma dissipatif tel qu'introduit dans la Définition 4.1.1 peut satisfaire une inégalité d'entropie discrète globale.

**Proposition 4.1.1** (Stabilité entropique globale des schémas numériques dissipatifs). *Considérons un schéma numérique volumes finis dissipatif (4.3) vérifiant l'inégalité (4.4) et qui approche les solutions du système d'équations (4.1). Considérons une séquence  $(w_i^n)_{i \in \mathbb{Z}}$  non constante, incluse dans un compact de  $\Omega$ .*

*Si l'entropie  $\eta$  associée au système (4.1) est une fonction de classe  $C^2(\Omega, \mathbb{R})$  strictement convexe et si pour tout  $\Delta x > 0$ ,  $(w_i^n)_{i \in \mathbb{Z}}$  est telle que*

$$\left| \sum_{i \in \mathbb{Z}} \eta(w_i^n) \Delta x \right| < +\infty, \quad (4.7a)$$

$$\left| \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot \frac{\mathcal{F}_{i+\frac{1}{2}} - \mathcal{F}_{i-\frac{1}{2}}}{\Delta x} \right| \Delta x < +\infty, \quad (4.7b)$$

$$\int_0^1 (1-s) \sum_{i \in \mathbb{Z}} \nabla^2 \eta \left( w_i^n - s \frac{\Delta t}{\Delta x} (\mathcal{F}_{i+\frac{1}{2}} - \mathcal{F}_{i-\frac{1}{2}}) \right) \frac{\mathcal{F}_{i+\frac{1}{2}} - \mathcal{F}_{i-\frac{1}{2}}}{\Delta x} \cdot \frac{\mathcal{F}_{i+\frac{1}{2}} - \mathcal{F}_{i-\frac{1}{2}}}{\Delta x} \Delta x \, ds < +\infty, \quad (4.7c)$$

*alors il existe  $\Delta t > 0$  suffisamment petit tel que la séquence mise à jour  $(w_i^{n+1})_{i \in \mathbb{Z}}$  est incluse dans  $\Omega$  et tel que l'inégalité d'entropie discrète globale suivante est satisfaite :*

$$\sum_{i \in \mathbb{Z}} \frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} \Delta x \leq 0. \quad (4.8)$$

*Démonstration.* D'après le Lemme 4.1.1, il existe  $\Delta t > 0$  tel que la séquence  $(w_i^{n+1})_{i \in \mathbb{Z}}$  est incluse dans  $\Omega$ . Supposons que cette condition est vérifiée et notons

$$\mathcal{R}_i = -(\mathcal{F}_{i+\frac{1}{2}} - \mathcal{F}_{i-\frac{1}{2}}).$$

Ainsi, l'application de  $\eta$  au schéma numérique (4.3) s'écrit  $\eta(w_i^{n+1}) = \eta(w_i^n + \Delta t \mathcal{R}_i / \Delta x)$ . Comme  $\eta$  est une fonction de classe  $C^2$ , un développement de Taylor avec reste intégral de l'expression ci-avant donne

$$\eta(w_i^{n+1}) = \eta(w_i^n) + \frac{\Delta t}{\Delta x} \nabla \eta(w_i^n) \cdot \mathcal{R}_i + \left( \frac{\Delta t}{\Delta x} \right)^2 \int_0^1 (1-s) \nabla^2 \eta \left( w_i^n + s \frac{\Delta t}{\Delta x} \mathcal{R}_i \right) \mathcal{R}_i \cdot \mathcal{R}_i \, ds.$$

En réorganisant ses termes puis en sommant sur  $i$ , l'égalité ci-dessus devient

$$\begin{aligned} \sum_{i \in \mathbb{Z}} \frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} \Delta x &= \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot \mathcal{R}_i \\ &+ \frac{\Delta t}{\Delta x} \sum_{i \in \mathbb{Z}} \int_0^1 (1-s) \nabla^2 \eta \left( w_i^n + s \frac{\Delta t}{\Delta x} \mathcal{R}_i \right) \mathcal{R}_i \cdot \mathcal{R}_i ds. \end{aligned} \quad (4.9)$$

Puisque la séquence  $(w_i^n)_{i \in \mathbb{Z}}$  est incluse dans un compact de  $\Omega$ , les hypothèses de sommabilités données par (4.7) assurent que les sommes ci-dessus sont toutes convergentes. De plus, comme cette séquence n'est pas constante, la convexité stricte de l'entropie  $\eta$  permet de déduire que l'inégalité stricte suivante est satisfaite :

$$\int_0^1 (1-s) \sum_{i \in \mathbb{Z}} \nabla^2 \eta \left( w_i^n + s \frac{\Delta t}{\Delta x} \mathcal{R}_i \right) \mathcal{R}_i \cdot \mathcal{R}_i ds > 0.$$

Comme le schéma numérique est dissipatif, l'inégalité stricte de dissipation (4.4) se lit dans ce cas

$$\sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot \mathcal{R}_i < 0.$$

Par conséquent, et d'après (4.9), l'inégalité d'entropie globale discrète (4.8) est satisfaite si  $\Delta t > 0$  vérifie

$$0 < \frac{\Delta t}{\Delta x} \leq \frac{- \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot \mathcal{R}_i}{\int_0^1 (1-s) \sum_{i \in \mathbb{Z}} \nabla^2 \eta \left( w_i^n + s \frac{\Delta t}{\Delta x} \mathcal{R}_i \right) \mathcal{R}_i \cdot \mathcal{R}_i ds}. \quad (4.10)$$

□

En supposant une régularité  $C^2$  de l'entropie  $\eta$ , la preuve ci-dessus donne l'inégalité (4.10) qui s'identifie à une condition CFL. En raison de la quantité  $\nabla^2 \eta(w_i^n + s \Delta t \mathcal{R}_i / \Delta x)$ , cette condition CFL demeure implicite sauf dans le cas d'une entropie quadratique pour laquelle  $\nabla^2 \eta = I_m$ . Par conséquent, l'utilisation pratique de la condition CFL (4.10) nécessite en général la résolution d'une équation non linéaire.

Cette condition CFL est globale ce qui signifie qu'elle dépend de quantités de la forme  $\sum_{i \in \mathbb{Z}}$  et cette dépendance justifie les hypothèses de sommabilité données dans (4.7). Ces hypothèses de sommabilité peuvent par exemple être satisfaites en introduisant des espaces de Sobolev discrets homogènes pour la séquence  $(w_i^n)_{i \in \mathbb{Z}}$  ou bien en supposant que la séquence  $(w_{i+1}^n - w_i^n)_{i \in \mathbb{Z}}$  est à support compact. Ces deux exemples seront présentés dans les sections suivantes.

La Proposition 4.1.1 donne un cadre suffisant à la stabilité entropique discrète globale (4.8). Ce cadre théorique repose en partie sur l'inégalité de dissipation (4.4) et l'équation (4.9) montre que cette inégalité est nécessaire à la stabilité entropique globale. Ainsi, les schémas numériques du second ordre proposés dans ce chapitre vont maintenant être définis pour satisfaire l'inégalité de dissipation (4.4).

## 4.2 Schémas non limités d'ordre deux

La famille de schémas du second ordre proposée est obtenue à partir du schéma HLL symétrique [118] aussi appelé schéma de Rusanov [185]. Pour un réel  $\lambda > 0$  donné, le flux numérique  $\mathcal{F}_\lambda^{O1}$  de ce schéma s'écrit

$$\mathcal{F}_\lambda^{O1}(w_i^n, w_{i+1}^n) = \frac{f(w_i^n) + f(w_{i+1}^n)}{2} - \frac{\lambda}{2} (w_{i+1}^n - w_i^n). \quad (4.11)$$

En considérant une régularité de  $w$  suffisante et en s'appuyant sur la consistance du flux numérique ci dessus, un développement de Taylor autour de l'interface  $x_{i+\frac{1}{2}}$  donne

$$\mathcal{F}_\lambda^{O1}(w_i^n, w_{i+1}^n) = f(w(x_{i+\frac{1}{2}}, t^n)) - \frac{\lambda \Delta x}{2} \partial_x w(x_{i+\frac{1}{2}}, t^n) + \mathcal{O}(\Delta x^2). \quad (4.12)$$

En utilisant la Définition 1.4.1.1 et d'après l'expression ci-dessus, une montée à l'ordre deux peut être obtenue en ajoutant au flux numérique (4.11) un terme consistant avec la quantité  $+\lambda\partial_x w(x_{i+\frac{1}{2}}, t^n)/2$ . Ce terme additionnel peut se construire à partir d'une séquence  $(\alpha_i^n)_{i \in \mathbb{Z}}$  de  $\mathbb{R}^m$  telle que

$$\alpha_i^n = \frac{\lambda\Delta x}{2} \partial_x w(x_i, t^n) + \mathcal{O}(\Delta x^2). \quad (4.13)$$

La moyenne arithmétique de  $\alpha_i^n$  et de  $\alpha_{i+1}^n$  est consistante avec  $+\lambda\partial_x w(x_{i+\frac{1}{2}}, t^n)/2$  et par conséquent, les schémas du second ordre proposés sont définis par un flux numérique  $\mathcal{F}_\lambda^{O2}$  de la forme

$$\begin{aligned} \mathcal{F}_\lambda^{O2}(w_{i-1}^n, w_i^n, w_{i+1}^n, w_{i+2}^n) &= \mathcal{F}_\lambda^{O1}(w_i^n, w_{i+1}^n) + \frac{1}{2}(\alpha_i^n + \alpha_{i+1}^n), \\ &= \frac{f(w_i^n) + f(w_{i+1}^n)}{2} - \frac{\lambda}{2}(w_{i+1}^n - w_i^n) + \frac{1}{2}(\alpha_i^n + \alpha_{i+1}^n). \end{aligned}$$

Par construction, le flux numérique  $\mathcal{F}_\lambda^{O2}$  contient une approximation de  $-\lambda\Delta x\partial_x w$  héritée du terme  $-\lambda(w_{i+1}^n - w_i^n)$ . Ce flux numérique contient de plus une approximation de  $+\lambda\Delta x\partial_x w$  en vertu de l'estimation (4.13) vérifiée par la quantité  $\alpha_i^n$ . Ainsi, le terme  $-\lambda(w_{i+1}^n - w_i^n) + \alpha_i^n + \alpha_{i+1}^n$  est un résidu de la forme  $\mathcal{O}(\Delta x^2)$ . Pour  $\Delta x > 0$  donné, ce résidu n'est en général pas nul et l'idée développée est de proposer une séquence  $(\alpha_i^n)_{i \in \mathbb{Z}}$  satisfaisant (4.13) et telle que  $-\lambda(w_{i+1}^n - w_i^n) + \alpha_i^n + \alpha_{i+1}^n$  assure l'inégalité stricte de dissipation (4.4).

**Proposition 4.2.1** (Schéma numérique du second ordre spatial). *Considérons une séquence  $(w_i^n)_{i \in \mathbb{Z}}$  non constante, à support compact et incluse dans  $\Omega$ . Soit  $(\Theta_i^n)_{i \in \mathbb{Z}}$  une séquence de matrices diagonales de taille  $m \times m$  et notons  $I_m$  la matrice identité de taille  $m$ . Soit  $(\alpha_i^n)_{i \in \mathbb{Z}}$  la séquence de  $\mathbb{R}^m$  donnée par*

$$\alpha_i^n = \frac{\lambda}{2}(\Theta_i^n (w_{i+1}^n - w_i^n) + (I_m - \Theta_i^n)(w_i^n - w_{i-1}^n)), \quad \forall i \in \mathbb{Z}. \quad (4.14)$$

Soit un schéma numérique volumes finis conservatif (4.3) défini par un flux numérique  $\mathcal{F}_{i+\frac{1}{2}}^n$  donné par

$$\mathcal{F}_{i+\frac{1}{2}}^n = \frac{f(w_i^n) + f(w_{i+1}^n)}{2} - \frac{\lambda}{2}(w_{i+1}^n - w_i^n) + \frac{1}{2}(\alpha_i^n + \alpha_{i+1}^n). \quad (4.15)$$

Si les matrices  $\Theta_i^n$  sont uniformément bornées selon l'indice  $i$  et bornées lorsque  $\Delta x \rightarrow 0$  alors le schéma numérique ainsi défini est du second ordre spatial au sens de la Définition formelle 1.4.1.1-iii).

*Démonstration.* Considérons  $u : \mathbb{R} \times \mathbb{R}^+ \rightarrow \mathbb{R}^m$  une fonction régulière à support compact et notons

$$u_i^n = \frac{1}{\Delta x} \int_{C_i} u(x, t^n) dx, \quad \forall i \in \mathbb{Z}.$$

En introduisant  $(u_i^n)_{i \in \mathbb{Z}}$  donnée ci-dessus dans la définition (4.14),  $\alpha_i^n$  s'écrit

$$\alpha_i^n = \frac{\lambda}{2}(\Theta_i^n (u_{i+1}^n - u_i^n) + (I_m - \Theta_i^n)(u_i^n - u_{i-1}^n)).$$

Comme la fonction  $u$  est régulière et comme les matrices  $\Theta_i^n$  sont bornées, un développement de Taylor dans l'expression ci-dessus donne

$$\alpha_i^n = \frac{\lambda\Delta x}{2} \partial_x u(x_i, t^n) + \mathcal{O}(\Delta x^2).$$

Par conséquent, le flux numérique (4.15) vérifie

$$\begin{aligned} \mathcal{F}_{i+\frac{1}{2}}^n &= f(u(x_{i+\frac{1}{2}}, t^n)) - \frac{\lambda\Delta x}{2} \partial_x u(x_{i+\frac{1}{2}}, t^n) + \frac{\lambda\Delta x}{2} \frac{\partial_x u(x_i, t^n) + \partial_x u(x_{i+1}, t^n)}{2} + \mathcal{O}(\Delta x^2), \\ &= f(u(x_{i+\frac{1}{2}}, t^n)) + \mathcal{O}(\Delta x^2). \end{aligned}$$

Puisque cette dernière égalité s'identifie à la condition du second ordre spatial donnée dans la Définition 1.4.1.1-iii), ceci conclut la preuve.  $\square$



À une multiplication par  $2/(\lambda\Delta x)$  près, la quantité  $\alpha_i^n$  donnée par (4.14) est une approximation de la dérivée de  $w$  à la cellule  $C_i$  mais contrairement aux schémas MUSCL [219], (W)ENO [171, 191, 194] ou DG [85, 125, 194], cette approximation ne nécessite pas de procédure de limitation. Par simplicité,  $\alpha_i^n$  est défini avec des matrices  $\Theta_i^n$  diagonales mais des matrices plus générales pourraient être considérées. Ces matrices sont formellement des paramètres libres qui vont maintenant être sélectionnés pour satisfaire l'inégalité stricte de dissipation (4.4). Cette inégalité entraînera ensuite, en vertu de la Proposition 4.1.1, la stabilité entropique globale du schéma du second ordre défini par son flux numérique (4.15).

## 4.3 Stabilité entropique discrète globale

### 4.3.1 Notations et sommabilité

Le schéma défini par la Proposition 4.2.1 est formellement du second ordre et cette section détaille ses propriétés de stabilité. Pour établir ces propriétés, il est pratique de définir les espaces de Sobolev discrets

$$\begin{aligned} l^2(\mathbb{Z}) &= \left\{ v \in \mathbb{R}^{\mathbb{Z}} \mid \sum_{i \in \mathbb{Z}} |v_i|^2 \Delta x < +\infty \right\}, \\ h^2(\mathbb{Z}) &= \left\{ v \in l^2(\mathbb{Z}) \mid \left( \frac{v_{i+1} - v_i}{\Delta x} \right)_{i \in \mathbb{Z}} \in l^2(\mathbb{Z}), \left( \frac{v_{i+1} - 2v_i + v_{i-1}}{\Delta x^2} \right)_{i \in \mathbb{Z}} \in l^2(\mathbb{Z}) \right\}, \\ \dot{h}^2(\mathbb{Z}, \Omega) &= \left\{ w \in \Omega^{\mathbb{Z}} : \left( \frac{w_{i+1} - w_i}{\Delta x} \right)_{i \in \mathbb{Z}} \in l^2(\mathbb{Z}), \left( \frac{w_{i+1} - 2w_i + w_{i-1}}{\Delta x^2} \right)_{i \in \mathbb{Z}} \in l^2(\mathbb{Z}) \right\}. \end{aligned}$$

Pour  $\eta$  une entropie strictement convexe de classe  $C^1(\Omega, \mathbb{R})$ , ces espaces de Sobolev sont également complétés par l'espace special

$$S_\eta^{O2}(\mathbb{Z}, \Omega) = \left\{ w \in \dot{h}^2(\mathbb{Z}, \Omega) : \sum_{j=1}^m \sum_{i \in \mathbb{Z}} |(\nabla \eta(w_{i+1}) - \nabla \eta(w_{i-1}))_j| |(w_{i+1} - 2w_i + w_{i-1}))_j| \neq 0 \right\}. \quad (4.16)$$

De plus, en notant  $\mathcal{M}_{2m,2m}(\mathbb{R})$  l'ensemble des matrices carrées de taille  $2m \times 2m$  à coefficients réels et avec  $\delta_{i+\frac{1}{2}}^n = w_{i+1}^n - w_i^n$  dans  $\mathbb{R}^m$ , des matrices par blocs  $D_i^n$ ,  $N_i^n$ ,  $P_i^n$  et  $F_i^n$  sont définies par

$$D_i^n = \begin{pmatrix} \delta_{i-\frac{1}{2}}^n \\ \delta_{i+\frac{1}{2}}^n \end{pmatrix} \in \mathbb{R}^{2m}, \quad (4.17a)$$

$$N_i^n(s) = \begin{pmatrix} \nabla^2 \eta(w_i^n - s\delta_{i-\frac{1}{2}}^n) & 0 \\ 0 & \nabla^2 \eta(w_i^n + s\delta_{i+\frac{1}{2}}^n) \end{pmatrix} \in \mathcal{M}_{2m,2m}(\mathbb{R}), \quad (4.17b)$$

$$P_i^n(s) = \begin{pmatrix} (1-2s)I + \Theta_i^n & -\Theta_i^n \\ \Theta_i^n - I & 2(1-s)I - \Theta_i^n \end{pmatrix} \in \mathcal{M}_{2m,2m}(\mathbb{R}), \quad (4.17c)$$

$$F_i^n(s) = \begin{pmatrix} -\nabla f(w_i^n - s\delta_{i-\frac{1}{2}}^n) & 0 \\ 0 & \nabla f(w_i^n + s\delta_{i+\frac{1}{2}}^n) \end{pmatrix} \in \mathcal{M}_{2m,2m}(\mathbb{R}), \quad \forall s \in [0, 1]. \quad (4.17d)$$

L'objectif est d'utiliser la Proposition 4.1.1 pour montrer que le schéma du second ordre défini par le flux numérique (4.15) vérifie une stabilité entropique discrète globale de la forme (4.8). Comme la Proposition 4.1.1 repose sur des formulations sommées sur l'indice  $i$  dans  $\mathbb{Z}$ , il est d'abord nécessaire démontrer la convergence de certaines quantités définies avec les matrices par blocs données par (4.17). Le premier résultat est un lemme technique qui traite des applications à valeurs dans  $\mathcal{M}_{m,m}(\mathbb{R})$

**Lemme 4.3.1.1.** *Considérons  $\Omega$  un ensemble convexe non vide inclus dans  $\mathbb{R}^m$ . Soient  $K$  un ensemble compact convexe inclus dans  $\Omega$  et  $(a, b)$  deux éléments de  $K$ . Si  $A : \Omega \rightarrow \mathcal{M}_{m,m}(\mathbb{R})$  est une application continue alors pour tout doublet  $(a, b)$  dans  $K^2$ , l'application  $s \in [0, 1] \mapsto A((1-s)a + sb)$  est continue sur le segment  $[0, 1]$ . De plus, en notant  $\|\cdot\|$  une norme matricielle quelconque, l'estimation suivante est vérifiée*

$$\sup_{s \in [0,1]} \|A((1-s)a + sb)\| \leq \max_{w \in K} \|A(w)\|.$$

*Démonstration.* Puisque  $a$  et  $b$  sont dans un compact convexe  $K$ , il suit immédiatement pour tout  $s$  dans  $[0, 1]$

$$(1-s)a + sb \in K \subset \Omega.$$

Par conséquent, l'application  $A : s \in [0, 1] \mapsto \|A((1-s)a + sb)\|$  est bien définie et cette application est continue car il s'agit d'une composée de fonctions continues.

Comme le segment  $[0, 1]$  est un compact de  $\mathbb{R}$ , il existe  $s^*$  dans  $[0, 1]$  tel que

$$\sup_{s \in [0,1]} \|A((1-s)a + sb)\| = \|A((1-s^*)a + s^*b)\|.$$

Ainsi, et puisque  $(1-s^*)a + s^*b$  est un élément de l'ensemble compact  $K$ , il vient

$$\|A((1-s^*)a + s^*b)\| \leq \sup_{w \in K} \|A(w)\| = \max_{w \in K} \|A(w)\|.$$

Cette dernière expression est exactement l'estimation annoncée et par conséquent, ceci conclut la preuve.  $\square$

En s'appuyant sur le lemme ci-dessus, il est maintenant possible de montrer la convergence de certaines sommes globales.

**Lemme 4.3.1.2.** *Soient  $\eta$  une entropie strictement convexe de classe  $C^2(\Omega, \mathbb{R})$  et  $(w_i^n)_{i \in \mathbb{Z}}$  une séquence dans  $S_\eta^{O2}(\mathbb{Z}, \Omega)$ . Considérons les matrices par blocs  $D_i^n$ ,  $N_i^n(s)$ ,  $P_i^n(s)$  et  $F_i^n(s)$  définies par (4.17) et supposons l'existence de  $K^n$  un ensemble compact convexe inclus dans  $\Omega$ . Si  $(w_i^n) \subset K^n$  alors*

$$\left| \sum_{i \in \mathbb{Z}} \int_0^1 s \left( \int_0^1 N_i^n(us) du \right) F_i^n(s) D_i^n \cdot D_i^n ds \right| < +\infty, \quad (4.18a)$$

$$\left| \sum_{i \in \mathbb{Z}} \int_0^1 N_i^n(s) P_i^n(s) D_i^n \cdot D_i^n ds \right| < +\infty. \quad (4.18b)$$

*Démonstration.* Concernant la sommabilité (4.18a), et en notant  $\|\cdot\|_2$  la norme euclidienne usuelle de  $\mathbb{R}^{2m}$ , l'inégalité de Cauchy-Schwarz donne

$$\left| \sum_{i \in \mathbb{Z}} \int_0^1 s \left( \int_0^1 N_i^n(us) du \right) F_i^n(s) D_i^n \cdot D_i^n ds \right| \leq \sum_{i \in \mathbb{Z}} \sup_{s \in [0,1]} \|N_i^n(s)\|_2 \sup_{s \in [0,1]} \|F_i^n(s)\|_2 \|D_i^n\|_2^2. \quad (4.19)$$

En s'appuyant sur l'équivalence des normes des espaces vectoriels normés en dimension finie, il existe une constante  $c(m) > 0$  dépendante uniquement de la dimension  $m$  et telle que

$$\sup_{s \in [0,1]} \|F_i^n(s)\|_2 \leq c(m) \sup_{s \in [0,1]} \|F_i^n(s)\|_\infty,$$

$$\sup_{s \in [0,1]} \|N_i^n(s)\|_2 \leq c(m) \sup_{s \in [0,1]} \|N_i^n(s)\|_\infty, \quad \forall i \in \mathbb{Z}.$$

Puisque  $(w_i^n)_{i \in \mathbb{Z}}$  est incluse dans le compact convexe  $K^n$ , l'application du Lemme 4.3.1.1 aux cas des matrices  $\nabla^2 \eta$  et  $\nabla f$  associées à la norme matricielle subordonnée  $\|\cdot\|_\infty$  donne

$$\sup_{s \in [0,1]} \|f((1-s)w_i + sw_{i-1})\|_\infty \leq \max_{w \in K^n} \|f(w)\|_\infty,$$

$$\sup_{s \in [0,1]} \|\nabla^2 \eta((1-s)w_i + sw_{i-1})\|_\infty \leq \max_{w \in K^n} \|\nabla^2 \eta(w)\|_\infty, \quad \forall i \in \mathbb{Z}.$$

Ainsi, en utilisant les définitions des matrices par blocs  $N_i^n(s)$  et  $F_i^n(s)$  respectivement données par (4.17b) et par (4.17d), il suit

$$\sup_{s \in [0,1]} \|N_i^n(s)\|_\infty \leq \max_{w \in K^n} \|\nabla^2 \eta(w)\|_\infty, \quad \sup_{s \in [0,1]} \|F_i^n(s)\|_\infty \leq \max_{w \in K^n} \|\nabla f(w)\|_\infty. \quad (4.20)$$

En introduisant ces deux expressions dans l'estimation (4.19), et puisque  $(w_i^n)$  est dans  $\dot{h}^2(\mathbb{Z}, \Omega)$ , on en déduit

$$\begin{aligned} & \left| \sum_{i \in \mathbb{Z}} \int_0^1 s \left( \int_0^1 N_i^n(us) du \right) F_i^n(s) D_i^n \cdot D_i^n ds \right| \\ & \leq \sum_{i \in \mathbb{Z}} \sup_{s \in [0,1]} \|N_i^n(s)\|_2 \sup_{s \in [0,1]} \|F_i^n(s)\|_2 \|D_i^n\|_2, \\ & \leq c(m)^2 \max_{w \in K^n} \|\nabla^2 \eta(w)\|_\infty \max_{w \in K^n} \|\nabla f(w)\|_\infty \sum_{i \in \mathbb{Z}} \|D_i^n\|_2^2 < +\infty. \end{aligned}$$

Ces dernières inégalités terminent de montrer le résultat (4.18a).

Pour la sommabilité donnée dans (4.18b), l'inégalité de Cauchy-Schwarz permet d'écrire

$$\left| \sum_{i \in \mathbb{Z}} \int_0^1 N_i^n(s) P_i^n(s) D_i^n \cdot D_i^n ds \right| \leq \sum_{i \in \mathbb{Z}} \sup_{s \in [0,1]} \|N_i^n(s)\|_2 \sup_{s \in [0,1]} \|P_i^n(s)\|_2 \|D_i^n\|_2^2. \quad (4.21)$$

Puisque les matrices  $(\Theta_i^n)_{i \in \mathbb{Z}}$  sont bornées selon l'indice  $i$ , et d'après la définition de la matrice par bloc  $P_i^n(s)$  donnée par (4.17c), il existe  $C > 0$  telle que

$$\sup_{s \in [0,1]} \|P_i^n(s)\|_2 \leq C, \quad \forall i \in \mathbb{Z}.$$

En introduisant le résultat ci-dessus ainsi que l'estimation (4.20) dans l'inégalité (4.21), on en déduit

$$\left| \sum_{i \in \mathbb{Z}} \int_0^1 N_i^n(s) P_i^n(s) D_i^n \cdot D_i^n ds \right| \leq c(m) \max_{w \in K^n} \|\nabla^2 \eta(w)\|_\infty C \sum_{i \in \mathbb{Z}} \|D_i^n\|_2^2 < +\infty.$$

Puisque cette inégalité termine de montrer le résultat (4.18b), ceci termine la preuve du lemme.  $\square$

Le lemme ci-dessus montre le caractère fini des quantités (4.18). Ces quantités sont maintenant utilisées pour reformuler l'inégalité de dissipation (4.4) dans le cas des schémas du second ordre définis par la Proposition 4.2.1.

### 4.3.2 Reformulation de l'inégalité de dissipation

En considérant le flux numérique du second ordre défini dans (4.15), le membre de gauche de l'inégalité de dissipation (4.4) s'écrit

$$\begin{aligned} \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot (\mathcal{F}_{i+\frac{1}{2}}^n - \mathcal{F}_{i-\frac{1}{2}}^n) &= \frac{1}{2} \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot (f(w_{i+1}^n) - f(w_{i-1}^n)) \\ & \quad \frac{1}{2} \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot (-\lambda(\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n) + \alpha_{i+1}^n - \alpha_{i-1}^n). \end{aligned}$$

Une reformulation des deux sommes ci-dessus à l'aide des définitions (4.17) est détaillée dans les deux prochains lemmes. Puisque la convergence des expressions est établie dans le Lemme 4.3.1.2, seules les reformulations sont énoncées.

**Lemme 4.3.2.1** (Reformulation du flux d'entropie global). *Soit  $\eta$  une entropie strictement convexe de classe  $C^2(\Omega, \mathbb{R})$ . Considérons une séquence  $(w_i^n)_{i \in \mathbb{Z}}$  dans  $S_\eta^{O2}(\mathbb{Z}, \Omega)$  et supposons l'existence de  $K^n$  un ensemble compact convexe inclus dans  $\Omega$  tel que  $(w_i^n)_{i \in \mathbb{Z}} \subset K^n$ . En considérant les matrices par blocs  $D_i^n$ ,  $N_i^n(s)$  et  $F_i^n(s)$  définies dans (4.17), l'égalité suivante est satisfaite :*

$$-\frac{1}{2} \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot (f(w_{i+1}^n) - f(w_{i-1}^n)) = \sum_{i \in \mathbb{Z}} \int_0^1 \frac{s}{2} \left( \int_0^1 N_i^n(us) du \right) F_i^n(s) D_i^n \cdot D_i^n ds.$$

*Démonstration.* Puisque les fonctions  $\eta$ ,  $G$  et  $f$  ont la régularité suffisante et que la séquence  $(w_i^n)_{i \in \mathbb{Z}}$  est incluse dans  $\Omega$ , un développement de Taylor avec reste intégral donne

$$\begin{aligned} G(w_{i+1}^n) &= G(w_i^n) + \int_0^1 \nabla G(w_i^n + s\delta_{i+\frac{1}{2}}^n) \cdot \delta_{i+\frac{1}{2}}^n ds, \\ \nabla \eta(w_i^n + s\delta_{i+\frac{1}{2}}^n) &= \nabla \eta(w_i^n) + s \int_0^1 \nabla^2 \eta(w_i^n + us\delta_{i+\frac{1}{2}}^n) \delta_{i+\frac{1}{2}}^n du, \\ f(w_{i+1}^n) &= f(w_i^n) + \int_0^1 \nabla f(w_i^n + s\delta_{i+\frac{1}{2}}^n) \delta_{i+\frac{1}{2}}^n ds. \end{aligned}$$

Par suite, et puisque le flux d'entropie  $G$  satisfait  $\nabla G(w)^\top = \nabla \eta(w)^\top \nabla f(w)$  pour tout  $w$  dans  $\Omega$ , les trois égalités ci-dessus permettent d'écrire

$$\begin{aligned} \nabla \eta(w_i^n) \cdot (f(w_{i+1}^n) - f(w_i^n)) &= \int_0^1 \nabla \eta(w_i^n) \cdot \nabla f(w_i^n + s\delta_{i+\frac{1}{2}}^n) \delta_{i+\frac{1}{2}}^n ds, \\ &= \int_0^1 \nabla G(w_i^n + s\delta_{i+\frac{1}{2}}^n) \cdot \delta_{i+\frac{1}{2}}^n ds - \int_0^1 \int_0^1 s \nabla^2 \eta(w_i^n + us\delta_{i+\frac{1}{2}}^n) \nabla f(w_i^n + s\delta_{i+\frac{1}{2}}^n) \delta_{i+\frac{1}{2}}^n \cdot \delta_{i+\frac{1}{2}}^n duds, \\ &= G(w_{i+1}^n) - G(w_i^n) - \int_0^1 \int_0^1 s \nabla^2 \eta(w_i^n + us\delta_{i+\frac{1}{2}}^n) \nabla f(w_i^n + s\delta_{i+\frac{1}{2}}^n) \delta_{i+\frac{1}{2}}^n \cdot \delta_{i+\frac{1}{2}}^n duds. \end{aligned} \tag{4.22}$$

De la même manière, il suit

$$\begin{aligned} \nabla \eta(w_i^n) \cdot (f(w_i^n) - f(w_{i-1}^n)) &= \\ G(w_i^n) - G(w_{i-1}^n) + \int_0^1 \int_0^1 s \nabla^2 \eta(w_i^n - us\delta_{i-\frac{1}{2}}^n) \nabla f(w_i^n - s\delta_{i-\frac{1}{2}}^n) \delta_{i-\frac{1}{2}}^n \cdot \delta_{i-\frac{1}{2}}^n duds. \end{aligned} \tag{4.23}$$

En sommant les égalités (4.22) et (4.23) sur l'indice  $i$  dans  $\mathbb{Z}$ , on obtient

$$\begin{aligned} &-\frac{1}{2} \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot (f(w_{i+1}^n) - f(w_{i-1}^n)) \\ &= -\frac{1}{2} \sum_{i \in \mathbb{Z}} \left( G(w_{i+1}^n) - G(w_i^n) - \int_0^1 \int_0^1 s \nabla^2 \eta(w_i^n + us\delta_{i+\frac{1}{2}}^n) \nabla f(w_i^n + s\delta_{i+\frac{1}{2}}^n) \delta_{i+\frac{1}{2}}^n \cdot \delta_{i+\frac{1}{2}}^n duds \right) \\ &\quad - \frac{1}{2} \sum_{i \in \mathbb{Z}} \left( G(w_i^n) - G(w_{i-1}^n) + \int_0^1 \int_0^1 s \nabla^2 \eta(w_i^n - us\delta_{i-\frac{1}{2}}^n) \nabla f(w_i^n - s\delta_{i-\frac{1}{2}}^n) \delta_{i-\frac{1}{2}}^n \cdot \delta_{i-\frac{1}{2}}^n duds \right), \\ &= \frac{1}{2} \sum_{i \in \mathbb{Z}} \int_0^1 \int_0^1 s \nabla^2 \eta(w_i^n + us\delta_{i+\frac{1}{2}}^n) \nabla f(w_i^n + s\delta_{i+\frac{1}{2}}^n) \delta_{i+\frac{1}{2}}^n \cdot \delta_{i+\frac{1}{2}}^n duds \\ &\quad - \frac{1}{2} \sum_{i \in \mathbb{Z}} \int_0^1 \int_0^1 s \nabla^2 \eta(w_i^n - us\delta_{i-\frac{1}{2}}^n) \nabla f(w_i^n - s\delta_{i-\frac{1}{2}}^n) \delta_{i-\frac{1}{2}}^n \cdot \delta_{i-\frac{1}{2}}^n duds. \end{aligned}$$

En récrivant l'égalité ci-dessus avec les définitions des matrices par blocs  $N_i^n(s)$ ,  $F_i^n(s)$  et  $D_i^n$  données dans (4.17), on en déduit le résultat annoncé ce qui termine la preuve.  $\square$

**Lemme 4.3.2.2** (Reformulation de la dissipation globale). *Soit  $\eta$  une entropie strictement convexe de classe  $C^2(\Omega, \mathbb{R})$ . Considérons une séquence  $(w_i^n)_{i \in \mathbb{Z}}$  dans  $S_\eta^{O2}(\mathbb{Z}, \Omega)$  et supposons l'existence de  $K^n$  un*

ensemble compact convexe inclus dans  $\Omega$  tel que  $(w_i^n)_{i \in \mathbb{Z}} \subset K^n$ . En considérant les matrices par blocs  $D_i^n$ ,  $N_i^n(s)$  et  $P_i^n(s)$  définies dans (4.17), l'égalité suivante est satisfaite :

$$\frac{1}{2} \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot (\lambda(\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n) - (\alpha_{i+1}^n - \alpha_{i-1}^n)) = -\frac{\lambda}{4} \sum_{i \in \mathbb{Z}} \int_0^1 N_i^n(s) P_i^n(s) D_i^n \cdot D_i^n ds.$$

*Démonstration.* Puisque la séquence  $(w_i^n)_{i \in \mathbb{Z}}$  est incluse dans  $\Omega$  et que l'entropie  $\eta$  est un fonction de classe  $C^2$ , un développement de Taylor avec reste intégral donne

$$\begin{aligned} \eta(w_{i+1}^n) &= \eta(w_i^n) + \nabla \eta(w_i^n) \cdot \delta_{i+\frac{1}{2}}^n + \int_0^1 (1-s) \nabla^2 \eta(w_i^n + s\delta_{i+\frac{1}{2}}^n) \delta_{i+\frac{1}{2}}^n \cdot \delta_{i+\frac{1}{2}}^n ds, \\ \eta(w_{i-1}^n) &= \eta(w_i^n) - \nabla \eta(w_i^n) \cdot \delta_{i-\frac{1}{2}}^n + \int_0^1 (1-s) \nabla^2 \eta(w_i^n - s\delta_{i-\frac{1}{2}}^n) \delta_{i-\frac{1}{2}}^n \cdot \delta_{i-\frac{1}{2}}^n ds. \end{aligned}$$

En utilisant les deux égalités ci-dessus, il suit

$$\begin{aligned} \frac{\lambda}{2} \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot (\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n) &= \frac{\lambda}{2} \sum_{i \in \mathbb{Z}} (\eta(w_{i+1}^n) - 2\eta(w_i^n) + \eta(w_{i-1}^n)) \\ &\quad - \frac{\lambda}{2} \int_0^1 (1-s) \sum_{i \in \mathbb{Z}} \nabla^2 \eta(w_i^n + s\delta_{i+\frac{1}{2}}^n) \delta_{i+\frac{1}{2}}^n \cdot \delta_{i+\frac{1}{2}}^n ds \\ &\quad - \frac{\lambda}{2} \int_0^1 (1-s) \sum_{i \in \mathbb{Z}} \nabla^2 \eta(w_i^n - s\delta_{i-\frac{1}{2}}^n) \delta_{i-\frac{1}{2}}^n \cdot \delta_{i-\frac{1}{2}}^n ds, \\ &= -\frac{\lambda}{2} \int_0^1 (1-s) \sum_{i \in \mathbb{Z}} N_i^n(s) D_i^n \cdot D_i^n ds. \end{aligned} \tag{4.24}$$

Cette dernière égalité donne une partie du résultat attendu et il est maintenant nécessaire de reformuler les quantités formées avec la suite  $(\alpha_i^n)_{i \in \mathbb{Z}}$  définie dans (4.14). Avec une translation d'indice, on obtient

$$\begin{aligned} &-\frac{1}{2} \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot (\alpha_{i+1}^n - \alpha_{i-1}^n) \\ &= \frac{1}{2} \sum_{i \in \mathbb{Z}} (\nabla \eta(w_{i+1}^n) - \nabla \eta(w_{i-1}^n)) \cdot \alpha_i^n, \\ &= \frac{\lambda}{4} \sum_{i \in \mathbb{Z}} (\nabla \eta(w_{i+1}^n) - \nabla \eta(w_{i-1}^n)) \cdot \Theta_i^n \delta_{i+\frac{1}{2}}^n + \frac{\lambda}{4} \sum_{i \in \mathbb{Z}} (\nabla \eta(w_{i+1}^n) - \nabla \eta(w_{i-1}^n)) \cdot (I_m - \Theta_i^n) \delta_{i-\frac{1}{2}}^n. \end{aligned} \tag{4.25}$$

En utilisant par ailleurs un développement de Taylor avec reste intégral au premier ordre, il est possible d'écrire

$$\begin{aligned} \nabla \eta(w_{i+1}^n) &= \nabla \eta(w_i^n) + \int_0^1 \nabla^2 \eta(w_i^n + s\delta_{i+\frac{1}{2}}^n) \delta_{i+\frac{1}{2}}^n ds, \\ \nabla \eta(w_{i-1}^n) &= \nabla \eta(w_i^n) - \int_0^1 \nabla^2 \eta(w_i^n - s\delta_{i-\frac{1}{2}}^n) \delta_{i-\frac{1}{2}}^n ds. \end{aligned}$$

L'introduction des deux égalités ci-dessus dans l'équation (4.25) permet de déduire

$$\begin{aligned}
& -\frac{1}{2} \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot (\alpha_{i+1}^n - \alpha_{i-1}^n) = \frac{\lambda}{4} \sum_{i \in \mathbb{Z}} \int_0^1 \nabla^2 \eta(w_i^n + s \delta_{i+\frac{1}{2}}^n) \Theta_i^n \delta_{i+\frac{1}{2}}^n \cdot \delta_{i+\frac{1}{2}}^n ds, \\
& + \frac{\lambda}{4} \sum_{i \in \mathbb{Z}} \int_0^1 \nabla^2 \eta(w_i^n + s \delta_{i+\frac{1}{2}}^n) (I_m - \Theta_i^n) \delta_{i-\frac{1}{2}}^n \cdot \delta_{i+\frac{1}{2}}^n ds \\
& + \frac{\lambda}{4} \sum_{i \in \mathbb{Z}} \int_0^1 \nabla^2 \eta(w_i^n - s \delta_{i-\frac{1}{2}}^n) \Theta_i^n \delta_{i+\frac{1}{2}}^n \cdot \delta_{i-\frac{1}{2}}^n ds \\
& + \frac{\lambda}{4} \sum_{i \in \mathbb{Z}} \int_0^1 \nabla^2 \eta(w_i^n - s \delta_{i-\frac{1}{2}}^n) (I_m - \Theta_i^n) \delta_{i-\frac{1}{2}}^n \cdot \delta_{i-\frac{1}{2}}^n ds.
\end{aligned}$$

En réécrivant l'égalité précédente à l'aide des matrices par blocs  $D_i^n$  et  $N_i^n(s)$  respectivement données dans (4.17a) et dans (4.17b) et en utilisant le résultat intermédiaire (4.24), il suit

$$\begin{aligned}
& \frac{1}{2} \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot (\lambda(\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n) - (\alpha_{i+1}^n - \alpha_{i-1}^n)) \\
& = -\frac{\lambda}{2} \int_0^1 (1-s) \sum_{i \in \mathbb{Z}} N_i^n(s) D_i^n \cdot D_i^n ds + \frac{\lambda}{4} \sum_{i \in \mathbb{Z}} \int_0^1 N_i^n(s) \begin{pmatrix} I_m - \Theta_i^n & \Theta_i^n \\ I_m - \Theta_i^n & \Theta_i^n \end{pmatrix} D_i^n \cdot D_i^n ds, \\
& = -\frac{\lambda}{4} \int_0^1 \sum_{i \in \mathbb{Z}} N_i^n(s) \left( \begin{pmatrix} (2-2s)I_m & 0 \\ 0 & (2-2s)I_m \end{pmatrix} - \begin{pmatrix} I_m - \Theta_i^n & \Theta_i^n \\ I_m - \Theta_i^n & \Theta_i^n \end{pmatrix} \right) D_i^n \cdot D_i^n ds.
\end{aligned}$$

L'introduction de la matrice par blocs  $P_i^n(s)$  définie par (4.17c) dans l'égalité ci-dessus permet de déduire la reformulation annoncée ce qui conclut la preuve.  $\square$

D'après les deux lemmes ci-dessus, l'inégalité de dissipation (4.4) du schéma du second ordre défini par le flux numérique (4.15) s'écrit

$$-\sum_{i \in \mathbb{Z}} \int_0^1 s \left( \int_0^1 N_i^n(us) du \right) F_i^n(s) D_i^n \cdot D_i^n ds + \frac{\lambda}{2} \sum_{i \in \mathbb{Z}} \int_0^1 N_i^n(s) P_i^n(s) D_i^n \cdot D_i^n ds > 0. \quad (4.26)$$

Par conséquent, et puisque toutes les sommes sont finies, si l'inégalité stricte

$$\sum_{i \in \mathbb{Z}} \int_0^1 N_i^n(s) P_i^n(s) D_i^n \cdot D_i^n ds > 0, \quad (4.27)$$

est satisfaite alors il existe  $\lambda > 0$  suffisamment grand tel que l'inégalité de dissipation (4.26) est satisfaite.

Par construction de la matrice par blocs  $P_i^n(s)$  définie par (4.17), le membre de gauche de l'inégalité (4.27) dépend des matrices  $\Theta_i^n$  qui sont formellement des paramètres libres. Ces paramètres libres doivent donc être choisis pour satisfaire l'inégalité stricte (4.27). L'existence d'une séquence de matrices  $(\Theta_i^n)_{i \in \mathbb{Z}}$  satisfaisant l'inégalité (4.27) est maintenant établie.

### 4.3.3 Existence de l'inégalité de dissipation

Cette section concerne l'existence d'une séquence de matrices  $(\Theta_i^n)_{i \in \mathbb{Z}}$  qui satisfait l'inégalité stricte (4.27). Ce résultat est donné dans le lemme ci-dessous.

**Lemme 4.3.3.1.** *Soit  $\eta$  une entropie strictement convexe de classe  $C^2(\Omega, \mathbb{R})$ . Considérons une séquence  $(w_i^n)_{i \in \mathbb{Z}}$  dans  $S_\eta^{O2}(\mathbb{Z}, \Omega)$  et supposons l'existence de  $K^n$  un ensemble compact convexe inclus dans  $\Omega$  tel que  $(w_i^n)_{i \in \mathbb{Z}} \subset K^n$ . En considérant les matrices par blocs  $D_i^n$ ,  $N_i^n(s)$  et  $P_i^n(s)$  définies dans (4.17), notons  $Q_i^n(s)$  et  $\theta$ , respectivement dans  $\mathcal{M}_{2m, 2m}(\mathbb{R})$  et dans  $\mathbb{R}$ , tels que*

$$Q_i^n(s) = \begin{pmatrix} (1-2s)\nabla^2 \eta(w_i^n - s \delta_{i-\frac{1}{2}}^n) & -\nabla^2 \eta(w_i^n + s \delta_{i+\frac{1}{2}}^n) \\ 0 & 2(1-s)\nabla^2 \eta(w_i^n + s \delta_{i+\frac{1}{2}}^n) \end{pmatrix}, \quad \forall s \in [0, 1], \quad (4.28)$$

$$\theta > \frac{-\min\left(0, \int_0^1 \sum_{i \in \mathbb{Z}} Q_i^n(s) D_i^n \cdot D_i^n ds\right)}{\sum_{\substack{i \in \mathbb{Z} \\ j \in \{1, \dots, m\}}} \left| (\nabla \eta(w_{i+1}^n) - \nabla \eta(w_{i-1}^n))_j (\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n)_j \right|}. \quad (4.29)$$

Si la séquence de matrices  $(\Theta_i^n)_{i \in \mathbb{Z}}$  est définie par

$$\Theta_i^n = -\theta \operatorname{diag}_{1 \leq j \leq m} \left( \operatorname{sign}((\nabla \eta(w_{i+1}^n) - \nabla \eta(w_{i-1}^n))_j (\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n)_j) \right), \quad (4.30)$$

alors l'inégalité stricte (4.27) est satisfaite.

*Démonstration.* Tout d'abord, et par définition de l'espace  $S_\eta(\Omega, \mathbb{Z})$  donnée dans (4.16), le dénominateur de (4.29) n'est pas nul. Il est maintenant nécessaire de montrer que le numérateur de (4.29) est fini.

En notant  $\|\cdot\|_2$  la norme euclidienne usuelle de  $\mathbb{R}^{2m}$ , l'inégalité de Cauchy-Schwarz permet d'écrire

$$\left| \int_0^1 \sum_{i \in \mathbb{Z}} Q_i^n(s) D_i^n \cdot D_i^n ds \right| \leq \sum_{i \in \mathbb{Z}} \left( \int_0^1 \|Q_i^n(s)\|_2 ds \right) \|D_i^n\|_2^2 \leq \sum_{i \in \mathbb{Z}} \sup_{s \in [0,1]} \|Q_i(s)\|_2 \|D_i^n\|_2^2. \quad (4.31)$$

En utilisant l'équivalence des normes dans les espaces vectoriels de dimension finie, il existe une constante  $c(m) > 0$  dépendante uniquement de la dimension  $m$  et telle que

$$\sup_{s \in [0,1]} \|Q_i(s)\|_2 \leq c(m) \sup_{s \in [0,1]} \|Q_i(s)\|_\infty.$$

Comme la séquence  $(w_i^n)_{i \in \mathbb{Z}}$  est incluse dans le compact convexe  $K^n$  et notant  $\|\cdot\|_\infty$  la norme matricielle dans  $\mathcal{M}_{m,m}(\mathbb{R})$ , il est possible d'utiliser le Lemme 4.3.1.1 qui donne une majoration uniforme de  $\sup_{s \in [0,1]} \|\nabla^2 \eta((1-s)w_i + sw_{i-1})\|_\infty$ . Cette majoration s'écrit pour tout  $i$  dans  $\mathbb{Z}$

$$\sup_{s \in [0,1]} \|\nabla^2 \eta((1-s)w_i + sw_{i-1})\|_\infty \leq \max_{w \in K^n} \|\nabla^2 \eta(w)\|_\infty.$$

Par conséquent, et d'après la définition de la matrice par blocs  $Q_i^n(s)$  donnée dans (4.28), on en déduit

$$\sup_{s \in [0,1]} \|Q_i(s)\|_\infty \leq 2 \max_{w \in K^n} \|\nabla^2 \eta(w)\|_\infty, \quad \forall i \in \mathbb{Z}.$$

En insérant l'inégalité ci-dessus dans l'estimation (4.31), et puisque  $(w_i^n)_{i \in \mathbb{Z}}$  est dans  $\dot{h}^2(\mathbb{Z}, \Omega)$ , il suit

$$\sum_{i \in \mathbb{Z}} \sup_{s \in [0,1]} \|Q_i(s)\|_2 \|D_i^n\|_2^2 \leq 2c(m) \max_{w \in K^n} \|\nabla^2 \eta(w)\|_\infty \sum_{i \in \mathbb{Z}} \|D_i^n\|_2^2 < +\infty.$$

Ainsi, le rapport des deux quantités qui définissent une borne inférieure de  $\theta$  dans (4.29) est nécessairement fini.

En ce qui concerne l'inégalité stricte (4.27), les matrices  $\Theta_i^n$  données par (4.30) sont symétriques et par conséquent, pour tout  $(a, b)$  dans  $(\mathbb{R}^m)^2$ , l'égalité  $\Theta_i^n a \cdot b = a \cdot \Theta_i^n b$  est vérifiée. Avec cette égalité, et en associant la définition de la matrice  $Q_i^n(s)$  donnée par (4.28) aux expressions des matrices  $D_i^n$ ,  $N_i^n(s)$  et  $P_i^n(s)$  énoncées dans (4.17), un calcul direct montre que

$$\begin{aligned} \sum_{i \in \mathbb{Z}} \int_0^1 N_i^n(s) P_i^n(s) D_i^n \cdot D_i^n ds = \\ \int_0^1 \sum_{i \in \mathbb{Z}} Q_i^n(s) D_i^n \cdot D_i^n ds - \sum_{i \in \mathbb{Z}} \Theta_i^n (\nabla \eta(w_{i+1}^n) - \nabla \eta(w_{i-1}^n)) \cdot (\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n). \end{aligned}$$

En utilisant la définition de  $\Theta_i^n$  donnée par (4.30), l'égalité ci-dessus se reformule de la manière suivante :

$$\sum_{i \in \mathbb{Z}} \int_0^1 N_i^n(s) P_i^n(s) D_i^n \cdot D_i^n ds = \int_0^1 \sum_{i \in \mathbb{Z}} Q_i^n(s) D_i^n \cdot D_i^n ds + \theta \sum_{\substack{i \in \mathbb{Z} \\ j \in \{1, \dots, m\}}} \left| ((\nabla \eta(w_{i+1}^n) - \nabla \eta(w_{i-1}^n))_j (\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n)_j) \right|.$$

Comme  $\theta$  vérifie l'inégalité (4.29), la quantité ci dessus est strictement positive, ce qui termine la preuve.  $\square$

La séquence de matrices  $(\Theta_i^n)_{i \in \mathbb{Z}}$  donnée par (4.30) permet de satisfaire l'inégalité stricte (4.27). Ainsi, le couplage de ces matrices à un choix de  $\lambda > 0$  suffisamment grand assure l'existence de l'inégalité de dissipation (4.26). Par conséquent, le schéma du second ordre (4.15) est dissipatif au sens de la Définition 4.1.1.

Outre ces aspects théoriques, l'inégalité (4.27) a également un intérêt pratique car elle est le critère de sélection des matrices  $(\Theta_i^n)_{i \in \mathbb{Z}}$ . Le lemme qui suit donne l'expression de cette inégalité pour le cas particulier mais usuel des équations de conservation scalaires munies de l'entropie quadratique.

**Lemme 4.3.3.2** (Cas des équations scalaires munies de l'entropie  $\eta(w) = w^2$ ). *Considérons une équation scalaire (4.1) ( $m = 1$ ) munie de l'entropie quadratique  $\eta(w) = w^2$ . Considérons de plus une séquence  $(w_i^n)_{i \in \mathbb{Z}}$  dans  $h^2(\mathbb{Z})$ , non constante et incluse dans  $\Omega$ . Dans ce cas,  $(\Theta_i^n)_{i \in \mathbb{Z}}$  est une séquence dans  $\mathbb{R}$  et l'inégalité stricte (4.27) s'écrit*

$$\sum_{i \in \mathbb{Z}} \left( (\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n)^2 - 2\Theta_i^n ((\delta_{i+\frac{1}{2}}^n)^2 - (\delta_{i-\frac{1}{2}}^n)^2) \right) > 0. \quad (4.32)$$

*Démonstration.* En utilisant le Lemme 4.3.2.2 dans le cas d'une équation scalaire et avec  $\eta(w) = w^2$ , le membre de gauche de l'inégalité (4.27) s'écrit

$$\sum_{i \in \mathbb{Z}} \int_0^1 N_i^n(s) P_i^n(s) D_i^n \cdot D_i^n ds = - \sum_{i \in \mathbb{Z}} 4w_i^n ((\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n) - \frac{\alpha_{i+1}^n - \alpha_{i-1}^n}{\lambda}).$$

En distribuant  $w_i^n$  puis en utilisant des translations de l'indice  $i$ , l'égalité ci-dessus se réécrit

$$\begin{aligned} \sum_{i \in \mathbb{Z}} \int_0^1 N_i^n(s) P_i^n(s) D_i^n \cdot D_i^n ds &= 4 \sum_{i \in \mathbb{Z}} (\delta_{i+\frac{1}{2}}^n)^2 + \sum_{i \in \mathbb{Z}} 4w_i^n \frac{\alpha_{i+1}^n - \alpha_{i-1}^n}{\lambda}, \\ &= 4 \sum_{i \in \mathbb{Z}} (\delta_{i+\frac{1}{2}}^n)^2 - \frac{4}{\lambda} \sum_{i \in \mathbb{Z}} (\delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n) \alpha_i^n. \end{aligned} \quad (4.33)$$

La définition de  $(\alpha_i^n)_{i \in \mathbb{Z}}$  donnée dans (4.14) et des translations de l'indice  $i$  permettent également d'écrire

$$\begin{aligned} -\frac{4}{\lambda} \sum_{i \in \mathbb{Z}} (\delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n) \alpha_i^n &= -2 \sum_{i \in \mathbb{Z}} (\delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n) (\Theta_i^n \delta_{i+\frac{1}{2}}^n + (1 - \Theta_i^n) \delta_{i-\frac{1}{2}}^n), \\ &= -2 \sum_{i \in \mathbb{Z}} \Theta_i^n (\delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n) (\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n) - 2 \sum_{i \in \mathbb{Z}} \delta_{i+\frac{1}{2}}^n \delta_{i-\frac{1}{2}}^n - 2 \sum_{i \in \mathbb{Z}} (\delta_{i+\frac{1}{2}}^n)^2. \end{aligned}$$

En utilisant le résultat ci-dessus dans l'égalité (4.33), il suit

$$\begin{aligned} \sum_{i \in \mathbb{Z}} \int_0^1 N_i^n(s) P_i^n(s) D_i^n \cdot D_i^n ds &= 2 \sum_{i \in \mathbb{Z}} (\delta_{i+\frac{1}{2}}^n)^2 - 2 \sum_{i \in \mathbb{Z}} \delta_{i+\frac{1}{2}}^n \delta_{i-\frac{1}{2}}^n - 2 \sum_{i \in \mathbb{Z}} \Theta_i^n (\delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n) (\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n), \\ &= \sum_{i \in \mathbb{Z}} ((\delta_{i+\frac{1}{2}}^n)^2 - 2\delta_{i+\frac{1}{2}}^n \delta_{i-\frac{1}{2}}^n + (\delta_{i-\frac{1}{2}}^n)^2) - 2 \sum_{i \in \mathbb{Z}} \Theta_i^n ((\delta_{i+\frac{1}{2}}^n)^2 - (\delta_{i-\frac{1}{2}}^n)^2), \\ &= \sum_{i \in \mathbb{Z}} (\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n)^2 - 2 \sum_{i \in \mathbb{Z}} \Theta_i^n ((\delta_{i+\frac{1}{2}}^n)^2 - (\delta_{i-\frac{1}{2}}^n)^2). \end{aligned}$$



Comme cette dernière expression correspond au résultat annoncé, ceci conclut la preuve.  $\square$

Dans le cas scalaire et d'après le lemme ci-dessus, plusieurs choix de  $\Theta_i^n$  sont possibles puisque  $\Theta_i^n = \text{cste}$  ou  $\Theta_i^n = -\text{sign}((w_{i+1}^n - w_{i-1}^n)(w_{i+1}^n - 2w_i^n + w_{i-1}^n))$  permettent de satisfaire l'inégalité stricte (4.32).

Avec l'existence de l'inégalité de dissipation (4.26) montrée par les Lemmes 4.3.3.1 et 4.3.3.2, la stabilité entropique globale des schémas du second ordre définis dans la Proposition 4.2.1 peut maintenant être présentée.

#### 4.3.4 Stabilité entropique discrète globale

Cette section établit la stabilité entropique globale du schéma du second ordre défini dans la Proposition 4.2.1. Cette stabilité est obtenue à l'aide du caractère dissipatif défini en Section 4.1. D'après la section ci-dessus, cette propriété de dissipation repose sur l'inégalité (4.27) qui est paramétrée par une séquence de matrices  $(\Theta_i^n)_{i \in \mathbb{Z}}$ . Les Lemmes 4.3.3.1 et 4.3.3.2 montrent que plusieurs choix de matrices  $(\Theta_i^n)_{i \in \mathbb{Z}}$  sont possibles et par conséquent, le théorème suivant concerne une famille de schémas du second ordre.

**Théorème 4.3.4.1** (Stabilité entropique discrète globale de schémas du second ordre). *Soit  $\eta$  une entropie strictement convexe de classe  $C^2(\Omega, \mathbb{R})$  associée au système (4.1). Considérons une séquence  $(w_i^n)_{i \in \mathbb{Z}}$  dans  $S_\eta^{O2}(\Omega, \mathbb{Z})$  telle que  $\sum_{i \in \mathbb{Z}} \eta(w_i^n) \Delta x$  est fini puis supposons l'existence de  $K^n$  un ensemble compact convexe inclus dans  $\Omega$  tel que  $(w_i^n)_{i \in \mathbb{Z}} \subset K^n$ . Considérons les matrices par blocs  $D_i^n$ ,  $N_i^n(s)$  et  $P_i^n(s)$  définies dans (4.17). Soit  $(\Theta_i^n)_{i \in \mathbb{Z}}$  une séquence de matrices diagonales de taille  $m \times m$ , uniformément bornées selon l'indice  $i$ , bornées lorsque  $\Delta x \rightarrow 0$  et satisfaisant l'inégalité stricte*

$$\sum_{i \in \mathbb{Z}} \int_0^1 N_i^n(s) P_i^n(s) D_i^n \cdot D_i^n ds > 0. \quad (4.34)$$

Pour  $\lambda > 0$ , considérons le schéma numérique du second ordre défini dans la Proposition 4.2.1 pour lequel le flux numérique s'écrit

$$\mathcal{F}_{i+\frac{1}{2}}^n = \frac{f(w_i^n) + f(w_{i+1}^n)}{2} - \frac{\lambda}{2} (w_{i+1}^n - w_i^n) + \frac{1}{2} (\alpha_i^n + \alpha_{i+1}^n), \quad (4.35a)$$

$$\alpha_i^n = \frac{\lambda}{2} (\Theta_i^n (w_{i+1}^n - w_i^n) + (I_m - \Theta_i^n) (w_i^n - w_{i-1}^n)), \quad \forall i \in \mathbb{Z}. \quad (4.35b)$$

Il existe  $\lambda > 0$  suffisamment grand et  $\Delta t > 0$  suffisamment petit tels que  $(w_i^{n+1})_{i \in \mathbb{Z}}$  est incluse dans  $\Omega$  et tels que

$$\sum_{i \in \mathbb{Z}} \frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} \Delta x \leq 0. \quad (4.36)$$

Le théorème ci-dessus énonce seulement l'existence de  $\lambda > 0$  et de  $\Delta t > 0$  tel que l'inégalité d'entropie discrète globale (4.36) est satisfaite mais une condition CFL plus précise est spécifiée dans (4.10) et une borne inférieure explicite du paramètre  $\lambda > 0$  est donnée dans la preuve qui suit.

*Démonstration du Théorème 4.3.4.1.* L'objectif est de montrer l'inégalité d'entropie discrète globale (4.36) en utilisant la Proposition 4.1.1. Par conséquent, il est d'abord nécessaire de montrer les sommabilités

$$\left| \sum_{i \in \mathbb{Z}} \eta(w_i^n) \Delta x \right| < +\infty, \quad (4.37a)$$

$$\left| \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot \frac{\mathcal{F}_{i+\frac{1}{2}}^n - \mathcal{F}_{i-\frac{1}{2}}^n}{\Delta x} \right| \Delta x < +\infty, \quad (4.37b)$$

$$\int_0^1 (1-s) \sum_{i \in \mathbb{Z}} \nabla^2 \eta \left( w_i^n - s \frac{\Delta t}{\Delta x} (\mathcal{F}_{i+\frac{1}{2}}^n - \mathcal{F}_{i-\frac{1}{2}}^n) \right) \frac{\mathcal{F}_{i+\frac{1}{2}}^n - \mathcal{F}_{i-\frac{1}{2}}^n}{\Delta x} \cdot \frac{\mathcal{F}_{i+\frac{1}{2}}^n - \mathcal{F}_{i-\frac{1}{2}}^n}{\Delta x} \Delta x ds < +\infty. \quad (4.37c)$$

La sommabilité (4.37a) est satisfaite par hypothèse. La sommabilité (4.37b) correspond à la sommabilité du membre de gauche de l'inégalité de dissipation (4.4). Les Lemmes 4.3.2.1 et 4.3.2.2 donnent une reformulation équivalente de cette quantité et les sommabilités établies dans le Lemme 4.3.1.2 permettent de déduire

$$\begin{aligned} & \left| \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot \frac{\mathcal{F}_{i+\frac{1}{2}}^n - \mathcal{F}_{i-\frac{1}{2}}^n}{\Delta x} \right| \Delta x \\ &= \left| - \sum_{i \in \mathbb{Z}} \int_0^1 s \left( \int_0^1 N_i^n(us) du \right) F_i^n(s) D_i^n \cdot D_i^n ds + \frac{\lambda}{2} \sum_{i \in \mathbb{Z}} \int_0^1 N_i^n(s) P_i^n(s) D_i^n \cdot D_i^n ds \right| \\ &\leq \left| \sum_{i \in \mathbb{Z}} \int_0^1 s \left( \int_0^1 N_i^n(us) du \right) F_i^n(s) D_i^n \cdot D_i^n ds \right| + \frac{\lambda}{2} \left| \sum_{i \in \mathbb{Z}} \int_0^1 N_i^n(s) P_i^n(s) D_i^n \cdot D_i^n ds \right| < +\infty. \end{aligned}$$

Il reste à montrer la sommabilité (4.37c) pour laquelle la notation  $\mathcal{R}_i^n = -(\mathcal{F}_{i+1/2}^n - \mathcal{F}_{i-1/2}^n)$  est adoptée. Puisque  $(w_i^n)_{i \in \mathbb{Z}}$  est incluse dans le compact  $K^n$ , le Lemme de robustesse 4.1.1 assure l'existence de  $\Delta t > 0$  et de  $R^n > 0$  tels qu'en notant  $\overline{B}(0, \frac{\Delta t}{\Delta x} R^n)$  la boule fermée contrée en 0, de rayon  $\frac{\Delta t}{\Delta x} R^n$  relativement à une norme  $\|\cdot\|$  de  $\mathbb{R}^m$ , l'inclusion suivante est vérifiée :

$$w_i^n + s \frac{\Delta t}{\Delta x} \mathcal{R}_i^n \in K^n + \overline{B}(0, \frac{\Delta t}{\Delta x} R^n) \subset \Omega, \quad \forall (s, i) \in [0, 1] \times \mathbb{Z}.$$

Puisque  $K^n$  et  $\overline{B}(0, \frac{\Delta t}{\Delta x} R^n)$  sont deux ensembles compacts, leur somme définie également un ensemble compact. Par conséquent, il est possible de redéployer les arguments avancés dans les Lemmes 4.3.1.2 et 4.3.3.1 pour déduire l'existence d'une constante  $c(m) > 0$  telle que

$$\left| \int_0^1 (1-s) \sum_{i \in \mathbb{Z}} \nabla^2 \eta \left( w_i^n + s \frac{\Delta t}{\Delta x} \mathcal{R}_i^n \right) \mathcal{R}_i^n \cdot \mathcal{R}_i^n ds \right| \leq c(m) \max_{w \in K^{n+1}} \|\nabla^2 \eta(w)\|_\infty \sum_{i \in \mathbb{Z}} \|\mathcal{R}_i^n\|_2^2.$$

D'après l'expression du flux numérique  $\mathcal{F}_{i+1/2}^n$  donnée par (4.35) et puisque les matrices  $(\Theta_i^n)_{i \in \mathbb{Z}}$  sont bornées et que  $(w_i^n)_{i \in \mathbb{Z}}$  est dans  $\dot{h}^2(\mathbb{Z}, \Omega)$ , la quantité  $\sum_{i \in \mathbb{Z}} \|\mathcal{R}_i^n\|_2^2$  est nécessairement finie. Par conséquent, la sommabilité (4.37c) est satisfaite.

Puisque les sommabilités (4.37) sont maintenant établies, il est nécessaire de montrer que le schéma numérique du second ordre défini par le flux numérique (4.35) est dissipatif. D'après les Lemmes 4.3.2.1 et 4.3.2.2, l'inégalité de dissipation (4.4) de ce schéma s'écrit

$$- \sum_{i \in \mathbb{Z}} \int_0^1 s \left( \int_0^1 N_i^n(us) du \right) F_i^n(s) D_i^n \cdot D_i^n ds + \frac{\lambda}{2} \sum_{i \in \mathbb{Z}} \int_0^1 N_i^n(s) P_i^n(s) D_i^n \cdot D_i^n ds > 0. \quad (4.38)$$

Comme  $(\Theta_i^n)_{i \in \mathbb{Z}}$  vérifie l'inégalité stricte (4.34), il est possible de définir un réel  $\lambda^n \geq 0$  de la manière suivante :

$$\lambda^n = \frac{2 \max \left( 0, \sum_{i \in \mathbb{Z}} \int_0^1 s \left( \int_0^1 N_i^n(us) du \right) F_i^n(s) D_i^n \cdot D_i^n ds \right)}{\sum_{i \in \mathbb{Z}} \int_0^1 N_i^n(s) P_i^n(s) D_i^n \cdot D_i^n ds} \geq 0. \quad (4.39)$$

D'après le Lemme 4.3.1.2, le rapport des quantités ci dessus est nécessairement fini. Par conséquent, et en imposant la contrainte  $\lambda > \lambda_n$ , l'inégalité de dissipation (4.38) est satisfaite. Ainsi, l'inégalité d'entropie discrète globale (4.36) se déduit de la Proposition 4.1.1, ce qui conclut la preuve.  $\square$

Le théorème ci-dessus montre l'existence d'un schéma du second ordre sans limiteur de pente et qui satisfait une inégalité d'entropie discrète globale de la forme (4.36).

La mise en œuvre pratique du Théorème 4.3.4.1 demeure simple et elle se résume en trois étapes. D'abord, il est nécessaire de choisir  $(\Theta_i^n)_{i \in \mathbb{Z}}$  selon l'inégalité (4.34) et les Lemmes 4.3.3.1 et 4.3.3.2

donnent des exemples possibles. Ensuite le paramètre  $\lambda > 0$  doit être sélectionné selon une borne inférieure explicite définie par (4.39) puis  $\Delta t > 0$  doit satisfaire la condition CFL établie dans la Proposition 4.1.1. En notant  $\mathcal{R}_i = -(\mathcal{F}_{i+\frac{1}{2}}^n - \mathcal{F}_{i-\frac{1}{2}}^n)$ , cette condition CFL s'écrit

$$0 < \frac{\Delta t}{\Delta x} \leq \frac{-\sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot \mathcal{R}_i}{\int_0^1 (1-s) \sum_{i \in \mathbb{Z}} \nabla^2 \eta \left( w_i^n + s \frac{\Delta t}{\Delta x} \mathcal{R}_i \right) \mathcal{R}_i \cdot \mathcal{R}_i ds}. \quad (4.40)$$

La condition CFL ci-dessus n'est sans doute pas optimale et son utilisation pratique nécessite, en général, la résolution d'une équation non linéaire. La construction de cette équation demande l'évaluation du flux numérique, du gradient et de la matrice hessienne de  $\eta$  pour chaque élément de la séquence  $(w_i^n)_{i \in \mathbb{Z}}$ . Par conséquent, l'obtention du paramètre  $\Delta t > 0$  à partir de l'inégalité (4.40) implique un temps de calcul supplémentaire non négligeable.

Par simplicité, le résultat du Théorème 4.3.4.1 est présenté avec une discrétisation temporelle du premier ordre mais il est d'usage d'associer des discrétisations temps et espace du même ordre [16, 125, 194, 230]. La section suivante justifie cette montée en ordre temporel par l'étude de la condition CFL (4.40) pour le cas de l'équation d'advection linéaire.

### 4.3.5 Étude de la condition CFL pour l'advection linéaire

L'objectif de cette section est de justifier la discrétisation en temps du second ordre pour le schéma numérique défini dans la Proposition 4.2.1. D'après le Théorème 4.3.4.1, la stabilité entropique globale de ce schéma repose en partie sur la condition CFL non standard (4.40). La justification de la montée en ordre temporel repose sur l'étude de cette condition CFL dans le cas de l'équation d'advection linéaire scalaire. Pour  $a$  dans  $\mathbb{R}_*$ , ce problème s'écrit

$$\begin{cases} \partial_t w + a \partial_x w = 0, & t > 0, \quad x \in \mathbb{R}, \\ w(x, t = 0) = w_0(x). \end{cases} \quad (4.41)$$

En considérant  $(\alpha_i^n)_{i \in \mathbb{Z}}$  définie dans (4.14) puis en notant  $\nu = \frac{\lambda \Delta t}{\Delta x} > 0$  et  $\delta_{i+\frac{1}{2}}^n = w_{i+1}^n - w_i^n$  dans  $\mathbb{R}$ , le schéma du second ordre en espace et d'ordre un en temps de la Proposition 4.2.1 s'écrit dans ce cas

$$w_i^{n+1} = w_i^n - \frac{\nu}{2} \left( \frac{a}{\lambda} (\delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n) - \delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n + \frac{1}{\lambda} (\alpha_{i+1}^n - \alpha_{i-1}^n) \right). \quad (4.42)$$

Puisque le problème de l'advection linéaire (4.41) est usuellement associé à l'entropie quadratique  $\eta(w) = w^2$ , il est maintenant possible de donner une formulation de la condition CFL (4.40) dans le cadre de la stabilité  $L^2$  su schéma (4.42).

**Lemme 4.3.5.1** (Stabilité quadratique sous condition CFL parabolique). *Considérons une séquence non nul  $(w_i^n)_{i \in \mathbb{Z}}$  dans  $h^2(\mathbb{Z})$  ainsi qu'une séquence de réels  $(\Theta_i^n)_{i \in \mathbb{Z}}$  telle que l'inégalité stricte (4.32) est satisfaite. Soit  $\lambda > 0$  un réel donné et notons  $\nu = \frac{\lambda \Delta t}{\Delta x}$ . Si  $\nu$  vérifie*

$$0 < \nu \leq \frac{\sum_{i \in \mathbb{Z}} \left( (\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n)^2 - 2\Theta_i^n ((\delta_{i+\frac{1}{2}}^n)^2 - (\delta_{i-\frac{1}{2}}^n)^2) \right)}{\sum_{i \in \mathbb{Z}} \left( \frac{a}{\lambda} (\delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n) - \delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n + \frac{1}{\lambda} (\alpha_{i+1}^n - \alpha_{i-1}^n) \right)^2}, \quad (4.43)$$

alors le schéma du premier ordre en temps et du second ordre en espace (4.42) est  $L^2$ -stable ce qui s'écrit

$$\sum_{i \in \mathbb{Z}} |w_i^{n+1}|^2 \Delta x \leq \sum_{i \in \mathbb{Z}} |w_i^n|^2 \Delta x.$$

*Démonstration.* En élevant le schéma (4.42) au carré puis en sommant sur les indices  $i$ , il est possible d'écrire

$$\begin{aligned} \sum_{i \in \mathbb{Z}} (w_i^{n+1})^2 \Delta x &= \sum_{i \in \mathbb{Z}} (w_i^n)^2 \Delta x \\ &- \nu \sum_{i \in \mathbb{Z}} w_i^n \left( \frac{a}{\lambda} (\delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n) - \delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n + \frac{1}{\lambda} (\alpha_{i+1}^n - \alpha_{i-1}^n) \right) \Delta x \\ &+ \frac{\nu^2}{4} \sum_{i \in \mathbb{Z}} \left( \frac{a}{\lambda} (\delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n) - \delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n + \frac{1}{\lambda} (\alpha_{i+1}^n - \alpha_{i-1}^n) \right)^2 \Delta x. \end{aligned} \quad (4.44)$$

En développant le second terme de l'équation ci-dessus et en utilisant des translations de l'indice  $i$ , il suit

$$\begin{aligned} &\sum_{i \in \mathbb{Z}} w_i^n \left( \frac{a}{\lambda} (\delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n) - \delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n + \frac{1}{\lambda} (\alpha_{i+1}^n - \alpha_{i-1}^n) \right) \\ &= \frac{a}{\lambda} \sum_{i \in \mathbb{Z}} w_i^n (w_{i+1}^n - w_{i-1}^n) - \sum_{i \in \mathbb{Z}} w_i^n \delta_{i+\frac{1}{2}}^n + \sum_{i \in \mathbb{Z}} w_i^n \delta_{i-\frac{1}{2}}^n - \frac{1}{\lambda} \sum_{i \in \mathbb{Z}} \alpha_i^n (\delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n). \end{aligned}$$

En utilisant l'égalité  $2\alpha_i^n/\lambda = \Theta_i^n \delta_{i+\frac{1}{2}}^n + (1 - \Theta_i^n) \delta_{i-\frac{1}{2}}^n$  ainsi qu'une nouvelle translation de l'indice  $i$ , l'égalité ci-dessus se réécrit

$$\begin{aligned} &\sum_{i \in \mathbb{Z}} w_i^n \left( \frac{a}{\lambda} (\delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n) - \delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n + \frac{1}{\lambda} (\alpha_{i+1}^n - \alpha_{i-1}^n) \right) \\ &= \frac{1}{2} \sum_{i \in \mathbb{Z}} (\delta_{i+\frac{1}{2}}^n)^2 - \frac{1}{2} \sum_{i \in \mathbb{Z}} \Theta_i^n ((\delta_{i+\frac{1}{2}}^n)^2 - (\delta_{i-\frac{1}{2}}^n)^2) - \frac{1}{2} \sum_{i \in \mathbb{Z}} \delta_{i+\frac{1}{2}}^n \delta_{i-\frac{1}{2}}^n, \\ &= \frac{1}{4} \sum_{i \in \mathbb{Z}} (\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n)^2 - \frac{1}{2} \sum_{i \in \mathbb{Z}} \Theta_i^n ((\delta_{i+\frac{1}{2}}^n)^2 - (\delta_{i-\frac{1}{2}}^n)^2). \end{aligned}$$

L'introduction de l'expression ci-dessus dans l'égalité (4.44) permet de déduire

$$\begin{aligned} \sum_{i \in \mathbb{Z}} (w_i^{n+1})^2 \Delta x &= \sum_{i \in \mathbb{Z}} (w_i^n)^2 \Delta x \\ &- \frac{\nu}{4} \sum_{i \in \mathbb{Z}} \left( (\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n)^2 - 2\Theta_i^n ((\delta_{i+\frac{1}{2}}^n)^2 - (\delta_{i-\frac{1}{2}}^n)^2) \right) \Delta x \\ &+ \frac{\nu^2}{4} \sum_{i \in \mathbb{Z}} \left( \frac{a}{\lambda} (\delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n) - \delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n + \frac{1}{\lambda} (\alpha_{i+1}^n - \alpha_{i-1}^n) \right)^2 \Delta x. \end{aligned}$$

Comme la séquence  $(\Theta_i^n)_{i \in \mathbb{Z}}$  assure que la somme multiplicative de  $\nu$  est strictement positive et comme  $\nu > 0$  vérifie la condition CFL (4.43), on en déduit le résultat de stabilité  $L^2$  annoncé ce qui termine la preuve.  $\square$

Dans le cas de solutions régulières, une analyse de consistance montre que le comportement asymptotique de la condition CFL (4.43) lorsque  $\Delta x \rightarrow 0$  est de la forme  $\frac{\Delta t}{\Delta x^2} = \mathcal{O}(1)$ . Par conséquent, le schéma du premier ordre en temps et du second ordre en espace (4.42) est  $L^2$ -stable sous une condition CFL *parabolique*. De telles conditions CFL ont déjà été observées dans la littérature [73, 74, 230] mais elles demeurent restrictives en pratique.

Pour les schémas du second ordre spatial proposés dans ce chapitre, il est possible de retrouver une condition CFL *hyperbolique* de la forme  $\frac{\Delta t}{\Delta x} = \mathcal{O}(1)$  en utilisant une discrétisation en temps du second ordre de type *Runge-Kutta Strong Stability Preserving* (RKSSP) [107, 109, 110]. Dans le cas du problème de l'advection linéaire (4.41), l'application d'une telle discrétisation temporelle au schéma de la Proposition 4.2.1 donne

$$\begin{aligned} w_i^{n+\frac{1}{2}} &= w_i^n - \frac{\nu}{2} \left( \frac{a}{\lambda} (\delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n) - \delta_{i+\frac{1}{2}}^n + \delta_{i-\frac{1}{2}}^n + \frac{1}{\lambda} (\alpha_{i+1}^n + \alpha_{i-1}^n) \right), \\ w_i^{n+1} &= \frac{w_i^n}{2} + \frac{1}{2} \left( w_i^{n+\frac{1}{2}} - \frac{\nu}{2} \left( \frac{a}{\lambda} (\delta_{i+\frac{1}{2}}^{n+\frac{1}{2}} + \delta_{i-\frac{1}{2}}^{n+\frac{1}{2}}) - \delta_{i+\frac{1}{2}}^{n+\frac{1}{2}} + \delta_{i-\frac{1}{2}}^{n+\frac{1}{2}} + \frac{1}{\lambda} (\alpha_{i+1}^{n+\frac{1}{2}} + \alpha_{i-1}^{n+\frac{1}{2}}) \right) \right). \end{aligned} \quad (4.45)$$

En choisissant des séquences  $(\Theta_i^r)_{(i,r) \in \mathbb{Z} \times \{n, n+1/2\}}$  constantes, l'analyse de Fourier-Von Neumann [53, 74] est maintenant utilisée pour déduire une condition CFL de stabilité  $L^2$  du schéma (4.45).

**Lemme 4.3.5.2** (Stabilité quadratique sous condition CFL hyperbolique). *Considérons une séquence non nul  $(w_i^n)_{i \in \mathbb{Z}}$  dans  $l^2(\mathbb{Z})$  ainsi que  $\Theta$  dans  $\mathbb{R}$  puis deux séquences de réels  $(\Theta_i^n)_{i \in \mathbb{Z}}$  et  $(\Theta_i^{n+1/2})_{i \in \mathbb{Z}}$  telles que*

$$\Theta_i^n = \Theta_i^{n+1/2} = \Theta, \quad \forall i \in \mathbb{Z}.$$

Soit  $\lambda > 0$  un réel donné et notons  $\nu = \frac{\lambda \Delta t}{\Delta x}$ . Si  $\nu$  vérifie

$$0 < \nu \max(1, (a/\lambda)^2, (1 - 2\Theta + a/\lambda)^2) \leq 1, \quad (4.46)$$

alors le schéma numérique du second ordre en temps et en espace (4.45) est  $L^2$ -stable ce qui s'écrit

$$\sum_{i \in \mathbb{Z}} |w_i^{n+1}|^2 \Delta x \leq \sum_{i \in \mathbb{Z}} |w_i^n|^2 \Delta x. \quad (4.47)$$

*Démonstration.* Pour  $\xi$  dans  $\mathbb{R}$ , notons  $\hat{w}_\Delta(\xi, t^n)$  la transformée de Fourier de la fonction  $w_\Delta(x, t^n)$  dont l'expression est donnée par

$$w_\Delta(x, t^n) = \sum_{j \in \mathbb{Z}} w_j^n \chi_{C_j}(x).$$

En notant  $i^2 = -1$ , un calcul usuel montre que la transformée de Fourier de la fonction ci-dessus s'écrit

$$\hat{w}_\Delta(\xi, t^n) = \frac{\sin(\pi \xi \Delta x)}{\pi \xi} \sum_{j \in \mathbb{Z}} w_j^n e^{-2\pi i \xi j \Delta x}.$$

Comme les séquences  $(\Theta_i^r)_{(i,r) \in \mathbb{Z} \times \{n, n+1/2\}}$  sont constantes, le schéma numérique (4.45) est linéaire. Par conséquent, il existe une fonction  $\xi \mapsto g(\xi \Delta x)$  telle que  $\hat{w}_\Delta(\xi, t^{n+1}) = g(\xi \Delta x) \hat{w}_\Delta(\xi, t^n)$ . Par isométrie de la transformée de Fourier, la stabilité  $L^2$  (4.47) est équivalente à l'inégalité

$$|g(\xi \Delta x)| \leq 1, \quad \forall \xi \in \mathbb{R}. \quad (4.48)$$

Ainsi, l'objectif de la preuve est de montrer l'inégalité ci-dessus. Dans le cas  $\Theta_i^r = \Theta$  et en notant  $\phi_\xi = \sin^2(\pi \xi \Delta x)$  et  $h_\Theta = (1 - 2\Theta) \sin(\pi \xi \Delta x) + a/\lambda$ , un calcul direct montre que la fonction  $g(\xi \Delta x)$  pour le schéma (4.45) vérifie

$$\begin{aligned} |g(\xi \Delta x)|^2 &= (1 - 2\nu \phi_\xi (\phi_\xi - \nu \phi_\xi^3 + \nu(1 - \phi_\xi) h_\Theta^2))^2 + 4\phi_\xi (1 - \phi_\xi) (\nu h_\Theta (1 - 2\nu \phi_\xi^2))^2, \\ &= 1 - 4\nu \phi_\xi (\phi_\xi - \nu \phi_\xi^3 + \nu(1 - \phi_\xi) h_\Theta^2) + 4\nu^2 \phi_\xi^2 (\phi_\xi - \nu \phi_\xi^3 + \nu(1 - \phi_\xi) h_\Theta^2)^2 \\ &\quad + 4\phi_\xi (1 - \phi_\xi) (\nu h_\Theta (1 - 2\nu \phi_\xi^2))^2, \\ &= 1 - 4\nu \phi_\xi^2 (1 - \nu \phi_\xi^2) + 4\nu^2 \phi_\xi (1 - \phi_\xi) h_\Theta^2 (-1 + (1 - 2\nu \phi_\xi^2)^2) \\ &\quad + 4\nu^2 \phi_\xi^2 (\phi_\xi - \nu \phi_\xi^3 + \nu(1 - \phi_\xi) h_\Theta^2)^2. \end{aligned}$$

En utilisant l'égalité  $-1 + (1 - 2\nu \phi_\xi^2)^2 = -4\nu \phi_\xi^2 (1 - \nu \phi_\xi^2)$ , l'expression ci-dessus se réécrit

$$|g(\xi \Delta x)|^2 = 1 - 4\nu \phi_\xi^2 (1 - \nu \phi_\xi^2) (1 + 4\nu^2 \phi_\xi (1 - \phi_\xi) h_\Theta^2) + 4\nu^2 \phi_\xi^2 (\phi_\xi - \nu \phi_\xi^3 + \nu(1 - \phi_\xi) h_\Theta^2)^2. \quad (4.49)$$

On peut tout d'abord remarquer que la restriction de  $|g(\xi \Delta x)|$  au cas  $\phi_\xi = 1$  s'écrit

$$\begin{aligned} |g(\xi \Delta x)|^2 \Big|_{\phi_\xi=1} &= 1 - 4\nu(1 - \nu) + 4\nu^2(1 - \nu)^2, \\ &= (1 - 2\nu(1 - \nu))^2. \end{aligned}$$

En conséquence, et pour ce cas particulier, l'inégalité (4.48) est satisfaite si  $\nu \leq 1$ . Ainsi, la condition nécessaire  $\nu \leq 1$  est désormais imposée.

Ensuite, et puisque  $0 \leq \phi_\xi \leq 1$  pour tout  $\xi$  dans  $\mathbb{R}$ , les deux inégalités suivantes sont satisfaites :

$$0 \leq 1 - \phi_\xi \leq 1 - \nu\phi_\xi^2, \quad (1 - \nu\phi_\xi^2)^2 \leq 1 - \nu\phi_\xi^2.$$

Avec les deux inégalités ci-dessus et comme  $\nu^2 \leq \nu$ , il est possible d'écrire

$$\begin{aligned} (\phi_\xi - \nu\phi_\xi^3 + \nu(1 - \phi_\xi)h_\Theta^2)^2 &= (\phi_\xi(1 - \nu\phi_\xi^2) + \nu(1 - \phi_\xi)h_\Theta^2)^2, \\ &= (1 - \nu\phi_\xi^2)^2\phi_\xi^2 + 2\nu\phi_\xi(1 - \nu\phi_\xi^2)(1 - \phi_\xi)h_\Theta^2 \\ &\quad + \nu^2(1 - \phi_\xi)^2h_\Theta^4, \\ &\leq (1 - \nu\phi_\xi^2)(\phi_\xi^2 + 2\nu\phi_\xi(1 - \phi_\xi)h_\Theta^2 + \nu^2(1 - \nu\phi_\xi^2)h_\Theta^4), \\ &\leq (1 - \nu\phi_\xi^2)(\phi_\xi^2 + 2\nu\phi_\xi(1 - \phi_\xi)h_\Theta^2 + \nu(1 - \nu\phi_\xi^2)h_\Theta^4). \end{aligned}$$

L'utilisation de la majoration ci-dessus dans l'égalité (4.49) donne

$$\begin{aligned} |g(\xi\Delta x)|^2 &\leq 1 - 4\nu\phi_\xi^2(1 - \nu\phi_\xi^2)h_\Theta^2(1 + 2\nu^2\phi_\xi(1 - \phi_\xi)h_\Theta^2 - \nu\phi_\xi^2 - \nu^2(1 - \nu\phi_\xi^2)h_\Theta^4), \\ &\leq 1 - 2\nu^3\phi_\xi^2\phi_{2\xi}(1 - \nu\phi_\xi^2)h_\Theta^2 - 4\nu\phi_\xi^2(1 - \nu\phi_\xi^2)^2(1 - \nu^2h_\Theta^4). \end{aligned} \quad (4.50)$$

Par ailleurs, la définition de  $h_\Theta = (1 - 2\Theta)\sin(\pi\xi\Delta x) + a/\lambda$  ainsi que la condition CFL (4.46) permettent de déduire

$$0 \leq 2\nu^3\phi_\xi^2\phi_{2\xi}(1 - \nu\phi_\xi^2)h_\Theta^2 + 4\nu\phi_\xi^2(1 - \nu\phi_\xi^2)^2(1 - \nu^2h_\Theta^4) \leq 1. \quad (4.51)$$

En utilisant les inégalités ci-dessus dans la majoration (4.50), on en déduit  $|g(\xi\Delta x)|^2 \leq 1$  ce qui conclut la preuve.  $\square$

Dans le cas de l'advection linéaire (4.41) et d'après le lemme ci-dessus, l'utilisation d'une discrétisation temporelle du second ordre de type RKSSP permet de modifier la condition CFL parabolique (4.43) en une condition CFL hyperbolique donnée par (4.46). Cet apport de la discrétisation temporelle est également observé numériquement dans le cas de fonctions de flux  $f$  non linéaire.

La dissipation des schémas du second ordre de la Proposition 4.2.1 est pilotée par la séquence  $(\alpha_i^n)$  à l'aide d'un choix de matrices  $(\Theta_i^n)_{i \in \mathbb{Z}}$ . Par conséquent, le rôle de  $(\alpha_i^n)_{i \in \mathbb{Z}}$  est double. D'après la section 4.2, cette séquence a d'abord un rôle de *correcteur* qui assure le second ordre spatial. Ensuite, d'après la section 4.3.3,  $(\alpha_i^n)_{i \in \mathbb{Z}}$  possède également un rôle de *dissipateur* qui conduit à la stabilité entropique discrète globale. En généralisant ces deux propriétés, il est possible de définir des schémas numériques d'ordre  $k$  en espace construits à partir de corrections dissipatives du même ordre.

## 4.4 Montée en ordre supérieur

### 4.4.1 Corrections dissipatives d'ordre $k$

Cette section propose une généralisation à l'ordre  $k \geq 2$  des schémas du second ordre présentés en Section 4.2. La première notion nécessaire à cette généralisation est celle de *correction dissipative d'ordre  $k$* .

**Définition 4.4.1** (Correction dissipative d'ordre  $k$ ). *Soient  $k$  un entier et  $\lambda > 0$  un réel. Considérons une séquence  $(w_i^n)_{i \in \mathbb{Z}}$  dans  $\Omega$  ainsi que  $\eta$  une entropie strictement convexe de classe  $C^2(\Omega, \mathbb{R})$ . Une séquence  $(\alpha_i^{Ok})_{i \in \mathbb{Z}}$  dans  $\mathbb{R}^m$  est appelée *correction dissipative d'ordre  $k$  (relativement à  $\eta$ )* si les deux propriétés suivantes sont satisfaites.*

- i) La moyenne arithmétique de  $\alpha_i^{Ok}$  et de  $\alpha_{i+1}^{Ok}$  corrige le terme  $-\lambda(w_{i+1}^n - w_i^n)/2$  à l'ordre  $k$ . Pour  $w$  régulière, cette correction d'ordre  $k$  s'écrit

$$-\frac{\lambda}{2}(w_{i+1}^n - w_i^n) + \frac{\alpha_{i+1}^{Ok} + \alpha_i^{Ok}}{2} = \mathcal{O}(\Delta x^k). \quad (4.52)$$

ii) La séquence  $(\alpha_i^{Ok})_{i \in \mathbb{Z}}$  vérifie une inégalité stricte globale donnée par

$$\int_0^1 \sum_{i \in \mathbb{Z}} \nabla^2 \eta(w_i^n + s(w_{i+1}^n - w_i^n))(w_{i+1}^n - w_i^n) \cdot \frac{\alpha_i^{Ok} + \alpha_{i+1}^{Ok}}{2} ds < 0. \quad (4.53)$$

L'inégalité ci dessus est appelée inégalité de dissipation associée à  $(\alpha_i^{Ok})_{i \in \mathbb{Z}}$ .

D'après les sections 4.2 et 4.3.3, une séquence de correction dissipative du second ordre  $(\alpha_i^{O2})_{i \in \mathbb{Z}}$  est donnée par (4.14) sous la contrainte (4.27) détaillée dans le Lemme 4.3.3.1.

L'inégalité de dissipation associée à  $(\alpha_i^{Ok})_{i \in \mathbb{Z}}$  donnée dans la Définition 4.4.1-ii) est une quantité globale pour laquelle il est nécessaire de montrer la convergence. Si la section 4.3 consacrée au schéma du second ordre utilise des espaces de Sobolev discrets homogènes pour établir ces sommabilités, cette section propose d'utiliser la notion de support compact étendu.

**Définition 4.4.2** (Support compact étendu dans un ensemble). *Soit  $\mathcal{A}$  un ensemble quelconque et  $(v_i)_{i \in \mathbb{Z}}$  une séquence à valeurs dans  $\mathcal{A}$ . Pour  $a$  dans  $\mathcal{A}$  quelconque, la séquence  $(v_i)_{i \in \mathbb{Z}}$  est dite à support compact étendu dans  $\mathcal{A}$  s'il existe un entier  $i^*$  tel que*

$$\forall i \in \mathbb{Z}, \quad \text{si } |i| \geq i^* \quad \text{alors } v_i = a. \quad (4.54)$$

La notion de support compact étendu donnée dans la définition ci-dessus a seulement un intérêt pratique dans ce mémoire de thèse. Tout d'abord, puisqu'elle prolonge formellement  $(w_i^n)_{i \in \mathbb{Z}}$  à des constantes de  $\Omega$ , cette notion assure que la séquence des différences  $(w_{i+1}^n - w_i^n)_{i \in \mathbb{Z}}$  est à support compact. Ensuite, et puisque les entropies  $\eta$  sont définies à une constante additive près, il est toujours possible de choisir  $a$  dans la définition (4.54) tel que  $(\eta(w_i^n))_{i \in \mathbb{Z}}$  est également à support compact. Ainsi, la Définition 4.4.2 entraîne la convergence de la plupart des sommes exposées dans la suite de ce chapitre et par conséquent, sauf si cela est nécessaire, la sommabilité des quantités proposées ne sera pas mentionnée.

Avec la définition de support compact étendu de la Définition 4.4.2, et en supposant l'existence d'une correction dissipative d'ordre  $k$  quelconque, il est possible de proposer une structure de flux numérique qui définit des schémas d'ordre  $k$  à stabilité entropique globale discrète.

**Proposition 4.4.1** (Stabilité entropique globale des schémas numériques d'ordre  $k$ ). *Soit  $K^n$  un ensemble compact de  $\Omega$ . Considérons une séquence  $(w_i^n)_{i \in \mathbb{Z}}$  non constante, à valeurs et à support compact étendu dans  $K^n$ . Considérons ensuite  $(\alpha_i^{Ok})_{i \in \mathbb{Z}}$  une séquence de correction dissipative d'ordre  $k$  et  $\mathfrak{F}_{i+1/2}^{Ok}$  un flux numérique consistant d'ordre  $k$  satisfaisants respectivement les Définitions 4.4.1 et 1.4.1.1. Pour un réel  $\lambda > 0$ , soit un schéma conservatif (4.3) défini par un flux numérique  $\mathcal{F}_{i+1/2}^{Ok}$  qui s'écrit*

$$\mathcal{F}_{i+1/2}^{Ok} = \mathfrak{F}_{i+1/2}^{Ok} - \frac{\lambda}{2}(w_{i+1}^n - w_i^n) + \frac{\alpha_i^{Ok} + \alpha_{i+1}^{Ok}}{2}. \quad (4.55)$$

Les assertions suivantes sont vérifiées.

i) Le schéma associé au flux numérique ci-dessus est d'ordre  $k$  en espace au sens de la Définition formelle 1.4.1.1-iii).

ii) Si

$$\left| \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot \frac{\mathfrak{F}_{i+1/2}^{Ok} - \mathfrak{F}_{i-1/2}^{Ok}}{\Delta x} \right| < +\infty, \quad \forall \lambda > 0, \quad \forall (w_i^n)_{i \in \mathbb{Z}} \subset \Omega, \quad (4.56)$$

et si l'entropie  $\eta$  associée au système (4.1) est une fonction de classe  $C^2(\Omega, \mathbb{R})$  strictement convexe alors il existe  $\lambda > 0$  suffisamment grand et  $\Delta t > 0$  suffisamment petit tels que le schéma numérique défini par (4.55) vérifie une stabilité entropique discrète globale qui s'écrit

$$\sum_{i \in \mathbb{Z}} \frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} \Delta x \leq 0. \quad (4.57)$$

*Démonstration.* Concernant l'assertion *i*), soit  $u : \mathbb{R} \rightarrow \mathbb{R}^m$  une fonction régulière et notons

$$u_i = \frac{1}{\Delta x} \int_{C_i} u(x) dx, \quad \forall i \in \mathbb{Z}.$$

Puisque le flux numérique  $\mathfrak{F}^{Ok}$  est formellement d'ordre  $k$ , la Définition 1.4.1.1-*iii*) assure l'existence d'un entier  $l$  tel que

$$\mathfrak{F}_{i+\frac{1}{2}}^{Ok} = \mathfrak{F}^{Ok}(u_{i-l}, \dots, u_{i+l+1}) = f(u(x_{i+\frac{1}{2}})) + \mathcal{O}(\Delta x^k).$$

Avec l'estimation ci-dessus et en vertu de la correction (4.52) de la séquence  $(\alpha_i^{Ok})_{i \in \mathbb{Z}}$ , le flux numérique  $\mathcal{F}_{i+\frac{1}{2}}^{Ok}$  vérifie

$$\begin{aligned} \mathcal{F}_{i+\frac{1}{2}}^{Ok} &= f(u(x_{i+\frac{1}{2}})) + \mathcal{O}(\Delta x^k) - \frac{\lambda}{2}(u_{i+1} - u_i) + \frac{\alpha_i^{Ok} + \alpha_{i+1}^{Ok}}{2}, \\ &= f(u(x_{i+\frac{1}{2}})) + \mathcal{O}(\Delta x^k). \end{aligned} \quad (4.58)$$

Puisque cette dernière estimation est la définition de l'ordre spatial donnée en Section 1.4.1, ceci montre la première assertion.

La preuve de l'assertion *ii*) repose sur la Proposition 4.1.1. Ainsi, et puisque les hypothèses de sommabilités (4.7) sont satisfaites, il est suffisant de montrer que le schéma numérique défini par (4.55) satisfait l'inégalité de dissipation relativement à  $\eta$  donnée par (4.4).

En notant  $\delta_{i+\frac{1}{2}}^n = w_{i+1}^n - w_i^n$  dans  $\mathbb{R}^m$ , le membre de gauche de l'inégalité de dissipation (4.4) pour les schémas (4.55) s'écrit

$$\begin{aligned} &\sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot (\mathcal{F}_{i+\frac{1}{2}}^{Ok} - \mathcal{F}_{i-\frac{1}{2}}^{Ok}) \\ &= \frac{1}{2} \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot (\mathfrak{F}_{i+\frac{1}{2}}^{Ok} - \mathfrak{F}_{i-\frac{1}{2}}^{Ok}) - \frac{\lambda}{2} \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot (\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n) + \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot \frac{\alpha_{i+1}^{Ok} - \alpha_{i-1}^{Ok}}{2}. \end{aligned} \quad (4.59)$$

L'objectif est de montrer que la quantité ci-dessus peut être strictement positive. Il est donc maintenant nécessaire de récrire les deux dernières sommes qui figurent dans le membre de droite de (4.59). En utilisant les définitions des matrices par blocs  $D_i^n$  et  $N_i^n(s)$  données par (4.17) ainsi que l'égalité (4.24) établie dans la preuve du Lemme 4.3.2.2), la seconde somme du membre de droite de l'égalité (4.59) se reformule de la manière suivante :

$$-\frac{\lambda}{2} \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot (\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n) = \frac{\lambda}{2} \int_0^1 (1-s) \sum_{i \in \mathbb{Z}} N_i^n(s) D_i^n \cdot D_i^n ds. \quad (4.60)$$

En utilisant une translation de l'indice  $i$  et un développement de Taylor avec reste intégral, la troisième somme du membre de droite de l'égalité (4.59) devient

$$\begin{aligned} \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot \frac{\alpha_{i+1}^{Ok} - \alpha_{i-1}^{Ok}}{2} &= - \sum_{i \in \mathbb{Z}} (\nabla \eta(w_{i+1}^n) - \nabla \eta(w_i^n)) \cdot \frac{\alpha_i^{Ok} + \alpha_{i+1}^{Ok}}{2}, \\ &= - \int_0^1 \sum_{i \in \mathbb{Z}} \nabla^2 \eta(w_i^n + s \delta_{i+\frac{1}{2}}^n) \delta_{i+\frac{1}{2}}^n \cdot \frac{\alpha_i^{Ok} + \alpha_{i+1}^{Ok}}{2} ds. \end{aligned} \quad (4.61)$$

D'après la propriété de dissipation (4.53), la quantité ci-dessus est strictement positive et par consé-



quent, l'introduction des reformulations (4.60)-(4.61) dans l'égalité (4.59) donne

$$\begin{aligned}
& \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot (\mathcal{F}_{i+\frac{1}{2}}^{Ok} - \mathcal{F}_{i-\frac{1}{2}}^{Ok}) \\
&= \frac{1}{2} \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot (\mathfrak{F}_{i+\frac{1}{2}}^{Ok} - \mathfrak{F}_{i-\frac{1}{2}}^{Ok}) + \frac{\lambda}{2} \int_0^1 (1-s) \sum_{i \in \mathbb{Z}} N_i^n(s) D_i^n \cdot D_i^n ds \\
&\quad - \int_0^1 \sum_{i \in \mathbb{Z}} \nabla^2 \eta(w_i^n + s \delta_{i+\frac{1}{2}}^n) \delta_{i+\frac{1}{2}}^n \cdot \frac{\alpha_i^{Ok} + \alpha_{i+1}^{Ok}}{2} ds, \\
&> \frac{1}{2} \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot (\mathfrak{F}_{i+\frac{1}{2}}^{Ok} - \mathfrak{F}_{i-\frac{1}{2}}^{Ok}) + \frac{\lambda}{2} \int_0^1 (1-s) \sum_{i \in \mathbb{Z}} N_i^n(s) D_i^n \cdot D_i^n ds.
\end{aligned}$$

Comme  $\eta$  est strictement convexe, la matrice par bloc  $N_i^n(s)$  est définie positive. Par conséquent, le terme multiplicatif de  $\lambda > 0$  dans l'expression ci-dessus est strictement positif. Comme la somme  $\sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot (\mathfrak{F}_{i+\frac{1}{2}}^{Ok} - \mathfrak{F}_{i-\frac{1}{2}}^{Ok})$  est nécessairement finie d'après l'hypothèse (4.56), il existe  $\lambda > 0$  suffisamment grand tel que

$$\sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot (\mathcal{F}_{i+\frac{1}{2}}^{Ok} - \mathcal{F}_{i-\frac{1}{2}}^{Ok}) > 0.$$

Puisque cette dernière expression s'identifie à l'inégalité de dissipation (4.4), l'inégalité d'entropie discrète globale (4.57) se déduit de la Proposition 4.1.1, ce qui conclut la preuve.  $\square$

D'après la proposition ci-dessus, un schéma numérique d'ordre  $k$  à stabilité entropique globale est entièrement défini par la donnée des quantités  $\alpha_i^{Ok}$  et  $\mathfrak{F}_{i+\frac{1}{2}}^{Ok}$ . Ainsi, la prochaine section propose une définition possible de ces objets pour le cas de schémas numériques d'ordre trois et quatre en espace.

#### 4.4.2 Schémas non limités d'ordre trois et quatre

Cette section propose des schémas numériques du troisième et du quatrième ordre définis à partir de la structure générale de flux numérique (4.55). Comme ces schémas sont d'ordre élevé, ils utilisent nécessairement des approximations des dérivées de  $w$  et de  $f$  pour lesquelles des formulations consistantes sont maintenant présentées.

**Lemme 4.4.2.1** (Approximations consistantes des termes dérivés). *Soit  $u : \mathbb{R} \rightarrow \Omega$  une fonction régulière et  $f : \Omega \rightarrow \mathbb{R}^m$  une fonction de flux de classe  $C^2(\Omega, \mathbb{R}^m)$ . Notons*

$$u_i = \frac{1}{\Delta x} \int_{C_i} u(x) dx, \quad \forall i \in \mathbb{Z}, \tag{4.62}$$

ainsi que  $\delta_{i+\frac{1}{2}} = u_{i+1} - u_i$  dans  $\mathbb{R}^m$  puis définissons les quantités

$$\Delta x \overline{\partial_x u}_{i+\frac{1}{2}}^{O3} = \frac{1}{3} (\delta_{i+\frac{3}{2}} + \delta_{i+\frac{1}{2}} + \delta_{i-\frac{1}{2}}), \tag{4.63a}$$

$$\Delta x \overline{\partial_x u}_{i+\frac{1}{2}}^{O4} = \frac{1}{24} (-\delta_{i+\frac{3}{2}} + 26 \delta_{i+\frac{1}{2}} - \delta_{i-\frac{1}{2}}), \tag{4.63b}$$

$$\Delta x^2 \overline{\partial_{xx} u}_i = \delta_{i+\frac{1}{2}} - \delta_{i-\frac{1}{2}}, \tag{4.63c}$$

$$\Delta x^3 \overline{\partial_{xxx} u}_i = (\delta_{i+\frac{3}{2}} - \delta_{i+\frac{1}{2}}) - (\delta_{i+\frac{1}{2}} - \delta_{i-\frac{1}{2}}), \tag{4.63d}$$

$$\Delta x^2 \overline{\partial_{xx} f(u)}_i = f(u_{i+1}) - 2f(u_i) + f(u_{i-1}), \tag{4.63e}$$

$$\Delta x^2 \overline{\nabla f(u) \partial_{xx} u}_i = \nabla f(u_i) (\delta_{i+\frac{1}{2}} - \delta_{i-\frac{1}{2}}). \tag{4.63f}$$

Alors, on a

$$\Delta x \overline{\partial_x u}_{i+\frac{1}{2}}^{O3} = \Delta x \partial_x u(x_{i+\frac{1}{2}}) + \mathcal{O}(\Delta x^3), \quad (4.64a)$$

$$\Delta x \overline{\partial_x u}_{i+\frac{1}{2}}^{O4} = \Delta x \partial_x u(x_{i+\frac{1}{2}}) + \mathcal{O}(\Delta x^4), \quad (4.64b)$$

$$\Delta x^2 \overline{\partial_{xx} u}_i = \Delta x^2 \partial_{xx} u(x_i) + \mathcal{O}(\Delta x^4), \quad (4.64c)$$

$$\Delta x^3 \overline{\partial_{xxx} u}_i = \Delta x^3 \partial_{xxx} u(x_i) + \mathcal{O}(\Delta x^4), \quad (4.64d)$$

$$\Delta x^2 \overline{\partial_{xx} f(u)}_i = \Delta x^2 \partial_{xx} f(u(x_i)) + \mathcal{O}(\Delta x^4), \quad (4.64e)$$

$$\Delta x^2 \overline{\nabla f(u) \partial_{xx} u}_i = \Delta x^2 \nabla f(u(x_i)) \partial_{xx} u(x_i) + \mathcal{O}(\Delta x^4). \quad (4.64f)$$

*Démonstration.* Concernant l'estimation (4.64a), et comme la cellule  $C_i$  est définie par  $C_i = ]x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}[$  pour tout  $i$  dans  $\mathbb{Z}$ , un changement de variable dans l'expression (4.63a) permet d'écrire

$$\begin{aligned} 3\Delta x \overline{\partial_x u}_{i+\frac{1}{2}}^{O3} &= u_{i+2} - u_{i-1} = \frac{1}{\Delta x} \int_{C_{i+2}} u(x) dx - \frac{1}{\Delta x} \int_{C_{i-1}} u(x) dx, \\ &= \frac{1}{\Delta x} \int_{\Delta x}^{2\Delta x} u(x + x_{i+\frac{1}{2}}) dx - \frac{1}{\Delta x} \int_{-2\Delta x}^{-\Delta x} u(x + x_{i+\frac{1}{2}}) dx. \end{aligned} \quad (4.65)$$

Puisque  $u$  est régulière, un développement de Taylor au voisinage de  $x_{i+\frac{1}{2}}$  donne

$$\begin{aligned} \frac{1}{\Delta x} \int_{\Delta x}^{2\Delta x} u(x + x_{i+\frac{1}{2}}) dx &= u(x_{i+\frac{1}{2}}) + \left( \frac{1}{\Delta x} \int_{\Delta x}^{2\Delta x} x dx \right) \partial_x u(x_{i+\frac{1}{2}}) \\ &\quad + \frac{1}{2} \left( \frac{1}{\Delta x} \int_{\Delta x}^{2\Delta x} x^2 dx \right) \partial_{xx} u(x_{i+\frac{1}{2}}) + \mathcal{O}(\Delta x^3), \\ &= u(x_{i+\frac{1}{2}}) + \frac{3\Delta x}{2} \partial_x u(x_{i+\frac{1}{2}}) + \frac{7\Delta x^2}{6} \partial_{xx} u(x_{i+\frac{1}{2}}) + \mathcal{O}(\Delta x^3). \end{aligned} \quad (4.66)$$

De la même manière,

$$\frac{1}{\Delta x} \int_{-2\Delta x}^{-\Delta x} u(x + x_{i+\frac{1}{2}}) dx = u(x_{i+\frac{1}{2}}) - \frac{3\Delta x}{2} \partial_x u(x_{i+\frac{1}{2}}) + \frac{7\Delta x^2}{6} \partial_{xx} u(x_{i+\frac{1}{2}}) + \mathcal{O}(\Delta x^3).$$

En retranchant membres à membres l'équation ci-dessus à l'équation (4.66) puis en insérant le résultat dans l'équation (4.65), il suit

$$3\Delta x \overline{\partial_x u}_{i+\frac{1}{2}}^{O3} = 3\Delta x \partial_x u(x_{i+\frac{1}{2}}) + \mathcal{O}(\Delta x^3).$$

À une multiplication par trois près, cette dernière égalité correspond au résultat (4.64a) attendu. Puisque l'ensemble des autres résultats (4.64) se montrent avec la même méthode de calcul, ceci conclut la preuve.  $\square$

Pour obtenir des schémas numériques d'ordre trois et quatre de la forme (4.55), il est d'abord nécessaire de définir deux flux numériques annexes  $\mathfrak{F}^{O3}$  et  $\mathfrak{F}^{O4}$  respectivement consistants aux ordres trois et quatre et qui vérifient tous deux l'hypothèse de sommabilité (4.56). Une définition possible de ces objets est donnée par une discrétisation centrée standard.

Ensuite,  $\mathfrak{F}^{O3}$  et  $\mathfrak{F}^{O4}$  doivent être respectivement complétées par des corrections dissipatives du troisième et du quatrième ordre chacune notée  $(\alpha_i^{O3})_{i \in \mathbb{Z}}$  et  $(\alpha_i^{O4})_{i \in \mathbb{Z}}$ . L'idée retenue est l'introduction de matrices  $(\Theta_i^n)_{i \in \mathbb{Z}}$  qui, comme pour le schéma du second ordre des Sections 4.2 et 4.3, permettront d'obtenir la propriété de dissipation.

Une proposition de schémas du troisième et quatrième ordre est maintenant établie selon ces considérations et à l'aide des approximations du Lemme 4.4.2.1.

**Proposition 4.4.2** (Schémas numériques du troisième et du quatrième ordre spatial). *Considérons une séquence  $(w_i^n)_{i \in \mathbb{Z}}$  incluse dans  $\Omega$ . Soient  $(\Theta_i^{O3})_{i \in \mathbb{Z}}$  et  $(\Theta_i^{O4})_{i \in \mathbb{Z}}$  deux séquences de matrices diagonales de taille  $m \times m$  et notons  $I_m$  la matrice identité de taille  $m$ . Considérons ensuite les approximations consistantes définies dans (4.63) puis notons pour tout  $i$  dans  $\mathbb{Z}$*

$$\Delta x \overline{\partial_x w_i^n}^{O3} = \Theta_i^{O3} \Delta x \overline{\partial_x w_{i+\frac{1}{2}}^n}^{O3} + (I_m - \Theta_i^{O3}) \Delta x \overline{\partial_x w_{i-\frac{1}{2}}^n}^{O3} + \frac{1}{2} (I_m - 2\Theta_i^{O3}) \Delta x^2 \overline{\partial_{xx} w_i^n}, \quad (4.67)$$

$$\begin{aligned} \Delta x \overline{\partial_x w_i^n}^{O4} &= \Theta_i^{O4} \Delta x \overline{\partial_x w_{i+\frac{1}{2}}^n}^{O4} + (I_m - \Theta_i^{O4}) \Delta x \overline{\partial_x w_{i-\frac{1}{2}}^n}^{O4} \\ &+ \frac{1}{4} (-\Theta_i^{O4} \Delta x^2 \overline{\partial_{xx} w_{i+1}^n} + (I_m - 2\Theta_i^{O4}) \Delta x^2 \overline{\partial_{xx} w_i^n} + (I_m - \Theta_i^{O4}) \Delta x^2 \overline{\partial_{xx} w_{i-1}^n}) \\ &+ \frac{\Delta x^3}{8} \overline{\partial_{xxx} w_i^n}. \end{aligned} \quad (4.68)$$

Soient  $k$  dans  $\{3, 4\}$ ,  $\lambda > 0$ ,  $(\alpha_i^{Ok})_{i \in \mathbb{Z}}$  des séquences de  $\mathbb{R}^m$  et  $\mathfrak{F}_{i+\frac{1}{2}}^{O4}$  un flux numérique chacun donné par

$$\alpha_i^{O3} = \frac{\lambda}{2} \Delta x \overline{\partial_x w_i^n}^{O3}, \quad (4.69)$$

$$\alpha_i^{O4} = \frac{\lambda}{2} \Delta x \overline{\partial_x w_i^n}^{O4} - \frac{\lambda}{24} \Delta x^3 \overline{\partial_{xxx} w_i^n}, \quad (4.70)$$

$$\begin{aligned} \mathfrak{F}_{i+\frac{1}{2}}^{O4} &= \frac{f(w_i^n) + f(w_{i+1}^n)}{2} - \Delta x^2 \frac{\overline{\partial_{xx} f(w^n)}_i + \overline{\partial_{xx} f(w^n)}_{i+1}}{16} \\ &- \Delta x^2 \frac{\overline{\nabla f(w^n) \partial_{xx} w_i^n} + \overline{\nabla f(w^n) \partial_{xx} w_{i+1}^n}}{48}. \end{aligned} \quad (4.71)$$

Avec la convention  $\mathfrak{F}^{O3} = \mathfrak{F}^{O4}$ , considérons un flux numérique  $\mathcal{F}_{i+\frac{1}{2}}^{Ok}$  de la forme (4.55) défini par les quantités ci-dessus. Si les matrices  $\Theta_i^{Ok}$  sont uniformément bornées selon l'indice  $i$  et bornées lorsque  $\Delta x \rightarrow 0$  alors un schéma numérique défini par  $\mathcal{F}_{i+\frac{1}{2}}^{Ok}$  est d'ordre  $k$  en espace au sens de la Définition formelle 1.4.1.1-iii).

*Démonstration.* Pour montrer le résultat, il est d'abord nécessaire de montrer que  $\mathfrak{F}^{O4}$  (4.71) est du quatrième ordre tel que décrit dans la Définition formelle 1.4.1.1-iii). En réutilisant la définition de  $(u_i)_{i \in \mathbb{Z}}$  donnée par (4.62) et comme  $u$  et  $f$  sont régulières, un développement de Taylor de la quantité  $f(u_i) + f(u_{i+1})$  au voisinage de  $x_{i+\frac{1}{2}}$  permet d'écrire

$$\begin{aligned} \mathfrak{F}_{i+\frac{1}{2}}^{O4} &= f(u(x_{i+\frac{1}{2}})) + \Delta x^2 \frac{\partial_{xx} f(u(x_{i+\frac{1}{2}}))}{8} + \frac{\Delta x^2}{24} \nabla f(u(x_{i+\frac{1}{2}})) \partial_{xx} u(x_{i+\frac{1}{2}}) \\ &- \Delta x^2 \frac{\overline{\partial_{xx} f(u)}_i + \overline{\partial_{xx} f(u)}_{i+1}}{16} - \Delta x^2 \frac{\overline{\nabla f(u) \partial_{xx} u}_i + \overline{\nabla f(u) \partial_{xx} u}_{i+1}}{48} + \mathcal{O}(\Delta x^4). \end{aligned}$$

En utilisant les estimations (4.64e)-(4.64f) dans l'égalité ci dessus, il suit

$$\begin{aligned} \mathfrak{F}_{i+\frac{1}{2}}^{O4} &= f(u(x_{i+\frac{1}{2}})) + \Delta x^2 \frac{\partial_{xx} f(u(x_{i+\frac{1}{2}}))}{8} + \frac{\Delta x^2}{24} \nabla f(u(x_{i+\frac{1}{2}})) \partial_{xx} u(x_{i+\frac{1}{2}}) \\ &- \Delta x^2 \frac{\partial_{xx} f(u(x_i)) + \partial_{xx} f(u(x_{i+1}))}{16} \\ &- \Delta x^2 \frac{\nabla f(u(x_i)) \partial_{xx} u(x_i) + \nabla f(u(x_{i+1})) \partial_{xx} u(x_{i+1})}{48} + \mathcal{O}(\Delta x^4). \end{aligned}$$

Puisque que pour toutes fonctions  $X$  régulières, l'estimation  $X(x_i) + X(x_{i+1}) = 2X(x_{i+\frac{1}{2}}) + \mathcal{O}(\Delta x^2)$  est vérifiée, l'égalité ci-dessus se récrit

$$\mathfrak{F}_{i+\frac{1}{2}}^{O4} = f(u(x_{i+\frac{1}{2}})) + \mathcal{O}(\Delta x^4). \quad (4.72)$$

D'après la Définition 1.4.1.1, cette dernière égalité montre la consistance à l'ordre quatre de la quantité  $\mathfrak{F}_{i+\frac{1}{2}}^{O4}$ .

Il est maintenant nécessaire de montrer que les définitions des séquences  $(\alpha_i^{Ok})_{i \in \mathbb{Z}}$  données par (4.69)-(4.70) assurent la propriété de correction (4.52). Puisque les matrices  $\Theta_i^{Ok}$  sont bornées, l'utilisation des estimations (4.64a) et (4.64c) dans l'expression de  $\alpha_i^{O3}$  donne

$$\frac{2\alpha_i^{O3}}{\lambda} = \Theta_i^{O3} \Delta x \partial_x u(x_{i+\frac{1}{2}}) + (I_m - \Theta_i^{O3}) \Delta x \partial_x u(x_{i-\frac{1}{2}}) + \frac{1}{2} (I_m - 2\Theta_i^{O3}) \Delta x^2 \partial_{xx} u(x_i) + \mathcal{O}(\Delta x^3).$$

En utilisant un développement de Taylor des dérivées premières de  $u$  autour de  $x_i$  l'estimation ci-dessus devient

$$\begin{aligned} \frac{2\alpha_i^{O3}}{\lambda} &= \Theta_i^{O3} \Delta x \left( \partial_x u(x_i) + \frac{\Delta x}{2} \partial_{xx} u(x_i) \right) + (I_m - \Theta_i^{O3}) \Delta x \left( \partial_x u(x_i) - \frac{\Delta x}{2} \partial_{xx} u(x_i) \right) \\ &\quad + \frac{1}{2} (I_m - 2\Theta_i^{O3}) \Delta x^2 \partial_{xx} u(x_i) + \mathcal{O}(\Delta x^3), \\ &= \Delta x \partial_x u(x_i) + \mathcal{O}(\Delta x^3). \end{aligned}$$

Comme des calculs identiques peuvent être conduits pour  $\alpha_{i+1}^{O3}$ , il suit

$$\begin{aligned} -\frac{\lambda}{2} (u_{i+1} - u_i) + \frac{\alpha_i^{O3} + \alpha_{i+1}^{O3}}{2} &= -\frac{\lambda}{2} \Delta x \partial_x u(x_{i+\frac{1}{2}}) + \frac{\lambda}{2} \Delta x \frac{\partial_x u(x_i) + \partial_x u(x_{i+1})}{2} + \mathcal{O}(\Delta x^3), \\ &= \mathcal{O}(\Delta x^3). \end{aligned}$$

L'association de l'égalité ci-dessus à l'estimation (4.72) termine de montrer que le flux numérique  $\mathcal{F}_{i+\frac{1}{2}}^{O3}$  définit un schéma du troisième ordre en espace. Comme la démonstration du quatrième ordre se mène *mutatis mutandis* ceci termine également la démonstration.  $\square$

Par simplicité, le choix  $\mathfrak{F}^{O3} = \mathfrak{F}^{O4}$  est retenu mais d'autres définitions restent possibles. La définition proposée pour  $\mathfrak{F}^{O4}$  dans (4.71) est une discrétisation centrée autour de l'interface or ces discrétisations sont en général instables. Par conséquent, il est désormais nécessaire de contraindre les matrices  $(\Theta_i^{Ok})_{i \in \mathbb{Z}}$  en vertu de la propriété de stabilité entropique attendue.

#### 4.4.3 Stabilité entropique discrète globale

La famille de schémas d'ordres trois et quatre définie par la Proposition 4.4.2 est paramétrée par la donnée d'une séquence de matrices diagonales  $(\Theta_i^{Ok})_{i \in \mathbb{Z}}$  dans  $\mathcal{M}_{m,m}(\mathbb{R})$ . L'objectif de cette section est de montrer que ces matrices peuvent être choisies pour assurer l'existence de l'inégalité d'entropie discrète globale 4.57. Puisque la preuve de cette existence repose sur la Proposition 4.4.1, il est maintenant nécessaire de montrer que les matrices  $\Theta_i^{Ok}$  permettent d'obtenir l'inégalité de dissipation (4.53) associée aux séquences  $(\alpha_i^{O3})_{i \in \mathbb{Z}}$  (4.69) et  $(\alpha_i^{O4})_{i \in \mathbb{Z}}$  (4.70).

Pour montrer ce résultat, un espace de séquence  $S_\eta^{O34}(\mathbb{Z}, \Omega)$  est introduit et il est donné par

$$\begin{aligned} S_\eta^{O34}(\mathbb{Z}, \Omega) &= \left\{ w \in \dot{h}^2(\mathbb{Z}, \Omega) : \right. \\ &\quad \left. \sum_{j=1}^m \sum_{i \in \mathbb{Z}} |(\nabla \eta(w_{i+1}) - \nabla \eta(w_{i-1}))_j| |(w_{i+2} - 4w_{i+1} + 6w_i - 4w_{i-1} + w_{i-2})_j| \neq 0 \right\}. \end{aligned} \quad (4.73)$$

Avec cet espace, un lemme sur l'existence des matrices  $\Theta_i^{Ok}$  peut maintenant être énoncé.

**Lemme 4.4.3.1** (Inégalité de dissipation associée à  $(\alpha_i^{Ok})_{(i,k) \in \mathbb{Z} \times \{3,4\}}$ ). Soient  $K^n$  un compact de  $\Omega$  et  $\eta$  une entropie strictement convexe de classe  $C^2(\Omega, \mathbb{R})$ . Considérons une séquence  $(w_i^n)_{i \in \mathbb{Z}}$  dans  $S_\eta^{O34}$ , non constante, à valeurs et à support compact étendu dans  $K^n$ . Considérons de plus les séquences  $(\alpha_i^{O3})_{i \in \mathbb{Z}}$  et  $(\alpha_i^{O4})_{i \in \mathbb{Z}}$  respectivement définies dans (4.69)-(4.70). Pour  $k$  dans  $\{3,4\}$ , il existe une séquence de matrices diagonales  $(\Theta_i^{Ok})_{i \in \mathbb{Z}}$  dans  $\mathcal{M}_{m,m}(\mathbb{R})$  telle que l'inégalité de dissipation (4.53) associée à  $(\alpha_i^{Ok})_{i \in \mathbb{Z}}$  est satisfaite.

*Démonstration.* D'après les hypothèses de compacité énoncées sur la séquence  $(w_i^n)_{i \in \mathbb{Z}}$ , le membre de gauche de l'inégalité de dissipation (4.53) associée à  $(\alpha_i^{Ok})_{i \in \mathbb{Z}}$  est convergent et il s'écrit

$$\begin{aligned} & \int_0^1 \sum_{i \in \mathbb{Z}} \nabla^2 \eta(w_i^n + s(w_{i+1}^n - w_i^n))(w_{i+1}^n - w_i^n) \cdot \frac{\alpha_i^{Ok} + \alpha_{i+1}^{Ok}}{2} ds \\ &= \sum_{i \in \mathbb{Z}} (\nabla \eta(w_{i+1}^n) - \nabla \eta(w_i^n)) \cdot \frac{\alpha_i^n + \alpha_{i+1}^{Ok}}{2}, \\ &= \frac{1}{2} \sum_{i \in \mathbb{Z}} (\nabla \eta(w_{i+1}^n) - \nabla \eta(w_{i-1}^n)) \cdot \alpha_i^{Ok}. \end{aligned} \quad (4.74)$$

L'objectif est de montrer que la quantité ci-dessus peut être strictement négative. Selon les expressions (4.67)-(4.68)-(4.69)-(4.70),  $\alpha_i^{Ok}$  est proportionnel à  $\lambda$  et affine en  $\Theta_i^{Ok}$ . Par conséquent, pour tout  $i$  dans  $\mathbb{Z}$ , il existe  $\mathcal{A}_i^{Ok}$  et  $\mathcal{B}_i^{Ok}$  des vecteurs de taille  $m$ , linéairement dépendants des quantités  $(\delta_{i+\nu+\frac{1}{2}}^n)_{-2 \leq \nu \leq 1}$  et tels que

$$\alpha_i^{Ok} = \frac{\lambda}{2} (\Theta_i^{Ok} \mathcal{A}_i^{Ok} + \mathcal{B}_i^{Ok}).$$

Comme  $\Theta_i^{Ok}$  est une matrice diagonale, l'introduction de l'égalité ci-dessus dans l'expression (4.74) donne

$$\begin{aligned} & \frac{2}{\lambda} \int_0^1 \sum_{i \in \mathbb{Z}} \nabla^2 \eta(w_i^n + s(w_{i+1}^n - w_i^n))(w_{i+1}^n - w_i^n) \cdot \frac{\alpha_i^{Ok} + \alpha_{i+1}^{Ok}}{2} ds \\ &= \sum_{i \in \mathbb{Z}} (\nabla \eta(w_{i+1}^n) - \nabla \eta(w_{i-1}^n)) \cdot (\Theta_i^{Ok} \mathcal{A}_i^{Ok} + \mathcal{B}_i^{Ok}), \\ &= \sum_{i \in \mathbb{Z}} (\nabla \eta(w_{i+1}^n) - \nabla \eta(w_{i-1}^n)) \cdot \mathcal{B}_i^{Ok} + \sum_{i \in \mathbb{Z}} \Theta_i^{Ok} (\nabla \eta(w_{i+1}^n) - \nabla \eta(w_{i-1}^n)) \cdot \mathcal{A}_i^{Ok}. \end{aligned} \quad (4.75)$$

Un calcul direct à partir des définitions de  $(\alpha_i^{Ok})_{i \in \mathbb{Z}}$  données par (4.67)-(4.68)-(4.69)-(4.70) montre que

$$\mathcal{A}_i^{O3} = \frac{w_{i+2}^n - 4w_{i+1}^n + 6w_i^n - 4w_{i-1}^n + w_{i-2}^n}{3}, \quad (4.76a)$$

$$\mathcal{A}_i^{O4} = -\frac{7}{8} \mathcal{A}_i^{O3}, \quad \forall i \in \mathbb{Z}. \quad (4.76b)$$

Par conséquent, et puisque la séquence  $(w_i^n)_{i \in \mathbb{Z}}$  est dans  $S_\eta^{O34}(\mathbb{Z}, \Omega)$ , on en déduit

$$\sum_{j=1}^m \sum_{i \in \mathbb{Z}} \left| (\nabla \eta(w_{i+1}^n) - \nabla \eta(w_{i-1}^n))_j (\mathcal{A}_i^{Ok})_j \right| > 0.$$

Ainsi, et en utilisant la compacité étendue de  $(w_i^n)_{i \in \mathbb{Z}}$ , il est possible de définir un réel  $\theta^{Ok}$  nécessairement fini et tel que

$$\theta^{Ok} > \frac{\max \left( 0, \sum_{i \in \mathbb{Z}} (\nabla \eta(w_{i+1}^n) - \nabla \eta(w_{i-1}^n)) \cdot \mathcal{B}_i^{Ok} \right)}{\sum_{j=1}^m \sum_{i \in \mathbb{Z}} \left| (\nabla \eta(w_{i+1}^n) - \nabla \eta(w_{i-1}^n))_j (\mathcal{A}_i^{Ok})_j \right|}. \quad (4.77)$$

Par suite, si les matrices diagonales  $\Theta_i^{Ok}$  de  $\mathcal{M}_{m,m}(\mathbb{R})$  sont telles que

$$\Theta_i^{Ok} = -\theta^{Ok} \operatorname{diag}_{1 \leq j \leq m} \left( \operatorname{sign}((\nabla \eta(w_{i+1}^n) - \nabla \eta(w_{i-1}^n))_j (\mathcal{A}_i^{Ok})_j) \right), \quad (4.78)$$

alors, d'après le résultat intermédiaire (4.75), l'inégalité stricte (4.53) est satisfaite ce qui termine la preuve.  $\square$

Les matrices  $\Theta_i^{Ok}$  définies dans (4.78) sont des extensions des matrices proposées en Section 4.3.3 pour le schéma du second ordre. Les vecteurs  $\mathcal{B}_i^{Ok}$  de l'inégalité (4.77) sont introduits formellement dans la démonstration ci-dessus mais un calcul direct montre que leurs expressions sont données par

$$\begin{aligned} \mathcal{B}_i^{O3} &= \frac{5\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n + 2\delta_{i-\frac{3}{2}}^n}{6}, \\ \mathcal{B}_i^{O4} &= \frac{2\delta_{i+\frac{3}{2}}^n + \delta_{i+\frac{1}{2}}^n + 4\delta_{i-\frac{1}{2}}^n - 7\delta_{i-\frac{3}{2}}^n}{24}, \quad \forall i \in \mathbb{Z}. \end{aligned}$$

Puisque l'inégalité stricte (4.53) est assurée, le Lemme 4.4.3.1 et la Proposition 4.4.2 montrent ainsi que les séquences  $(\alpha_i^{O3})_{i \in \mathbb{Z}}$  et  $(\alpha_i^{O4})_{i \in \mathbb{Z}}$  respectivement définies dans (4.69) et dans (4.70) sont des corrections dissipatives d'ordre trois et quatre au sens de la Définition 4.4.1. Par conséquent, un résultat de stabilité entropique globale est désormais accessible.

**Théorème 4.4.3.1** (Stabilité entropique discrète globale des schémas du troisième et du quatrième ordre). *Soient  $K^n$  un compact de  $\Omega$  et  $\eta$  une entropie strictement convexe de classe  $C^2(\Omega, \mathbb{R})$ . Considérons une séquence  $(w_i^n)_{i \in \mathbb{Z}}$  dans  $S_\eta^{O34}(\mathbb{Z}, \Omega)$ , non constante, à valeurs et à support compact étendu dans  $K^n$ . Pour  $k$  dans  $\{3, 4\}$ , supposons également que les séquences de matrices diagonales  $(\Theta_i^{Ok})_{i \in \mathbb{Z}}$  dans  $\mathcal{M}_{m,m}(\mathbb{R})$  utilisées dans (4.67)-(4.68) sont uniformément bornées selon l'indice  $i$ , bornées lorsque  $\Delta x \rightarrow 0$  et telles que l'inégalité stricte (4.53) est satisfaite.*

*Il existe  $\lambda > 0$  suffisamment grand et  $\Delta t > 0$  suffisamment petit tels que les schémas du troisième et du quatrième ordre respectivement définis par (4.55), (4.71), (4.69) et (4.55), (4.71), (4.69) vérifient chacun une stabilité entropique discrète globale de la forme.*

$$\sum_{i \in \mathbb{Z}} \frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} \Delta x \leq 0. \quad (4.79)$$

*Démonstration.* Pour  $k$  dans  $\{3, 4\}$  et d'après la preuve de la Proposition 4.4.2,  $(\alpha_i^{Ok})_{i \in \mathbb{Z}}$  vérifient

$$-\frac{\lambda}{2}(w_{i+1}^n - w_i^n) + \frac{\alpha_{i+1}^{Ok} + \alpha_i^{Ok}}{2} = \mathcal{O}(\Delta x^k).$$

Ainsi, et comme les matrices  $\Theta_i^{Ok}$  satisfont de plus l'inégalité stricte (4.53) (voir le Lemme 4.4.3.1), on en déduit que les séquences  $(\alpha_i^{Ok})_{i \in \mathbb{Z}}$  sont des corrections dissipatives d'ordre  $k$ .

Avec les hypothèses énoncées sur la séquence  $(w_i^n)_{i \in \mathbb{Z}}$  et d'après la définition de  $\mathfrak{F}_{i+1/2}^{O4}$  donnée dans (4.71), la condition de sommabilité

$$\left| \sum_{i \in \mathbb{Z}} \nabla \eta(w_i^n) \cdot \frac{\mathfrak{F}_{i+\frac{1}{2}}^{O4} - \mathfrak{F}_{i-\frac{1}{2}}^{O4}}{\Delta x} \right| < +\infty, \quad \forall \lambda > 0,$$

est satisfaite.

Puisque toutes les hypothèses de la Proposition 4.4.1 sont vérifiées, on en déduit l'inégalité d'entropie discrète globale (4.79) ce qui termine la preuve.  $\square$

Les schémas numériques d'ordre élevé définis dans cette section sont les extensions directes du schéma du second ordre introduit en Section 4.2. Par conséquent, la nécessité d'une discrétisation en temps d'ordre élevé demeure également pour ces schémas.

Les mises en œuvre du Théorème 4.4.3.1 et du Théorème 4.3.4.1 qui concerne le schéma du second ordre sont maintenant illustrées par des expériences numériques.

## 4.5 Résultats de simulations numériques

### 4.5.1 Méthodes et paramétrages

Cette section présente les résultats numériques des schémas du second, du troisième et du quatrième ordre définis en Sections 4.2 et 4.4.2. D'après ces sections, et pour chacun des ordres  $k$ , une famille de schémas paramétrés par des matrices  $\Theta_i^{O_k}$  dans  $\mathcal{M}_{m,m}(\mathbb{R})$  est proposée. Ainsi, pour conserver la clarté des graphiques, les résultats sont répartis en deux sous sections.

La première sous section est consacrée au second ordre. D'après le Théorème 4.3.4.1, les schémas proposés nécessitent la sélection de trois paramètres :  $\Theta_i^{O_2}$  dans  $\mathcal{M}_{m,m}(\mathbb{R})$ ,  $\lambda > 0$  et  $\Delta t > 0$ . En s'appuyant sur les Lemmes 4.3.3.1 et 4.3.3.2, plusieurs séquences de matrices  $(\Theta_i^{O_2})_{i \in \mathbb{Z}}$  sont proposées. Ensuite, le Théorème 4.3.4.1 donne des formulations explicites pour  $\lambda$  et pour  $\Delta t$  respectivement détaillées dans (4.39) et dans (4.40). Néanmoins, et notamment d'après [105, 118], ces formulations ne sont pas usuelles, puisqu'en notant  $\{\mu_k(w)\}_{k \in \{1, \dots, m\}}$  l'ensemble des  $m$  valeurs propres de  $\nabla f(w)$ , les paramètres  $\lambda > 0$  et  $\Delta t > 0$  sont généralement sélectionnés selon les conditions

$$\lambda = \max_{i,k} |\mu_k(w_i^n)|, \quad \frac{\lambda \Delta t}{\Delta x} \leq \frac{1}{2}. \quad (4.80)$$

Par conséquent, un comparatif des deux paramétrages est conduit dans le cas scalaire de l'équation de Burgers.

Ensuite, et conformément à l'analyse de la section 4.3.5, une montée à l'ordre deux en temps est conduite à l'aide d'une discrétisation *Runge-Kutta Strong Stability Preserving* (RKSSP) [107, 109, 110]. D'après [76, 230], cette montée en ordre temporel préserve le résultat de stabilité entropique discret global (4.36). Avec cette discrétisation, des résultats numériques pour les équations d'Euler sont présentés avec  $\lambda > 0$  et  $\Delta t > 0$  choisis selon les conditions usuelles données par (4.80). Pour contrôler la validité de ces choix, la quantité

$$\frac{\Delta \mathcal{S}}{\Delta t} = \sum_{i \in \mathbb{Z}} \frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} \Delta x, \quad (4.81)$$

est représentée graphiquement. D'après le Théorème 4.3.4.1, si le triplet  $(\Theta_i^{O_2}, \lambda, \Delta t)$  est convenablement choisis alors  $\Delta \mathcal{S}/\Delta t$  défini par (4.81) est négatif.

Enfin, les résultats des schémas du second ordre proposés sont comparés aux résultats du schéma MUSCL [219] associé au flux numérique de Rusanov [185] et équipé d'un limiteur de pente minmod. Dans le cas scalaire de l'équation de Burgers, ce schéma MUSCL est associé à une discrétisation temporelle du premier ordre et pour le cas des équations d'Euler il est associé à une discrétisation en temps d'ordre deux de type RKSSP.

Le seconde sous section concerne les schémas d'ordre trois et quatre. Ces schémas sont tout d'abord associés à des discrétisations temporelles RKSSP du même ordre. Ensuite, plusieurs choix de matrices  $\Theta_i^{O_k}$  sont adoptés à l'aide du Lemme 4.4.3.1. Pour un choix de matrice  $\Theta_i^{O_k}$  donné, les paramètres  $\lambda > 0$  et  $\Delta t > 0$  sont sélectionnés selon les conditions (4.80) puis la validité du triplet  $(\Theta_i^{O_k}, \lambda, \Delta t)$  est vérifiée numériquement avec le tracé de  $\Delta \mathcal{S}/\Delta t$  défini par (4.81)

Dans les deux sous sections, les erreurs  $L^1$ ,  $L^2$  et  $L^\infty$  entre la solution exacte et la solution numérique obtenue sont présentées sous forme de tableaux. Ces erreurs sont calculées avec les schémas équipés des discrétisations en temps RKSSP du même ordre. De plus, pour les comparaisons graphiques, une référence de la quantité  $\Delta \mathcal{S}/\Delta t$  donnée par (4.81) est calculée avec le schéma du premier ordre de Rusanov [185] sur un maillage fin constitué de 50 000 cellules.

### 4.5.2 Schéma numérique du second ordre

#### 4.5.2.1 L'équation de Burgers scalaire

Cette section s'intéresse à l'équation de Burgers scalaire dans laquelle l'inconnue  $w$  prend ses valeurs dans  $\mathbb{R}$  et la fonction de flux  $f$  vérifie  $f(w) = w^2/2$ . Avec ces définitions, l'équation de Burgers est équipée de l'entropie quadratique  $\eta(w) = w^2/2$ .

Le schéma du second ordre est défini par son flux numérique (4.35) dont l'expression est donnée par

$$\mathcal{F}_{i+\frac{1}{2}}^n = \frac{f(w_i^n) + f(w_{i+1}^n)}{2} - \frac{\lambda}{2} (w_{i+1}^n - w_i^n) + \frac{1}{2} (\alpha_i^n + \alpha_{i+1}^n), \quad (4.82a)$$

$$\alpha_i^n = \frac{\lambda}{2} (\Theta_i^n (w_{i+1}^n - w_i^n) + (I_m - \Theta_i^n) (w_i^n - w_{i-1}^n)), \quad \forall i \in \mathbb{Z}. \quad (4.82b)$$

En considérant un réel  $\theta$  satisfaisant l'inégalité stricte (4.29) et  $\varepsilon = 10^{-12}$ , les matrices  $\Theta_i^{O2}$  proposées pour l'équation de Burgers sont dans ce cas des nombres réels et ils sont donnés par

$$\Theta_{a,i}^{O2} = -\theta \operatorname{sign}((\delta_{i+\frac{1}{2}}^n)^2 - (\delta_{i-\frac{1}{2}}^n)^2), \quad (4.83a)$$

$$\Theta_{b,i}^{O2} = -\theta \tanh((\delta_{i+\frac{1}{2}}^n)^2 - (\delta_{i-\frac{1}{2}}^n)^2), \quad (4.83b)$$

$$\Theta_{c,i}^{O2} = \frac{((\delta_{i-\frac{1}{2}}^n)^2 - (\delta_{i+\frac{1}{2}}^n)^2)((\delta_{i-\frac{1}{2}}^n)^2 + (\delta_{i+\frac{1}{2}}^n)^2)}{((\delta_{i-\frac{1}{2}}^n)^2 + (\delta_{i+\frac{1}{2}}^n)^2)^2 + \varepsilon}, \quad (4.83c)$$

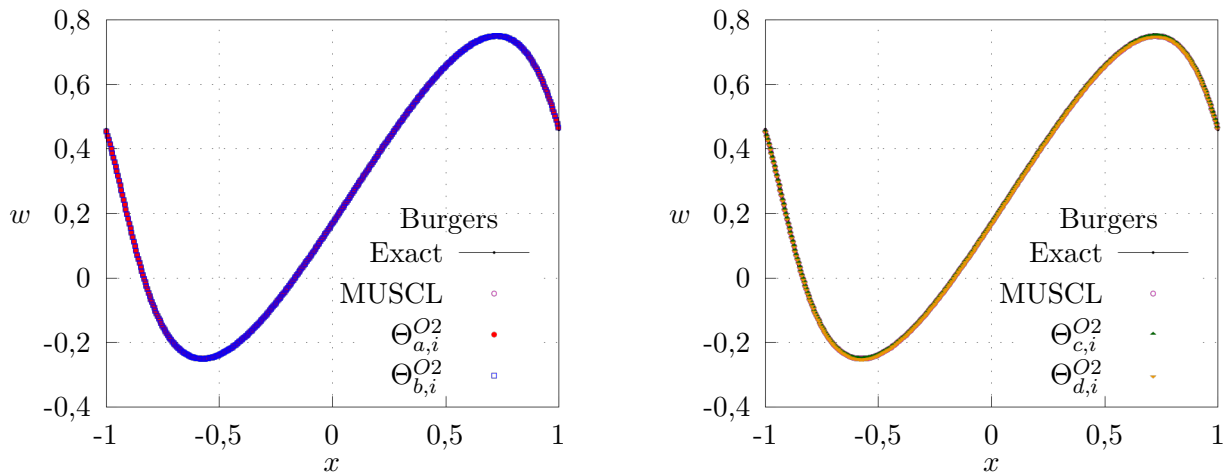
$$\Theta_{d,i}^{O2} = \frac{1}{2}. \quad (4.83d)$$

Pour le cas de l'équation de Burgers, les paramètres  $\lambda > 0$  et  $\Delta t > 0$  sont sélectionnés selon les conditions (4.39)-(4.40). Pour conduire un comparatif, ces paramètres sont par ailleurs calculés selon les conditions (4.80) puis appliqués à un schéma MUSCL tel que décrit en Section 4.5.1.

Le domaine spatial  $[-1, 1]$  du premier cas test est discrétisé avec 400 cellules et la condition initiale est

$$w_0(x) = 0,25 + 0,5 \sin(\pi x). \quad (4.84)$$

Des conditions de bord périodiques sont imposées aux deux frontières du domaine. Sous ces conditions et en sélectionnant un temps final de simulation suffisamment court, la solution reste régulière. Par conséquent, les ordres de précision spatiale peuvent être mesurés. La Figure 4.1 montre les résultats comparés au temps final de simulation  $t = 0,3$  et les tableaux 4.1 rapportent les erreurs numériques ainsi que les ordres pour différents maillages.





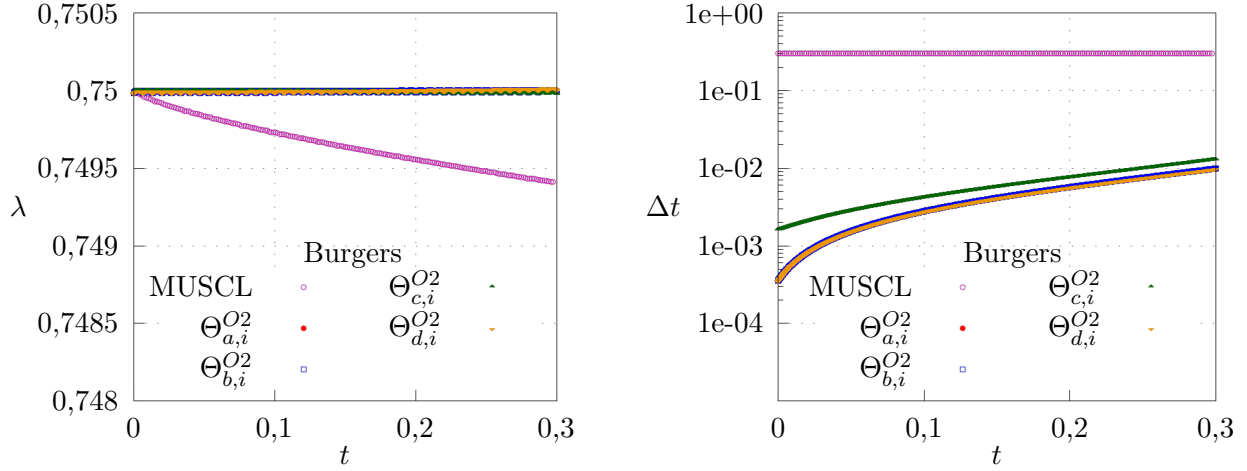


FIGURE 4.1 – En haut, résultats numériques produits par le schéma du second ordre du Théorème 4.3.4.1 sur un maillage de 400 cellules, avec les triplets  $((\Theta_{\alpha,i}^{O2})_{\alpha \in \{a,b,c,d\}}, \lambda, \Delta t)$  donnés dans (4.83)-(4.39)-(4.40) et pour des solutions régulières de l'équation de Burgers au temps final  $t = 0,3$ . En bas, comparatif entre les paramètres  $(\lambda, \Delta t)$  donnés par les contraintes (4.39)-(4.40) et les restrictions usuelles (4.80) appliquées au schéma MUSCL décrit en Section 4.5.1.

D'après les graphiques de la Figure 4.1, les résultats sont très satisfaisants mais la condition CFL (4.40) est plus restrictive que la condition CFL usuelle (4.80). Cette observation peut se quantifier à l'aide de  $\tau > 0$  le rapport des temps CPU nécessaires à la simulation pour les schémas proposés et pour le schéma MUSCL standard. Après mesures, les valeurs de  $\tau$  sont

$$\tau_{\Theta_{i,a}^{O2}} = 3111,71, \quad \tau_{\Theta_{i,b}^{O2}} = 3250,71, \quad \tau_{\Theta_{i,c}^{O2}} = 1633,14, \quad \tau_{\Theta_{i,d}^{O2}} = 3053,29.$$

Selon les résultats ci-dessus, les schémas proposés associés aux paramétrages (4.39)-(4.40) demandent un temps de calcul nettement supérieur aux schémas MUSCL. Néanmoins, une mesure directe révèle qu'environ 97% du temps est employé au calcul du couple  $(\lambda, \Delta t)$  selon les contraintes (4.39)-(4.40) or, à condition d'utiliser une discrétisation temporelle du second ordre, la section 4.3.5 suggère de substituer ces contraintes par (4.80). Comme les conditions usuelles (4.80) se calculent facilement, le temps CPU peut donc largement se réduire à l'aide d'une discrétisation en temps du second ordre de type RKSSP. En considérant une telle discrétisation couplée aux conditions usuelles (4.80), une nouvelle mesure de  $\tau$  donne

$$\tau_{\Theta_{i,a}^{O2}} = 0,78, \quad \tau_{\Theta_{i,b}^{O2}} = 0,94, \quad \tau_{\Theta_{i,c}^{O2}} = 0,61, \quad \tau_{\Theta_{i,d}^{O2}} = 0,59.$$

Dans ce cas, les temps de calcul sont comparables à ceux du schéma MUSCL standard décrit en Section 4.5.1.

Erreurs du schéma du second ordre avec $\Theta_i^n = \Theta_{a,i}^{O2}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	5,7E-04	-	6,8E-04	-	1,8E-03	-
200	1,4E-04	2,0	1,7E-04	2,0	4,5E-04	2,0
400	3,5E-05	2,0	4,1E-05	2,0	1,1E-04	2,0
800	8,8E-06	2,0	1,0E-05	2,0	2,8E-05	2,0
1600	2,2E-06	2,0	2,6E-06	2,0	6,9E-06	2,0
Erreurs du schéma du second ordre avec $\Theta_i^n = \Theta_{b,i}^{O2}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	5,7E-04	-	6,8E-04	-	1,8E-03	-
200	1,4E-04	2,0	1,7E-04	2,0	4,5E-04	2,0
400	3,5E-05	2,0	4,1E-05	2,0	1,1E-04	2,0
800	8,8E-06	2,0	1,0E-05	2,0	2,8E-05	2,0
1600	2,2E-06	2,0	2,6E-06	2,0	6,9E-06	2,0

Erreurs du schéma du second ordre avec $\Theta_i^n = \Theta_{c,i}^{O2}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	1,4E-03	-	1,6E-03	-	4,1E-03	-
200	2,4E-04	2,5	2,8E-04	2,6	7,6E-04	2,4
400	3,9E-05	2,7	4,3E-05	2,7	1,1E-04	2,8
800	8,7E-06	2,1	1,0E-05	2,1	2,7E-05	2,0
1600	2,2E-06	2,0	2,6E-06	2,0	6,8E-06	2,0
Erreurs du schéma du second ordre avec $\Theta_i^n = \Theta_{d,i}^{O2}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	4,4E-04	-	4,8E-04	-	1,1E-03	-
200	1,1E-04	2,0	1,1E-04	2,1	2,7E-04	2,1
400	2,6E-05	2,0	2,8E-05	2,0	6,6E-05	2,0
800	6,5E-06	2,0	6,8E-06	2,0	1,6E-05	2,0
1600	1,6E-06	2,0	1,7E-06	2,0	4,0E-06	2,0

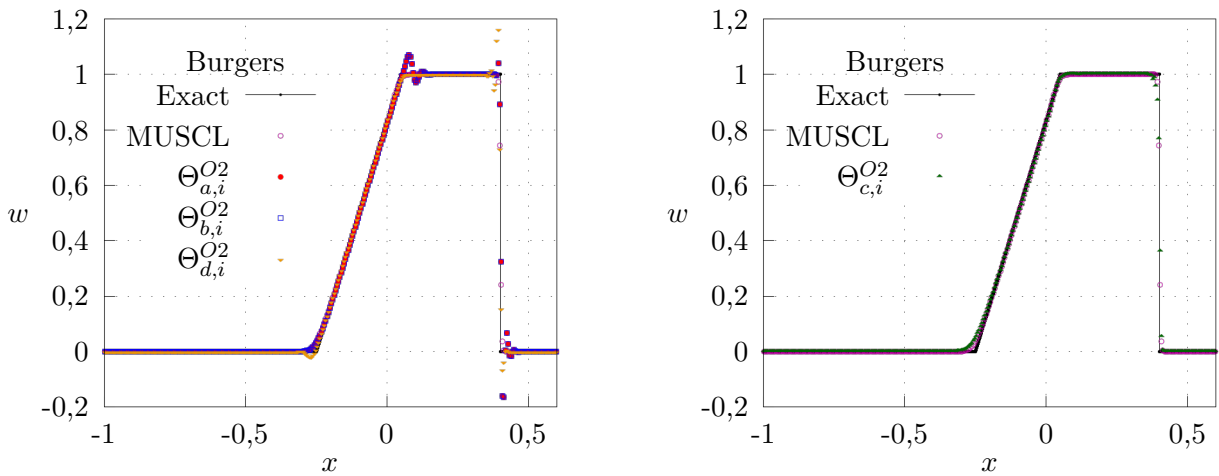
TABLE 4.1 – Évaluations des erreurs et des ordres numériques produits par le schéma du second ordre temps et espace du Théorème 4.3.4.1 avec avec les triplets  $((\Theta_{\alpha,i}^{O2})_{\alpha \in \{a,b,c,d\}}, \lambda, \Delta t)$  donnés dans (4.83)-(4.80) et à partir de solution régulière de l'équation de Burgers au temps final 0,3.

D'après les tableaux 4.1, et comme les schémas proposés n'utilisent pas de technique de limitation, le second ordre spatial est exactement retrouvé dans les expériences numériques. Ce second ordre n'est pas observé avec le schéma MUSCL décrit en Section 4.5.1 puisque, pour le même cas test, la mesure de ses ordres donne : 1,89 pour la norme  $L^1$ , 1,65 pour la norme  $L^2$  et 1,29 pour la norme  $L^\infty$ .

Le second cas test concerne une solution discontinue. Le domaine  $[-1, 1]$  est discrétisé avec 400 cellules et la condition initiale est

$$w_0(x) = \begin{cases} 1 & \text{si } -0,25 \leq x \leq 0,25, \\ 0 & \text{sinon.} \end{cases} \quad (4.85)$$

Des conditions de bord périodiques sont imposées aux deux frontières du domaine. Si le temps final de simulation est suffisamment court, la solution exacte se compose d'une onde de détente et d'une onde de choc qui n'interagissent pas entre elles. La Figure 4.2 montre les résultats comparés au temps final de simulation  $t = 0,3$  et les tableaux 4.2 rapportent les erreurs numériques ainsi que les ordres pour différents maillages.



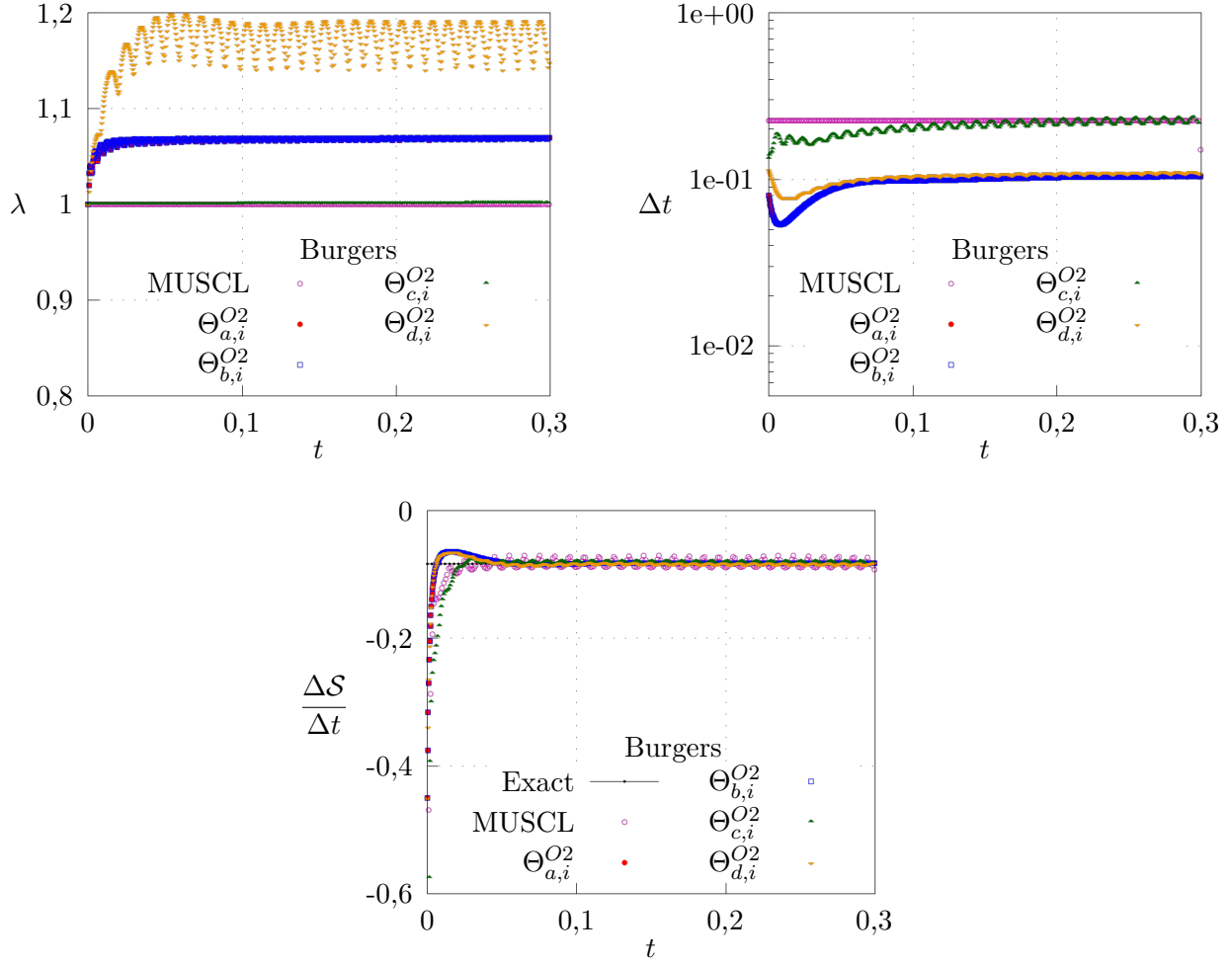


FIGURE 4.2 – En haut, résultats numériques produits par le schéma du second ordre du Théorème 4.3.4.1 sur un maillage de 400 cellules, avec les triplets  $((\Theta_{\alpha,i}^{O2})_{\alpha \in \{a,b,c,d\}}, \lambda, \Delta t)$  donnés dans (4.83)-(4.39)-(4.40) et pour des solutions discontinues de l'équation de Burgers au temps final  $t = 0,3$ . Au milieu, comparatif entre les paramètres  $(\lambda, \Delta t)$  donnés par les contraintes (4.39)-(4.40) et les restrictions usuelles (4.80) appliquées à un schéma MUSCL décrit en Section 4.5.1. En bas, évolution temporelle de  $\Delta S/\Delta t$  défini par (4.81).

D'après la Figure 4.2, pour tous les choix de  $\Theta_i^{O2}$  proposés dans (4.83), la quantité  $\Delta S/\Delta t$  définie par (4.81) est négative. Par conséquent, la stabilité entropique discrète globale est vérifiée. En particulier, le choix  $\Theta_{c,i}^{O2}$  (4.83c) donne des résultats remarquables sans oscillations parasites autour de la discontinuité. La présence d'une telle discontinuité semble également améliorer la condition CFL (4.40) puisqu'une mesure de  $\tau$  pour ce cas test donne

$$\tau_{\Theta_{i,a}^n} = 79,87, \quad \tau_{\Theta_{i,b}^n} = 77,24, \quad \tau_{\Theta_{i,c}^n} = 36,01, \quad \tau_{\Theta_{i,d}^n} = 69,33.$$

Les schémas proposés associés aux paramétrages (4.39)-(4.40) demandent donc une nouvelle fois un temps de calcul nettement supérieur aux schémas MUSCL mais, comme dans le cas d'une solution régulière, environ 97% du temps CPU est utilisé au calcul du couple  $(\lambda, \Delta t)$  selon les contraintes (4.39)-(4.40). Cette situation se corrige en utilisant une discrétisation en temps d'ordre deux RKSSP couplée aux conditions usuelles (4.80). Dans ce cas, une nouvelle mesure de  $\tau$  donne

$$\tau_{\Theta_{i,a}^n} = 0,89, \quad \tau_{\Theta_{i,b}^n} = 0,93, \quad \tau_{\Theta_{i,c}^n} = 0,65, \quad \tau_{\Theta_{i,d}^n} = 0,73.$$

Erreurs du schéma du second ordre avec $\Theta_i^n = \Theta_{a,i}^{O2}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	3,4E-02	-	6,4E-02	-	3,3E-01	-
200	1,7E-02	1,0	4,3E-02	0,6	3,3E-01	0,0
400	8,5E-03	1,0	3,0E-02	0,5	3,2E-01	0,0
800	4,3E-03	1,0	2,1E-02	0,5	3,2E-01	0,0
1600	2,1E-03	1,0	1,5E-02	0,5	3,2E-01	0,0
Erreurs du schéma du second ordre avec $\Theta_i^n = \Theta_{b,i}^{O2}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	5,4E-02	-	9,0E-02	-	3,8E-01	-
200	2,7E-02	1,0	5,9E-02	0,6	3,8E-01	0,0
400	1,4E-02	1,0	4,0E-02	0,6	3,8E-01	0,0
800	7,0E-03	1,0	2,7E-02	0,6	3,8E-01	0,0
1600	3,5E-03	1,0	1,8E-02	0,5	3,8E-01	0,0
Erreurs du schéma du second ordre avec $\Theta_i^n = \Theta_{c,i}^{O2}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	3,5E-02	-	7,1E-02	-	3,6E-01	-
200	1,8E-02	0,9	4,9E-02	0,5	3,6E-01	0,0
400	9,2E-03	1,0	3,4E-02	0,5	3,6E-01	0,0
800	4,6E-03	1,0	2,3E-02	0,5	3,6E-01	0,0
1600	2,3E-03	1,0	1,6E-02	0,5	3,6E-01	0,0
Erreurs du schéma du second ordre avec $\Theta_i^n = \Theta_{d,i}^{O2}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	3,1E-02	-	5,8E-02	-	2,8E-01	-
200	1,4E-02	1,1	4,0E-02	0,6	2,8E-01	0,0
400	7,1E-03	1,0	2,8E-02	0,5	2,8E-01	0,0
800	3,5E-03	1,0	1,9E-02	0,5	2,8E-01	0,0
1600	1,7E-03	1,0	1,4E-02	0,5	2,8E-01	0,0

TABLE 4.2 – Évaluations des erreurs et des ordres numériques produits par le schéma du second ordre temps et espace du Théorème 4.3.4.1 avec avec les triplets  $((\Theta_{\alpha,i}^{O2})_{\alpha \in \{a,b,c,d\}}, \lambda, \Delta t)$  donnés dans (4.83)-(4.80) et à partir de solution discontinue de l'équation de Burgers au temps final 0,3.

#### 4.5.2.2 Les équations d'Euler

Cette section est consacrée aux équations d'Euler pour lesquelles l'inconnue  $w$  s'écrit  $(\rho, \rho u, \rho E)^T$  dans  $\mathbb{R}^3$  et elle prend ses valeurs dans  $\Omega = \{(\rho, \rho u, \rho E) \in \mathbb{R}^3 : \rho > 0, E - u^2/2 > 0\}$ . La fonction de flux est  $f(w) = (\rho u, \rho u^2 + p, \rho E u + p u)^T$  et comme un gaz parfait diatomique est considéré, la fonction  $p$  est donnée par  $p(w) = (\gamma - 1)(\rho E - \rho u^2/2)$  avec  $\gamma = 1,4$ . Parmi les entropies possibles [22, 88, 138], l'entropie convexe retenue pour les équations d'Euler est  $\eta(w) = -\rho \ln(p/\rho^\gamma)$ .

Avec ces définitions, quatre simulations numériques sont réalisées. La première concerne une solution régulière à partir de laquelle les ordres spatiaux sont mesurés. Les deux suivantes présentent des solutions discontinues et la dernière illustre le comportement des schémas dans les zones *proches* du bord de  $\Omega$  caractérisées par  $\rho = 0$  et  $p = 0$ .

Le schéma du second ordre est défini par son flux numérique (4.35) qui s'écrit

$$\mathcal{F}_{i+\frac{1}{2}}^n = \frac{f(w_i^n) + f(w_{i+1}^n)}{2} - \frac{\lambda}{2} (w_{i+1}^n - w_i^n) + \frac{1}{2} (\alpha_i^n + \alpha_{i+1}^n), \quad (4.86a)$$

$$\alpha_i^n = \frac{\lambda}{2} (\Theta_i^n (w_{i+1}^n - w_i^n) + (I_m - \Theta_i^n) (w_i^n - w_{i-1}^n)), \quad \forall i \in \mathbb{Z}. \quad (4.86b)$$

En considérant un réel  $\theta$  satisfaisant l'inégalité stricte (4.29) et  $\varepsilon = 10^{-12}$ , les matrices diagonales  $\Theta_i^{O2}$

dans  $\mathcal{M}_{3,3}(\mathbb{R})$  proposées pour les équations d'Euler sont

$$\Theta_{a,i}^{O2} = -\theta \operatorname{diag}_{1 \leq j \leq 3} \left( \operatorname{sign}((\nabla \eta(w_{i+1}^n) - \nabla \eta(w_{i-1}^n))_j (\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n)_j) \right), \quad (4.87a)$$

$$\Theta_{b,i}^{O2} = -\theta \operatorname{diag}_{1 \leq j \leq 3} \left( \tanh((\nabla \eta(w_{i+1}^n) - \nabla \eta(w_{i-1}^n))_j (\delta_{i+\frac{1}{2}}^n - \delta_{i-\frac{1}{2}}^n)_j) \right), \quad (4.87b)$$

$$\Theta_{c,\varepsilon,i}^{O2} = \operatorname{diag}_{1 \leq j \leq 3} \left( \frac{\left( (\delta_{i-\frac{1}{2}}^n)_j^2 - (\delta_{i+\frac{1}{2}}^n)_j^2 \right) \left( (\delta_{i-\frac{1}{2}}^n)_j^2 + (\delta_{i+\frac{1}{2}}^n)_j^2 \right)}{\left( (\delta_{i-\frac{1}{2}}^n)_j^2 + (\delta_{i+\frac{1}{2}}^n)_j^2 \right) + \varepsilon} \right). \quad (4.87c)$$

Le domaine spatial  $[-1, 1]$  du premier cas test est discrétisé avec 400 cellules et la condition initiale est

$$\rho_0(x) = 1 + 0,5 \sin^2(\pi x), \quad u_0(x) = 0,5, \quad p_0(x) = 1. \quad (4.88)$$

Les conditions de bord sont périodiques aux deux extrémités du domaine spatial. Avec ces conditions, les équations d'Euler se résument à un transport linéaire et la solution demeure régulière. La Figure 4.3 montre les résultats comparés au temps final de simulation  $t = 0,2$  et les tableaux 4.3 rapportent les erreurs numériques ainsi que les ordres pour différents maillages.

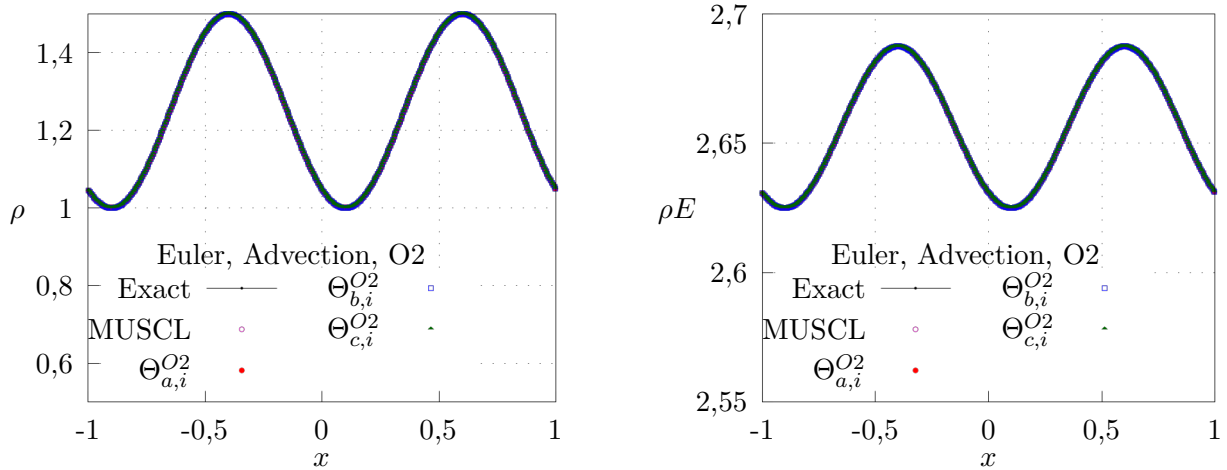


FIGURE 4.3 – Résultats numériques du schéma du second ordre temps et espace du Théorème 4.3.4.1 sur un maillage de 400 cellules, avec les triplets  $((\Theta_{\alpha,i}^{O2})_{\alpha \in \{a,b,c\}}, \lambda, \Delta t)$  donnés dans (4.87)-(4.80) et pour des solutions régulières des équations d'Euler au temps final  $t = 0,2$ .

Erreurs du schéma du second ordre avec $\Theta_i^n = \Theta_{a,i}^{O2}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	2,1E-03	-	1,7E-03	-	1,7E-03	-
200	5,3E-04	2,0	4,2E-04	2,0	4,2E-04	2,0
400	1,3E-04	2,0	1,1E-04	2,0	1,1E-04	2,0
800	3,3E-05	2,0	2,6E-05	2,0	2,6E-05	2,0
1600	8,4E-06	2,0	6,6E-06	2,0	6,6E-06	2,0
Erreurs du schéma du second ordre avec $\Theta_i^n = \Theta_{b,i}^{O2}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	2,1E-03	-	1,7E-03	-	1,7E-03	-
200	5,3E-04	2,0	4,2E-04	2,0	4,2E-04	2,0
400	1,3E-04	2,0	1,1E-04	2,0	1,1E-04	2,0
800	3,3E-05	2,0	2,6E-05	2,0	2,6E-05	2,0
1600	8,4E-06	2,0	6,6E-06	2,0	6,6E-06	2,0

cellule	Erreurs du schéma du second ordre avec $\Theta_i^n = \Theta_{c,i}^{O2}$					
	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	7,3E-03	-	6,8E-03	-	1,1E-02	-
200	1,5E-03	2,3	1,5E-03	2,2	2,9E-03	1,9
400	2,1E-04	2,8	1,8E-04	3,0	3,7E-04	3,0
800	3,7E-05	2,5	2,8E-05	2,7	2,6E-05	3,8
1600	8,4E-06	2,1	6,6E-06	2,1	6,5E-06	2,0

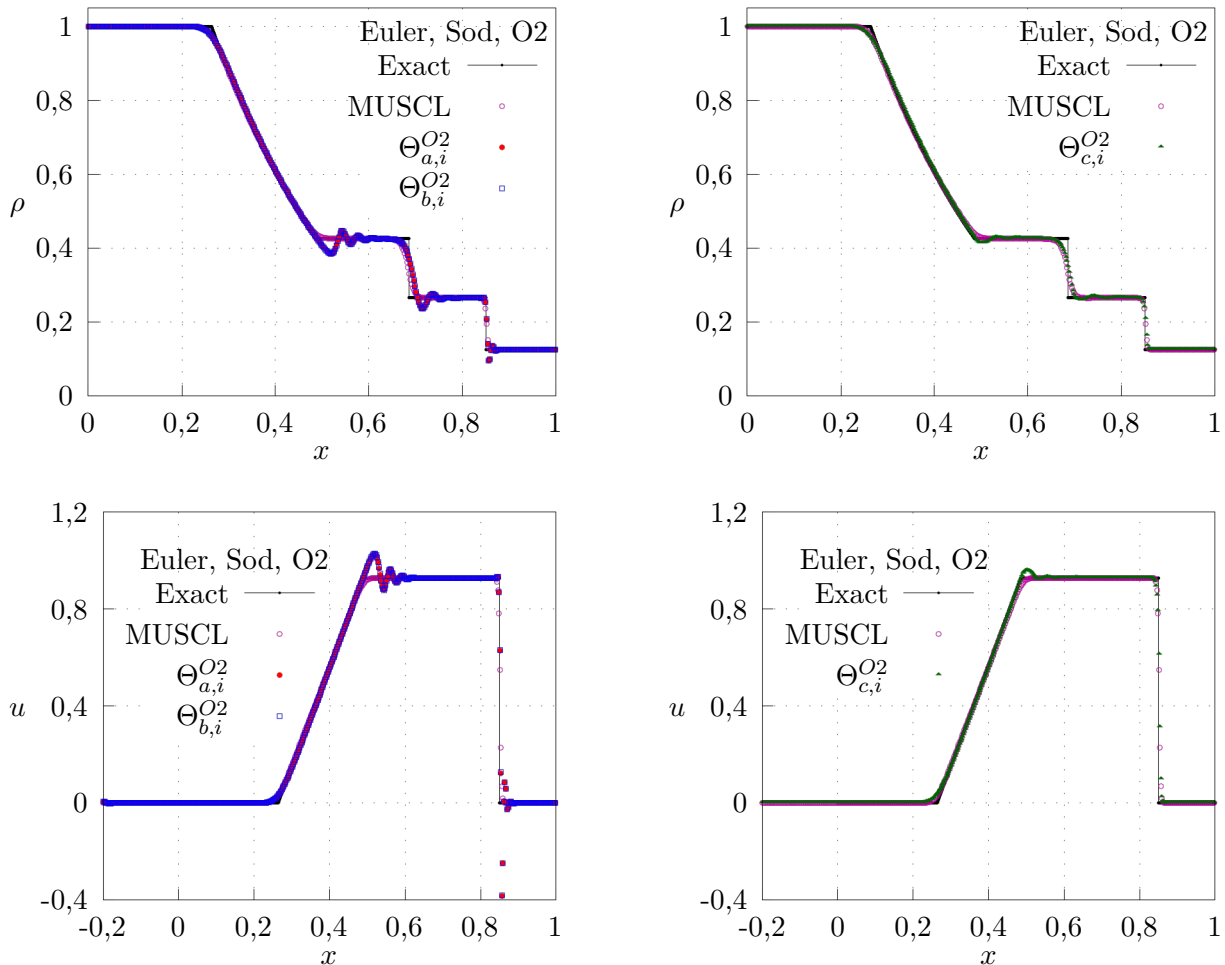
TABLE 4.3 – Évaluations des erreurs et des ordres numériques produits par le schéma du second ordre temps et espace du Théorème 4.3.4.1 avec les triplets  $((\Theta_{\alpha,i}^{O2})_{\alpha \in \{a,b,c\}}, \lambda, \Delta t)$  donnés dans (4.87)-(4.80) et à partir de solution régulière des équations d'Euler au temps final  $t = 0,2$ .

Les résultats sont très satisfaisants et le second ordre est exactement observé alors que, pour le même cas test, les ordres spatiaux du schéma MUSCL sont : 1,88 pour la norme  $L^1$ , 1,63 pour la norme  $L^2$  et 1,27 pour la norme  $L^\infty$ .

Le second cas test est le problème du tube à chocs de Sod [198]. Dans ce cas test, le domaine spatial  $[0, 1]$  est discrétisé avec 400 cellules et la condition initiale est

$$\rho_0(x) = \begin{cases} 1 & \text{si } x < 0,5, \\ 0,125 & \text{sinon,} \end{cases} \quad u_0(x) = 0, \quad p_0(x) = \begin{cases} 1 & \text{si } x < 0,5, \\ 0,1 & \text{sinon.} \end{cases} \quad (4.89)$$

Pour imposer des conditions de bord périodiques, le problème de Sod ci-dessus est symétrisé sur le domaine  $[-1, 0]$ . Le temps final de simulation est  $t = 0,2$ . La Figure 4.4 montre les résultats comparés et les tableaux 4.4 rapportent les erreurs numériques ainsi que les ordres pour différents maillages.



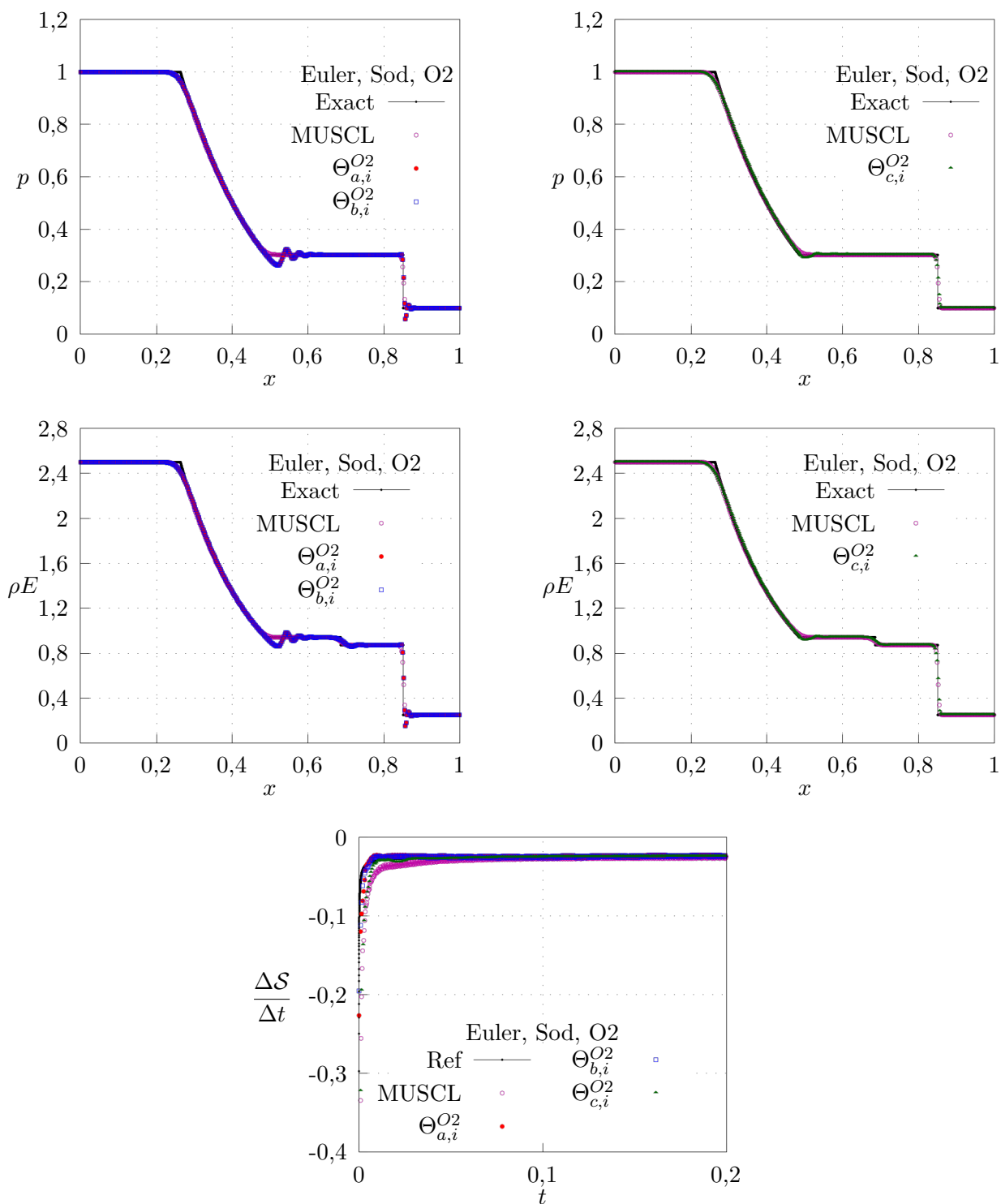


FIGURE 4.4 – Résultats numériques du schéma du second ordre temps et espace du Théorème 4.3.4.1 sur un maillage de 400 cellules, avec les triplets  $((\Theta_{\alpha,i}^{O2})_{\alpha \in \{a,b,c\}}, \lambda, \Delta t)$  donnés dans (4.87)-(4.80) et pour le problème de Sod (4.89) au temps final  $t = 0,2$ . En bas, évolution temporelle de  $\Delta S / \Delta t$  défini dans (4.81).

D'après la Figure 4.4, les triplets  $((\Theta_{\alpha,i}^{O2})_{\alpha \in \{a,b,c\}}, \lambda, \Delta t)$  donnés dans (4.87)-(4.80) assurent des valeurs de  $\Delta S / \Delta t$  négatives. Par conséquent, la stabilité entropique discrète globale est vérifiée. En particulier, le choix  $\Theta_{c,i}^{O2}$  (4.87c) donne une nouvelle fois des résultats remarquables avec peu d'oscillations parasites.

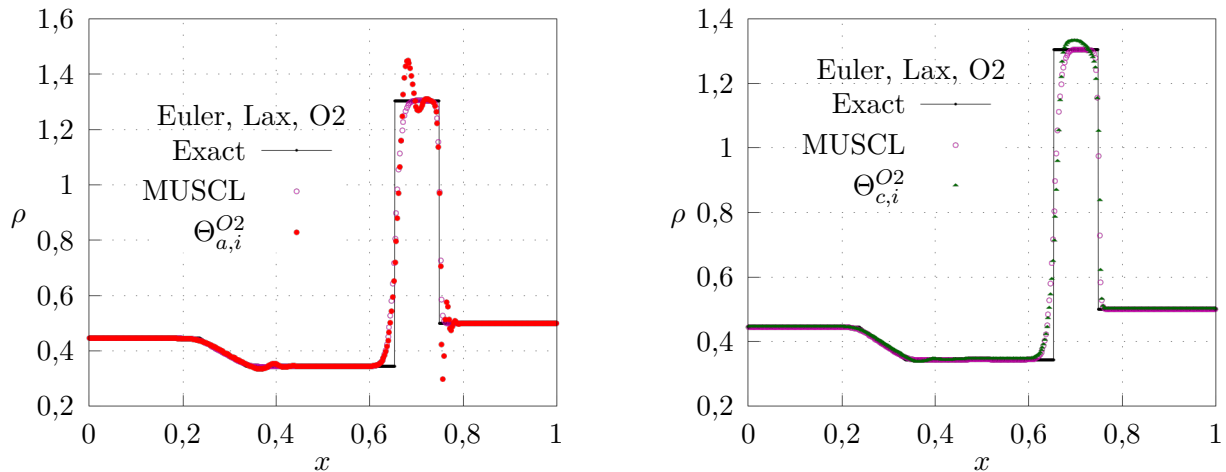
Erreurs du schéma du second ordre avec $\Theta_i^n = \Theta_{a,i}^{O2}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	1,9E-02	-	2,8E-02	-	9,5E-02	-
200	1,1E-02	0,8	1,9E-02	0,6	8,6E-02	0,1
400	6,3E-03	0,8	1,3E-02	0,5	9,2E-02	0,1
800	3,5E-03	0,8	9,1E-03	0,5	9,7E-02	0,1
1600	2,0E-03	0,8	7,0E-03	0,4	1,0E-01	0,1
Erreurs du schéma du second ordre avec $\Theta_i^n = \Theta_{b,i}^{O2}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	1,9E-02	-	2,8E-02	-	9,6E-02	-
200	1,1E-02	0,8	1,9E-02	0,6	8,7E-02	0,1
400	6,3E-03	0,8	1,3E-02	0,5	9,2E-02	0,1
800	3,5E-03	0,8	9,2E-03	0,5	9,7E-02	0,1
1600	2,0E-03	0,8	7,0E-03	0,4	1,0E-01	0,1
Erreurs du schéma du second ordre avec $\Theta_i^n = \Theta_{c,i}^{O2}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	1,5E-02	-	2,2E-02	-	7,9E-02	-
200	8,2E-03	0,8	1,5E-02	0,6	6,7E-02	0,2
400	4,3E-03	0,9	9,9E-03	0,6	6,9E-02	0,0
800	2,3E-03	0,9	6,9E-03	0,5	7,6E-02	0,1
1600	1,3E-03	0,9	5,4E-03	0,4	8,4E-02	0,1

TABLE 4.4 – Évaluations des erreurs et des ordres numériques produits par le schéma du second ordre temps et espace du Théorème 4.3.4.1 avec les triplets  $((\Theta_{\alpha,i}^{O2})_{\alpha \in \{a,b,c\}}, \lambda, \Delta t)$  donnés dans (4.87)-(4.80) et pour le problème de Sod (4.89) au temps final  $t = 0,2$ .

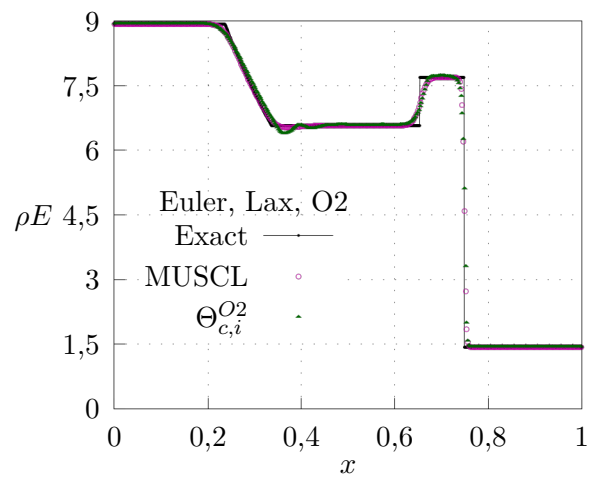
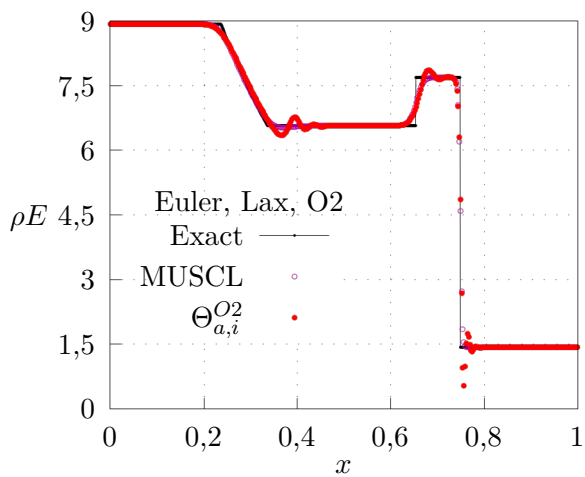
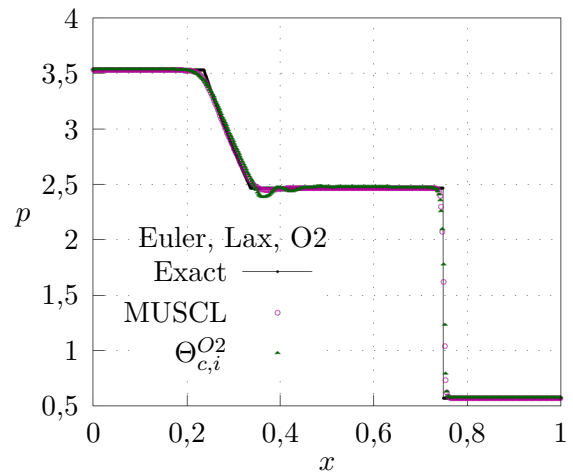
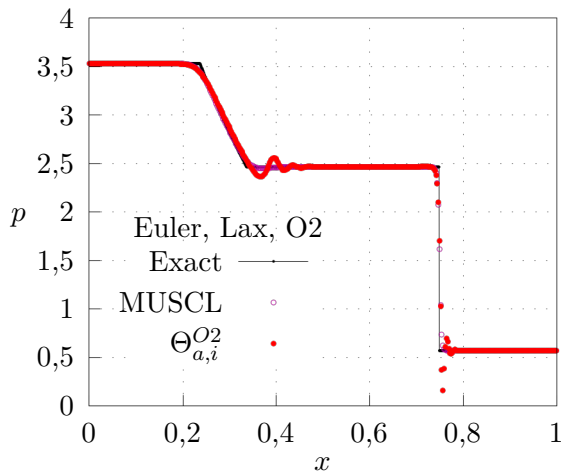
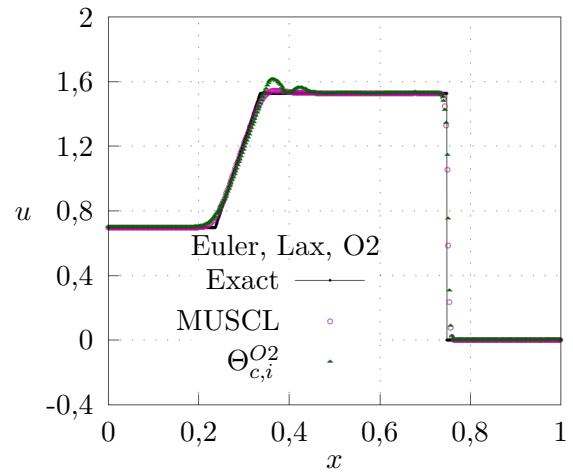
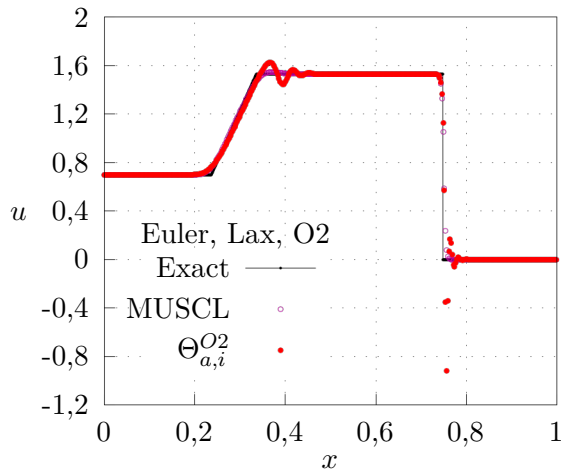
Le troisième cas test concerne le problème du tube à chocs de Lax [141]. Pour ce cas test, le domaine spatial  $[0, 1]$  est discrétisé avec 400 cellules et la condition initiale est

$$\rho_0(x) = \begin{cases} 0,445 & \text{si } x < 0,5, \\ 0,5 & \text{sinon,} \end{cases} \quad u_0(x) = \begin{cases} 0,698 & \text{si } x < 0,5, \\ 0 & \text{sinon,} \end{cases} \quad p_0(x) = \begin{cases} 3,528 & \text{si } x < 0,5, \\ 0,5710 & \text{sinon,} \end{cases} \quad (4.90)$$

Pour imposer une nouvelle fois des conditions de bord périodiques, le problème de Lax ci-dessus est symétrisé sur le domaine  $[-1, 0]$ . Le temps final de la simulation est  $t = 0,1$ . La Figure 4.5 montre les résultats comparés et les tableaux 4.5 rapportent les erreurs numériques ainsi que les ordres pour différents maillages.







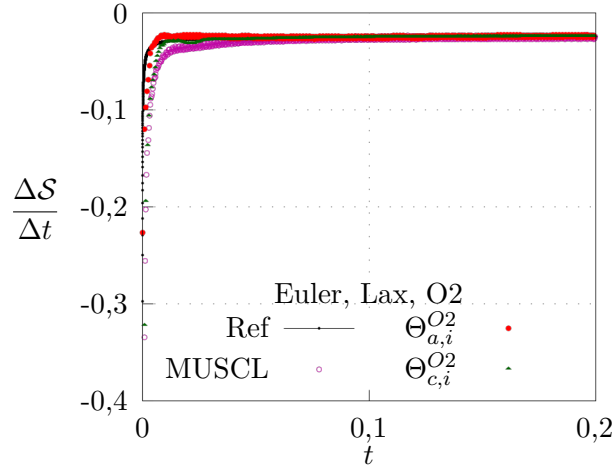


FIGURE 4.5 – Résultats numériques du schéma du second ordre temps et espace du Théorème 4.3.4.1 sur un maillage de 400 cellules, avec les triplets  $((\Theta_{\alpha,i}^{O2})_{\alpha \in \{a,c\}}, \lambda, \Delta t)$  donnés dans (4.87)-(4.80) et pour le problème de Lax (4.90) au temps final  $t = 0,1$ . En bas, évolution temporelle de  $\Delta S/\Delta t$  donné par (4.81).

Pour le cas test de Lax, le choix  $\Theta_{b,i}^{O2}$  donné dans (4.87b) n'assure pas la stabilité entropique globale du schéma proposé. En revanche, le choix  $\Theta_{c,i}^{O2}$  défini dans (4.87c) donne des résultats avec peu d'oscillations parasites.

Erreurs du schéma du second ordre avec $\Theta_i^n = \Theta_{a,i}^{O2}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	4,0E-02	-	1,0E-01	-	5,7E-01	-
200	2,9E-02	0,5	9,2E-02	0,1	6,1E-01	0,1
400	1,7E-02	0,7	6,6E-02	0,5	5,9E-01	0,1
800	1,1E-02	0,7	5,3E-02	0,3	6,0E-01	0,0
1600	6,8E-03	0,7	4,4E-02	0,3	6,2E-01	0,0
Erreurs du schéma du second ordre avec $\Theta_i^n = \Theta_{c,i}^{O2}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	4,9E-02	-	1,3E-01	-	6,3E-01	-
200	2,7E-02	0,9	9,7E-02	0,4	6,6E-01	0,1
400	1,5E-02	0,8	6,6E-02	0,5	5,9E-01	0,2
800	8,9E-03	0,8	5,1E-02	0,4	6,0E-01	0,0
1600	5,2E-03	0,8	4,0E-02	0,3	6,2E-01	0,0

TABLE 4.5 – Évaluations des erreurs et des ordres numériques produits par le schéma du second ordre temps et espace du Théorème 4.3.4.1 avec les triplets  $((\Theta_{\alpha,i}^{O2})_{\alpha \in \{a,c\}}, \lambda, \Delta t)$  donnés dans (4.87)-(4.80) et pour le problème de Lax (4.90) au temps final  $t = 0,1$ .

Le dernier cas test concerne les solutions des équations d'Euler *proches* du bord de  $\Omega$  caractérisé par  $\rho$  et  $p$  nuls. Le domaine spatial  $[0, 1]$  est discrétisé avec 400 cellules et la condition initiale est donnée par

$$\rho_0(x) = 1, \quad u_0(x) = \begin{cases} -2,7 & \text{si } x < 0,5, \\ 2,7 & \text{sinon,} \end{cases} \quad p_0(x) = 1. \quad (4.91)$$

La condition initiale ci-dessus est symétrisée sur le domaine  $[-1, 0]$  et des conditions de bord périodiques sont imposées. La solution exacte se compose de deux ondes de détente qui entraînent les variables  $\rho$ ,  $\rho E$  et  $p$  vers zéro. Pour ce cas test  $\lambda > 0$  et  $\Delta t > 0$  vérifient

$$\lambda = 1,05 \max_i (u_i^n \pm \sqrt{\gamma p(w_i^n)/\rho_i^n}, u_i^n), \quad \Delta t = \frac{\Delta x}{20\lambda}. \quad (4.92)$$

Le temps final de simulation est  $t = 0,05$ . La Figure 4.6 montre les résultats comparés et les tableaux 4.6 rapportent les erreurs numériques ainsi que les ordres pour différents maillages.

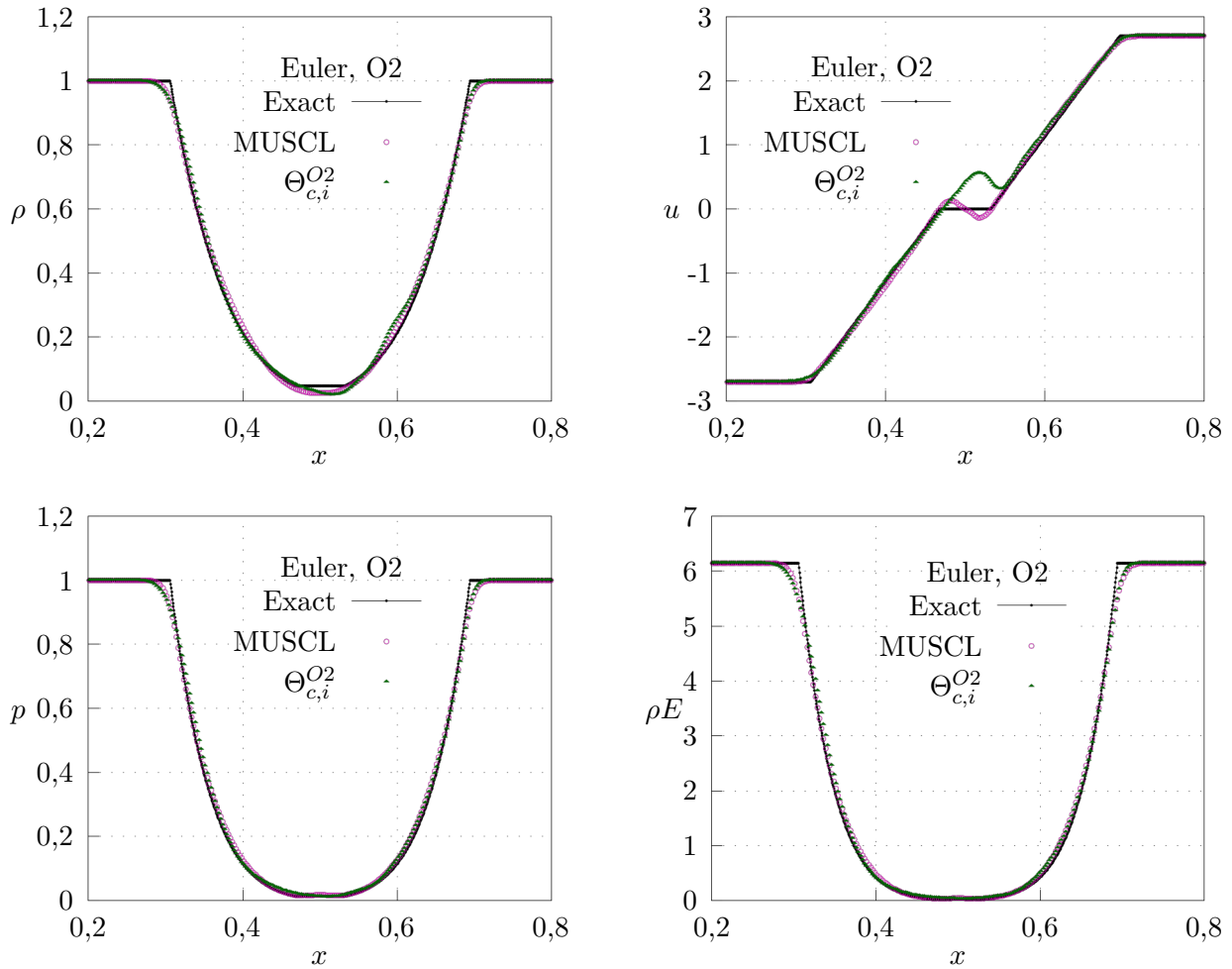


FIGURE 4.6 – Résultats numériques du schéma du second ordre temps et espace du Théorème 4.3.4.1 sur un maillage de 400 cellules, avec le triplet  $(\Theta_{c,i}^{O2}, \lambda, \Delta t)$  donné dans (4.87c)-(4.92) et pour le problème de la double détente (4.91) au temps final  $t = 0,05$ .

Pour le problème de la double détente (4.91), les matrices  $(\Theta_{a,i}^{O2})_{i \in \mathbb{Z}}$  et  $(\Theta_{b,i}^{O2})_{i \in \mathbb{Z}}$  respectivement données dans (4.87a)-(4.87b) génèrent des oscillations parasites qui conduisent à des valeurs de  $p$  négatives. Par conséquent, ces deux choix ne sont pas fonctionnels pour ce cas test. Les matrices  $\Theta_{c,i}^{O2}$  (4.87c) confirment leurs robustesses puisqu'au temps final de simulation  $t = 0,05$ , la quantité  $p$  demeure positive et sa valeur minimale est de l'ordre de  $10^{-3}$ .

cellule	Erreurs du schéma du second ordre avec $\Theta_i^n = \Theta_{c,i}^{O2}$					
	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	3,1E-02	-	5,4E-02	-	2,0E-01	-
200	2,5E-02	0,3	4,7E-02	0,2	2,1E-01	0,0
400	1,4E-02	0,8	2,9E-02	0,7	1,4E-01	0,6
800	7,9E-03	0,9	1,6E-02	0,9	7,3E-02	0,9
1600	4,1E-03	1,0	8,2E-03	0,9	4,3E-02	0,8

TABLE 4.6 – Évaluations des erreurs et des ordres numériques produits par le schéma du second ordre temps et espace du Théorème 4.3.4.1 avec le triplet  $(\Theta_{c,i}^{O2}, \lambda, \Delta t)$  donné dans (4.87c)-(4.92) et pour le problème de la double détente (4.91) au temps final  $t = 0,05$ .

### 4.5.2.3 Le cas des fonctions de flux non convexes

Cette section aborde le cas d'une fonction de flux scalaire  $f$  non convexe. Dans ce cas, et d'après [143], l'association d'une seule inégalité d'entropie à l'équation de conservation n'assure pas l'unicité des solutions. Par conséquent, des solutions constituées d'ondes de chocs *non classiques* peuvent apparaître. Pour illustrer ce phénomène, le cas test proposé dans [147, Section 3.1] est maintenant reproduit. En considérant l'entropie quadratique  $\eta(w) = w^2/2$ , le problème de Cauchy dans  $\mathbb{R}$  considéré pour ce cas test est

$$\begin{cases} \partial_t w + \partial_x w^3 = 0, & t > 0, \quad x \in ]-1, 1[, \\ w(x, t = 0) = -\sin(\pi x). \end{cases} \quad (4.93)$$

Le problème ci-dessus est associé à des conditions de bord de périodiques aux deux frontières du domaine  $[-1, 1]$ . La Figure 4.7 montre les résultats comparés au temps final de simulation  $t = 1$ .

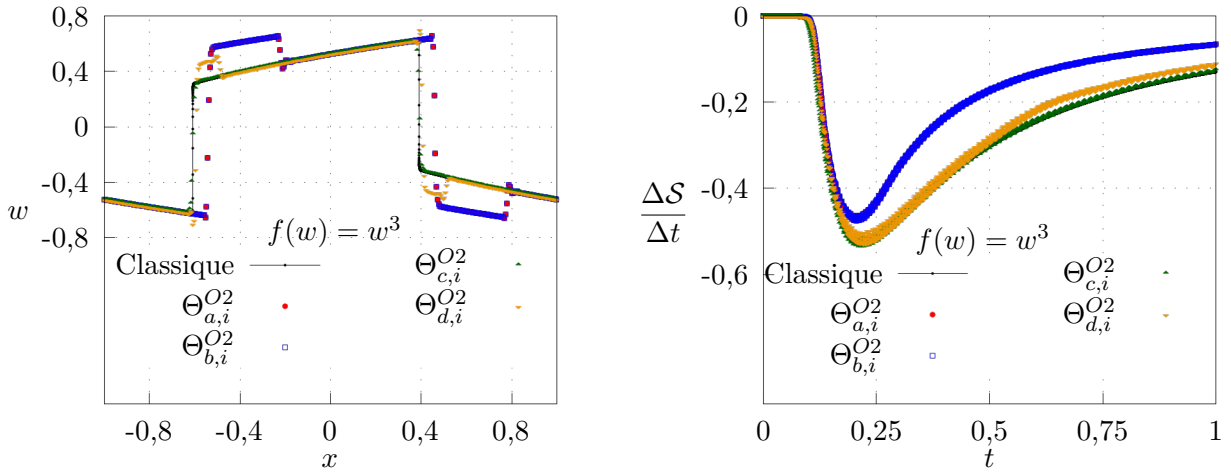


FIGURE 4.7 – À gauche, résultats numériques produits par le schéma du second ordre temps et espace du Théorème 4.3.4.1 sur un maillage de 400 cellules, avec les triplets  $((\Theta_{\alpha,i}^{O2})_{\alpha \in \{a,b,c,d\}}, \lambda, \Delta t)$  donnés dans (4.83)-(4.80) et pour les solutions du problème (4.93) au temps final  $t = 1$ . À droite, évolution temporelle du critère  $\Delta S/\Delta t$  défini dans (4.81).

D'après la Figure 4.7, si les choix des paramètres  $((\Theta_{\alpha,i}^{O2})_{\alpha \in \{a,b,c,d\}}, \lambda, \Delta t)$  donnés dans (4.83)-(4.80) assurent tous la stabilité entropique discrète globale, les solutions numériques calculées diffèrent. Le choix de  $(\Theta_{c,i}^{O2})_{i \in \mathbb{Z}}$  (4.83c) donne une solution *classique* alors que pour les choix  $(\Theta_{\alpha,i}^{O2})_{\alpha \in \{a,b,d\}}$ , le schéma du second ordre proposé converge vers une solution *non classique*. Cette observation est conforme aux résultats récents de LeFloch *et al* qui établissent dans [147] qu'un schéma numérique d'ordre élevé peut arbitrairement converger vers une solution non classique dans le cas d'une fonction de flux non convexe.

## 4.5.3 Schémas numériques du troisième et du quatrième ordre

### 4.5.3.1 L'équation de Burgers scalaire

Cette section dresse les résultats numériques des schémas d'ordre trois et quatre du Théorème 4.4.3.1 et pour le cas de l'équation de Burgers scalaire telle que rappelée en Section 4.5.2.1. Les résultats des deux schémas sont présentés conjointement et la notation  $Ok$  avec  $k$  dans  $\{3, 4\}$  désigne le schéma du même ordre.

En considérant un réel  $\theta^{Ok}$  satisfaisant l'inégalité stricte (4.77) puis les quantités  $\mathcal{A}_i^{Ok}$  définies dans

(4.76) et  $\varepsilon = 10^{-12}$ , les réels  $\Theta_i^{Ok}$  proposés pour l'équation de Burgers sont

$$\Theta_{a,i}^{Ok} = -\theta^{Ok} \operatorname{sign} \left( (\nabla \eta(w_{i+1}^n) - \nabla \eta(w_{i-1}^n)) (\mathcal{A}_i^{Ok}) \right), \quad (4.94a)$$

$$\Theta_{b,i}^{O4} = \frac{((\delta_{i-\frac{1}{2}}^n)^2 - (\delta_{i+\frac{1}{2}}^n)^2)((\delta_{i-\frac{1}{2}}^n)^2 + (\delta_{i+\frac{1}{2}}^n)^2)}{((\delta_{i-\frac{1}{2}}^n)^2 + (\delta_{i+\frac{1}{2}}^n)^2)^2 + \varepsilon}, \quad (4.94b)$$

$$\Theta_{c,i}^{Ok} = \frac{1}{2}. \quad (4.94c)$$

En utilisant ces paramétrages, les cas tests réalisés en Section 4.5.2.1 pour le schéma du second ordre sont maintenant reproduits avec les schémas d'ordre  $k$ . Les définitions des cas tests données en Section 4.5.2.1 demeurent mais comme des montées ordre temporel sont effectuées pour les schémas d'ordre  $k$ , les paramètres  $\lambda > 0$  et  $\Delta t > 0$  sont sélectionnés selon les restrictions usuelles (4.80).

Le premier cas test mesure l'ordre des schémas à partir d'une solution régulière obtenue avec la condition initiale (4.84). La Figure 4.8 montre les résultats au temps final de simulation  $t = 0,3$  et les tableaux 4.7-4.8-4.9 rapportent les erreurs numériques ainsi que les ordres pour différents maillages.

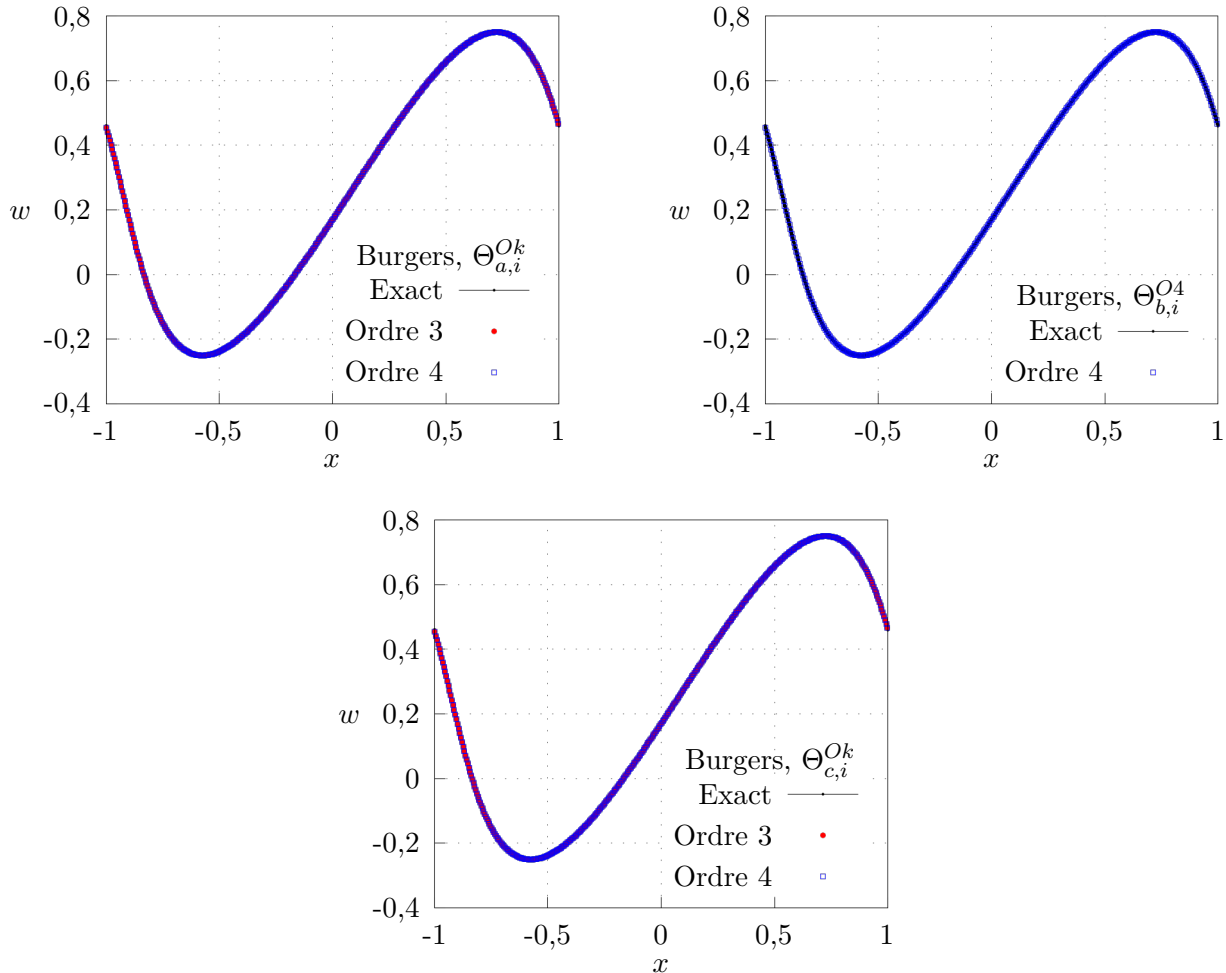


FIGURE 4.8 – Résultats numériques produits par les schémas du troisième et du quatrième ordre temps et espace du Théorème 4.4.3.1 sur un maillage de 400 cellules, avec les triplets  $((\Theta_{\alpha,i}^{Ok})_{\alpha \in \{a,b,c\}}, \lambda, \Delta t)$  donnés dans (4.94)-(4.80) et pour le problème (4.84) au temps final  $t = 0,3$ .

Erreurs du schéma du troisième ordre avec $\Theta_i^n = \Theta_{a,i}^{O3}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	8,2E-05	-	1,3E-04	-	3,4E-05	-
200	1,0E-05	3,0	1,6E-05	3,0	2,4E-06	3,8
400	1,3E-06	3,0	2,0E-06	3,0	1,9E-07	3,7
800	1,6E-07	3,0	2,5E-07	3,0	1,6E-08	3,6
1600	2,0E-08	3,0	3,1E-08	3,0	1,4E-09	3,6
Erreurs du schéma du quatrième ordre avec $\Theta_i^n = \Theta_{a,i}^{O4}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	6,1E-06	-	1,0E-05	-	1,2E-05	-
200	3,8E-07	4,0	6,4E-07	4,0	7,9E-07	4,0
400	2,4E-08	4,0	4,0E-08	4,0	4,9E-08	4,0
800	1,5E-09	4,0	2,5E-09	4,0	3,0E-09	4,0
1600	9,1E-11	4,0	1,5E-10	4,0	1,9E-10	4,0

TABLE 4.7 – Évaluations des erreurs et des ordres numériques produits par les schémas du troisième et du quatrième ordre temps et espace du Théorème 4.4.3.1 avec le triplet  $(\Theta_{a,i}^{O_k}, \lambda, \Delta t)$  donné dans (4.94a)-(4.80) et pour le problème (4.84) au temps final  $t = 0,3$ .

Erreurs du schéma du quatrième ordre avec $\Theta_i^n = \Theta_{b,i}^{O4}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	6,0E-06	-	9,8E-06	-	1,2E-05	-
200	3,7E-07	4,0	6,3E-07	4,0	7,6E-07	3,9
400	2,3E-08	4,0	3,9E-08	4,0	4,8E-08	4,0
800	1,5E-09	4,0	2,5E-09	4,0	3,0E-09	4,0
1600	9,1E-11	4,0	1,5E-10	4,0	1,9E-10	4,0

TABLE 4.8 – Évaluations des erreurs et des ordres numériques produits par le schéma du quatrième ordre temps et espace du Théorème 4.4.3.1 avec le triplet  $(\Theta_{b,i}^{O4}, \lambda, \Delta t)$  donné dans (4.94b)-(4.80) et pour le problème (4.84) au temps final  $t = 0,3$ .

Erreurs du schéma du troisième ordre avec $\Theta_i^n = \Theta_{c,i}^{O3}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	8,2E-05	-	1,3E-04	-	1,3E-05	-
200	1,0E-05	3,0	1,6E-05	3,0	1,0E-06	3,7
400	1,3E-06	3,0	2,0E-06	3,0	8,7E-08	3,6
800	1,6E-07	3,0	2,5E-07	3,0	7,6E-09	3,5
1600	2,0E-08	3,0	3,1E-08	3,0	6,8E-10	3,5
Erreurs du schéma du quatrième ordre avec $\Theta_i^n = \Theta_{c,i}^{O4}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	3,5E-06	-	5,9E-06	-	1,8E-06	-
200	2,0E-07	4,1	3,3E-07	4,1	1,1E-07	4,0
400	1,2E-08	4,1	1,9E-08	4,1	6,6E-09	4,1
800	7,2E-10	4,0	1,2E-09	4,1	4,0E-10	4,0
1600	4,5E-11	4,0	7,1E-11	4,0	2,2E-11	4,2

TABLE 4.9 – Évaluations des erreurs et des ordres numériques produits par les schémas du troisième et du quatrième ordre temps et espace du Théorème 4.4.3.1 avec le triplet  $(\Theta_{c,i}^{O_k}, \lambda, \Delta t)$  donné dans (4.94c)-(4.80) et pour le problème (4.84) au temps final  $t = 0,3$ .

D'après les tableaux 4.7-4.8-4.9, les ordres spatiaux attendus sont exactement retrouvés dans les expériences numériques.

Le second cas test se réfère à une solution discontinue de l'équation de Burgers telle qu'obtenue avec la condition initiale (4.85). La Figure 4.9 montre les résultats au temps final de simulation  $t = 0,3$  et les tableaux 4.10-4.11-4.12 rapportent les erreurs numériques ainsi que les ordres pour différents maillages.

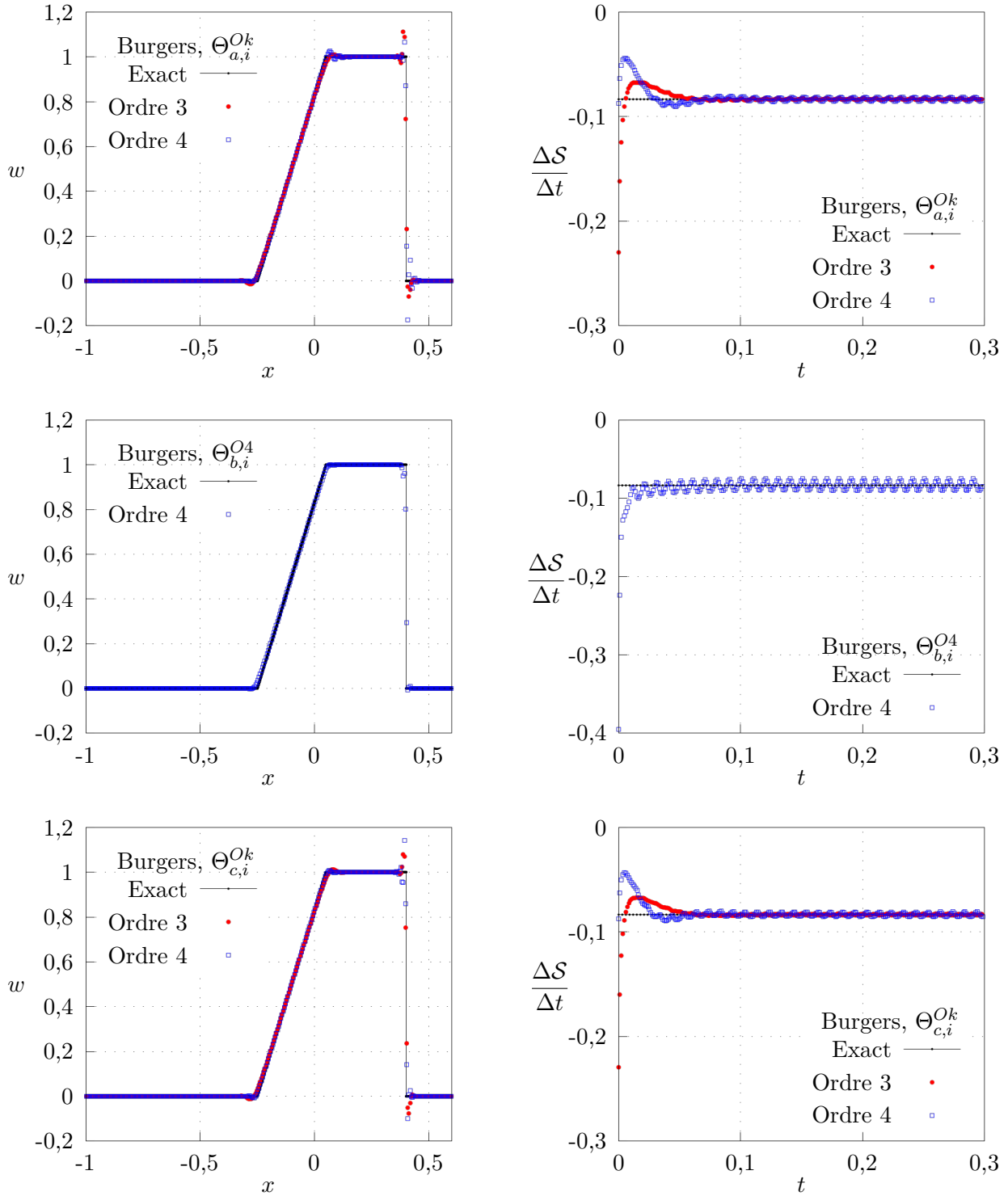


FIGURE 4.9 – À gauche, résultats numériques produits par les schémas du troisième et du quatrième ordre temps et espace du Théorème 4.4.3.1 sur un maillage de 400 cellules, avec les triplets  $((\Theta_{\alpha,i}^{Ok})_{\alpha \in \{a,b,c\}}, \lambda, \Delta t)$  donnés dans (4.94)-(4.80) et pour le problème (4.85) au temps final  $t = 0,3$ . À droite, évolution temporelle de  $\Delta S/\Delta t$  défini par (4.81).

D'après les graphiques à droite de la Figure 4.9, la quantité  $\Delta S/\Delta t$  définie par (4.81) est négative

pour tous les triplets  $((\Theta_{\alpha,i}^{Ok})_{\alpha \in \{a,b,c\}}, \lambda, \Delta t)$  donnés dans (4.94)-(4.80). Par conséquent, la stabilité entropique discrète globale est satisfaite. Le choix particulier  $\Theta_{b,i}^{O4}$  défini par (4.94b) donne un résultat très satisfaisant avec très peu d'oscillation parasite autour de la discontinuité. Ce choix particulier du paramètre  $\Theta_i$  donne également des résultats satisfaisants pour le schéma du second ordre exposé en Section 4.5.2.1 mais il n'est pas fonctionnel pour le schéma du troisième ordre.

Erreurs du schéma du troisième ordre avec $\Theta_i^n = \Theta_{a,i}^{O3}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	3,6E-02	-	6,3E-02	-	5,3E-02	-
200	1,7E-02	1,1	4,2E-02	0,6	4,2E-02	0,3
400	8,4E-03	1,0	2,9E-02	0,5	3,3E-02	0,3
800	4,2E-03	1,0	2,0E-02	0,5	2,6E-02	0,3
1600	2,1E-03	1,0	1,4E-02	0,5	2,1E-02	0,3
Erreurs du schéma du quatrième ordre avec $\Theta_i^n = \Theta_{a,i}^{O4}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	2,6E-02	-	4,5E-02	-	3,7E-02	-
200	1,2E-02	1,1	3,0E-02	0,6	2,9E-02	0,3
400	5,7E-03	1,0	2,1E-02	0,5	2,4E-02	0,3
800	2,8E-03	1,0	1,5E-02	0,5	1,9E-02	0,3
1600	1,4E-03	1,0	1,1E-02	0,5	1,5E-02	0,3

TABLE 4.10 – Évaluations des erreurs et des ordres numériques produits par les schémas du troisième et du quatrième ordre temps et espace du Théorème 4.4.3.1 avec le triplet  $(\Theta_{a,i}^{Ok}, \lambda, \Delta t)$  donné dans (4.94a)-(4.80) et pour le problème (4.85) au temps final  $t = 0,3$ .

Erreurs du schéma du quatrième ordre avec $\Theta_i^n = \Theta_{b,i}^{O4}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	2,3E-02	-	5,5E-02	-	7,3E-02	-
200	1,3E-02	0,8	3,9E-02	0,5	5,6E-02	0,4
400	6,5E-03	1,0	2,7E-02	0,5	4,4E-02	0,3
800	3,3E-03	1,0	1,9E-02	0,5	3,5E-02	0,3
1600	1,6E-03	1,0	1,3E-02	0,5	2,8E-02	0,3

TABLE 4.11 – Évaluations des erreurs et des ordres numériques produits par le schéma du quatrième ordre temps et espace du Théorème 4.4.3.1 avec le triplet  $(\Theta_{b,i}^{O4}, \lambda, \Delta t)$  donné dans (4.94b)-(4.80) et pour le problème (4.85) au temps final  $t = 0,3$ .

Erreurs du schéma du troisième ordre avec $\Theta_i^n = \Theta_{c,i}^{O3}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	3,4E-02	-	5,9E-02	-	3,2E-02	-
200	1,6E-02	1,1	3,9E-02	0,6	2,6E-02	0,3
400	7,8E-03	1,0	2,7E-02	0,5	2,0E-02	0,3
800	3,9E-03	1,0	1,9E-02	0,5	1,6E-02	0,3
1600	1,9E-03	1,0	1,3E-02	0,5	1,3E-02	0,3



Erreurs du schéma du quatrième ordre avec $\Theta_i^n = \Theta_{c,i}^{O4}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	2,3E-02	-	4,2E-02	-	3,1E-02	-
200	9,7E-03	1,2	2,8E-02	0,6	2,5E-02	0,3
400	4,8E-03	1,0	2,0E-02	0,5	2,0E-02	0,3
800	2,4E-03	1,0	1,4E-02	0,5	1,6E-02	0,3
1600	1,2E-03	1,0	9,7E-03	0,5	1,3E-02	0,3

TABLE 4.12 – Évaluations des erreurs et des ordres numériques produits par les schémas du troisième et du quatrième ordre temps et espace du Théorème 4.4.3.1 avec le triplet  $(\Theta_{c,i}^{Ok}, \lambda, \Delta t)$  donné dans (4.94c)-(4.80) et pour le problème (4.85) au temps final  $t = 0,3$ .

### 4.5.3.2 Les équations d'Euler

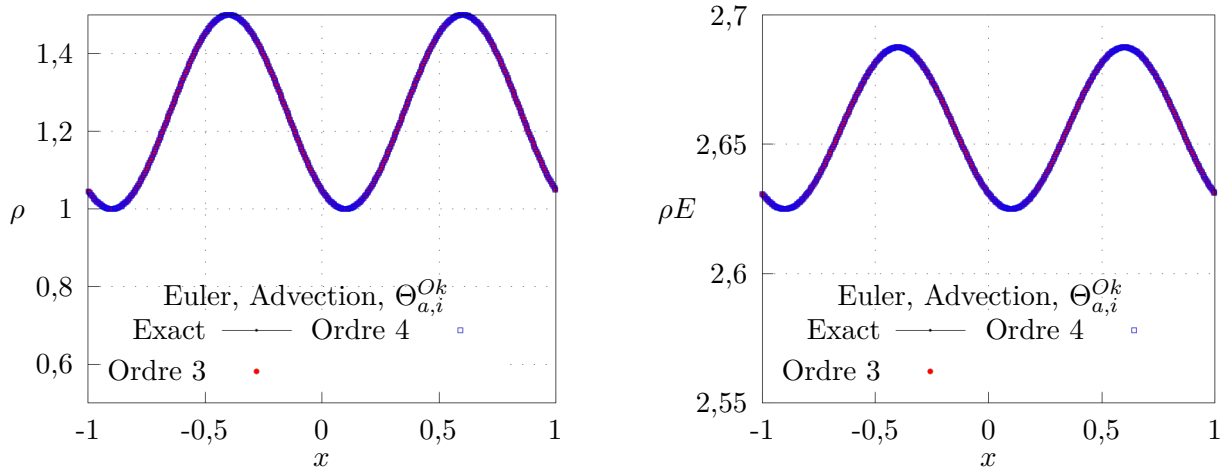
Cette section est consacrée aux résultats numériques des schémas d'ordre trois et quatre en espace et en temps du Théorème 4.4.3.1 appliqués aux équations d'Euler telle que décrites en Section 4.5.2.2. Ces résultats se composent de deux simulations numériques qui reproduisent pour les schémas d'ordre trois et quatre deux cas tests définis en Section 4.5.2.2. Les deux cas test choisis sont le transport linéaire (4.88) et le tube à chocs de Sod [198]. Pour chaque cas test, les résultats sont présentés conjointement et la notation  $Ok$  pour  $k$  dans  $\{3, 4\}$  désigne le schéma du même ordre.

En considérant un réel  $\theta^{Ok}$  satisfaisant l'inégalité stricte (4.77) ainsi que les vecteurs  $\mathcal{A}_i^{Ok}$  définis dans (4.76) et  $\varepsilon = 10^{-12}$ , les matrices  $\Theta_i^{Ok}$  retenues pour les équations d'Euler sont

$$\Theta_{a,i}^{Ok} = -\theta^{Ok} \text{diag}_{1 \leq j \leq 3} \left( \text{sign} \left( (\nabla \eta(w_{i+1}^n) - \nabla \eta(w_{i-1}^n))_j (\mathcal{A}_i^{Ok})_j \right) \right), \quad (4.95a)$$

$$\Theta_{b,\varepsilon,i}^{O4} = \text{diag}_{1 \leq j \leq 3} \left( \frac{\left( (\delta_{i-\frac{1}{2}}^n)_j^2 - (\delta_{i+\frac{1}{2}}^n)_j^2 \right) \left( (\delta_{i-\frac{1}{2}}^n)_j^2 + (\delta_{i+\frac{1}{2}}^n)_j^2 \right)}{\left( (\delta_{i-\frac{1}{2}}^n)_j^2 + (\delta_{i+\frac{1}{2}}^n)_j^2 \right)^2 + \varepsilon} \right). \quad (4.95b)$$

Le premier cas test concerne la mesure de l'ordre des schémas à partir d'un transport linéaire obtenu avec la condition initiale (4.88). La Figure 4.10 montre les résultats au temps final de simulation  $t = 0,2$  et les tableaux 4.13-4.14 rapportent les erreurs numériques ainsi que les ordres pour différents maillages.



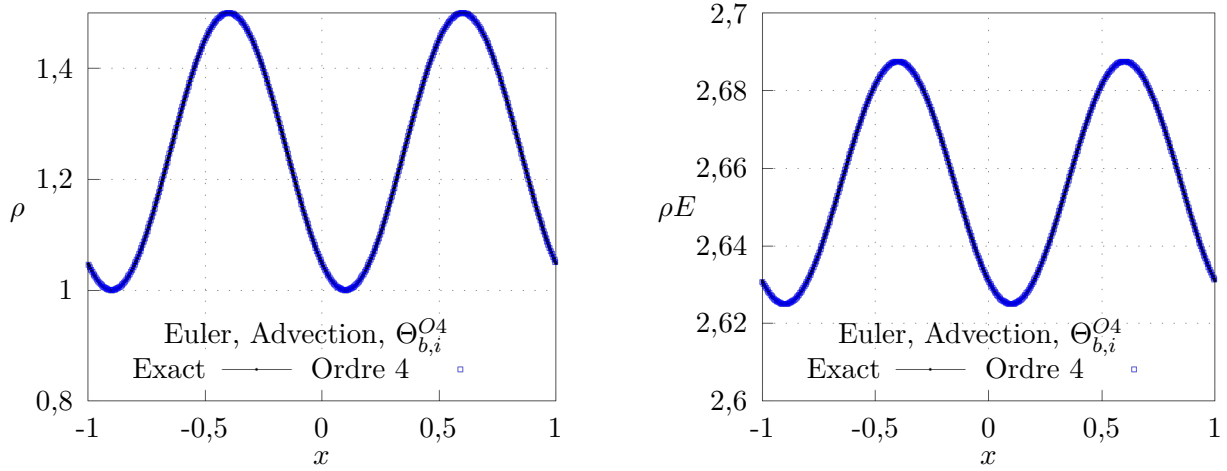


FIGURE 4.10 – Résultats numériques produits par les schémas du troisième et du quatrième ordre temps et espace du Théorème 4.4.3.1 sur un maillage de 400 cellules, avec les triplets  $((\Theta_{\alpha,i}^{O_k})_{\alpha \in \{a,b\}}, \lambda, \Delta t)$  donnés dans (4.95)-(4.80) et pour le problème (4.88) au temps final  $t = 0,2$ .

Erreurs du schéma du troisième ordre avec $\Theta_i^n = \Theta_{a,i}^{O_3}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	6,3E-04	-	3,4E-04	-	3,0E-04	-
200	7,9E-05	3,0	4,3E-05	3,0	3,8E-05	3,0
400	9,9E-06	3,0	5,4E-06	3,0	4,8E-06	3,0
800	1,2E-06	3,0	6,7E-07	3,0	6,0E-07	3,0
1600	1,5E-07	3,0	8,4E-08	3,0	7,5E-08	3,0
Erreurs du schéma du quatrième ordre avec $\Theta_i^n = \Theta_{a,i}^{O_4}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	2,0E-05	-	1,1E-05	-	9,6E-06	-
200	1,3E-06	4,0	6,8E-07	4,0	6,0E-07	4,0
400	7,8E-08	4,0	4,3E-08	4,0	3,8E-08	4,0
800	4,9E-09	4,0	2,7E-09	4,0	2,4E-09	4,0
1600	3,1E-10	4,0	1,7E-10	4,0	1,5E-10	4,0

TABLE 4.13 – Évaluations des erreurs et des ordres numériques produits par les schémas du troisième et du quatrième ordre temps et espace du Théorème 4.4.3.1 avec les triplets  $(\Theta_{a,i}^{O_k}, \lambda, \Delta t)$  donnés dans (4.95a)-(4.80) et pour le problème (4.88) au temps final  $t = 0,2$ .

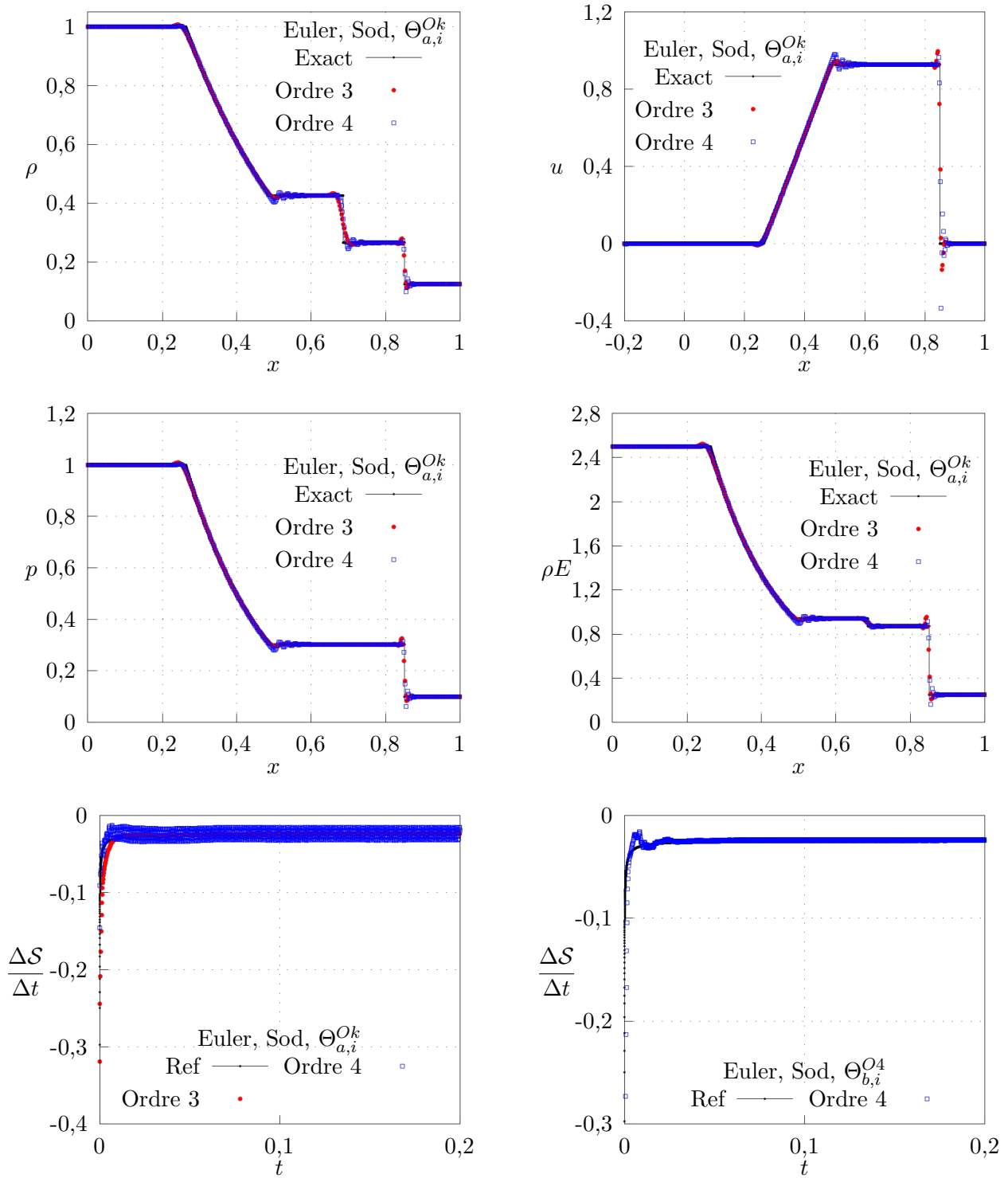
Erreurs du schéma du quatrième ordre avec $\Theta_i^n = \Theta_{b,i}^{O_4}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	5,5E-05	-	3,9E-05	-	7,2E-05	-
200	3,4E-06	4,0	2,7E-06	3,9	6,4E-06	3,5
400	1,2E-07	4,9	6,8E-08	5,3	1,4E-07	5,5
800	5,3E-09	4,4	2,8E-09	4,6	2,4E-09	5,9
1600	3,1E-10	4,1	1,7E-10	4,1	1,5E-10	4,0

TABLE 4.14 – Évaluations des erreurs et des ordres numériques produits par le schéma du quatrième ordre temps et espace du Théorème 4.4.3.1 avec le triplet  $(\Theta_{b,i}^{O_4}, \lambda, \Delta t)$  donné dans (4.95b)-(4.80) et pour le problème (4.88) au temps final  $t = 0,2$ .

D'après les tableaux 4.13-4.14, et puisque les schémas proposés n'utilisent pas de limiteurs de pente, les ordres théoriques attendus sont une nouvelle fois retrouvés dans les expériences numériques.

Pour conclure, le cas test du tube à chocs de Sod [198] est réalisé avec les schémas du troisième et du quatrième ordre temps et espace. Les conditions de ce cas test sont rappelées en Section 4.5.2.2.

La Figure 4.11 montre les résultats au temps final de simulation  $t = 0,2$  et les tableaux 4.15-4.16 rapportent les erreurs numériques ainsi que les ordres pour différents maillages.



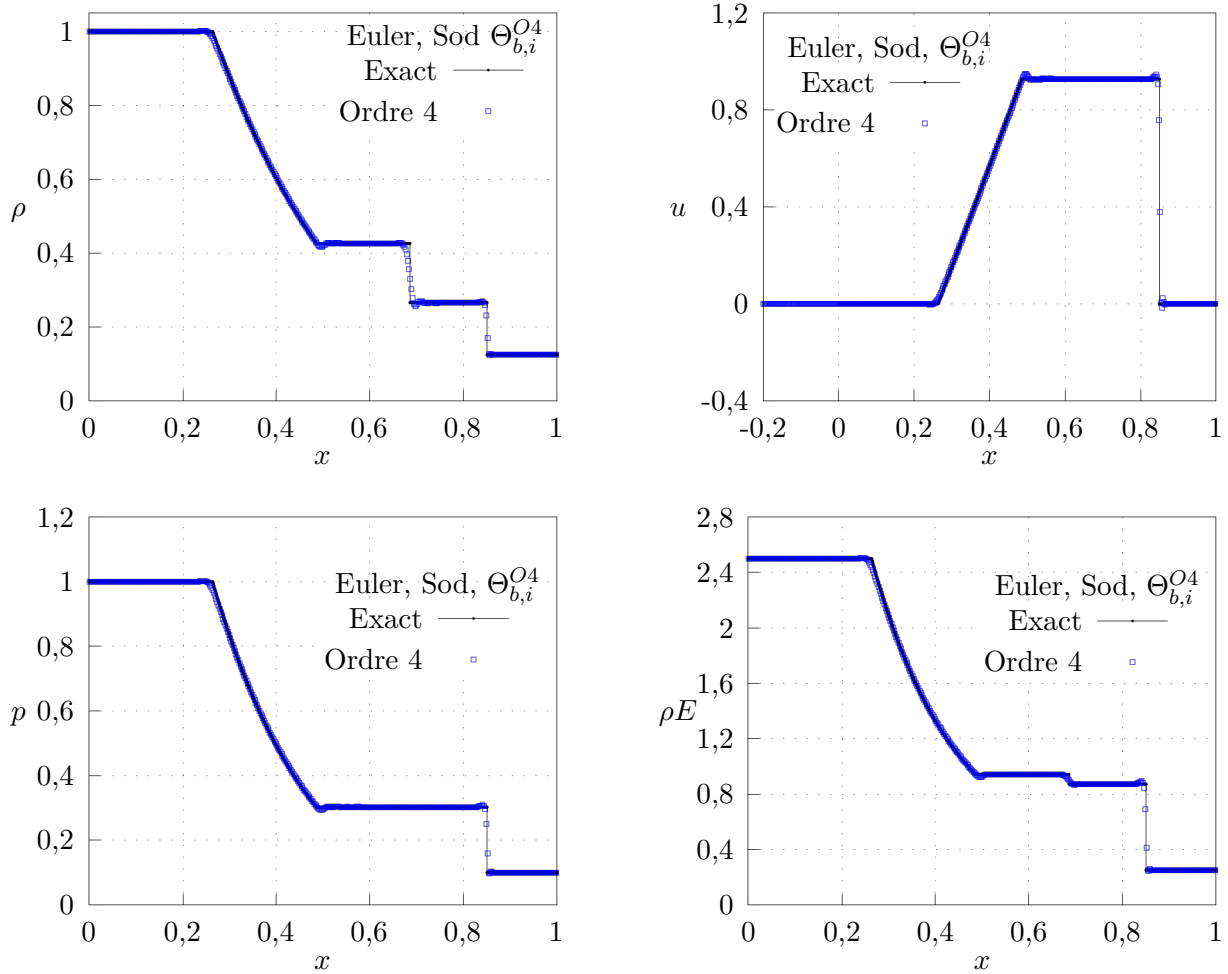


FIGURE 4.11 – Résultats numériques et évolution temporelle de  $\Delta S/\Delta t$  défini par (4.81) avec les schémas du troisième et du quatrième ordre temps et espace du Théorème 4.4.3.1 sur un maillage de 400 cellules, avec les triplets  $((\Theta_{\alpha,i}^{O_k})_{\alpha \in \{a,b\}}, \lambda, \Delta t)$  donnés dans (4.95)-(4.80) et pour le problème de Sod (4.89) au temps final  $t = 0,2$ .

Les paramètres  $((\Theta_{\alpha,i}^{O_k})_{\alpha \in \{a,b\}}, \lambda, \Delta t)$  donnés dans (4.95)-(4.80) assurent la stabilité entropique discrète globale. Les résultats sont corrects pour toutes les matrices  $(\Theta_{\alpha,i}^{O_k})_{\alpha \in \{a,b\}}$  mais ils sont moins oscillants avec le choix  $\Theta_{b,i}^{O4}$  (4.95b) dans le cas du schéma du quatrième ordre. Ce choix de matrices n'est pas possible pour le schéma du troisième ordre car il n'assure pas la stabilité entropique.

cellule	Erreurs du schéma du troisième ordre avec $\Theta_i^n = \Theta_{a,i}^{O3}$					
	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	4,5E-02	-	5,0E-02	-	2,0E-01	-
200	2,3E-02	0,9	3,1E-02	0,7	1,9E-01	0,1
400	1,2E-02	1,0	2,0E-02	0,7	1,7E-01	0,1
800	5,9E-03	1,0	1,2E-02	0,7	1,3E-01	0,4
1600	3,2E-03	0,9	8,5E-03	0,5	1,0E-01	0,4

Erreurs du schéma du quatrième ordre avec $\Theta_i^n = \Theta_{a,i}^{O4}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	3,5E-02	-	3,8E-02	-	1,4E-01	-
200	1,9E-02	0,9	2,4E-02	0,6	1,2E-01	0,3
400	9,9E-03	0,9	1,6E-02	0,6	1,2E-01	0,0
800	4,9E-03	1,0	9,5E-03	0,7	8,1E-02	0,6
1600	2,5E-03	1,0	6,5E-03	0,6	8,5E-02	0,1

TABLE 4.15 – Évaluations des erreurs et des ordres numériques produits par les schémas du troisième et du quatrième ordre temps et espace du Théorème 4.4.3.1 avec le triplet  $(\Theta_{a,i}^{O4}, \lambda, \Delta t)$  donné dans (4.95a)-(4.80) et pour le problème de Sod (4.89) au temps final  $t = 0,2$ .

Erreurs du schéma du quatrième ordre avec $\Theta_i^n = \Theta_{b,i}^{O4}$						
cellule	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	3,2E-02	-	3,9E-02	-	1,7E-01	-
200	1,6E-02	1,0	2,5E-02	0,7	1,6E-01	0,1
400	8,0E-03	1,0	1,5E-02	0,7	1,4E-01	0,2
800	3,8E-03	1,1	8,8E-03	0,8	1,1E-01	0,4
1600	2,0E-03	0,9	6,0E-03	0,5	8,9E-02	0,3

TABLE 4.16 – Évaluations des erreurs et des ordres numériques produits par le schéma du quatrième ordre temps et espace du Théorème 4.4.3.1 avec le triplet  $(\Theta_{b,i}^{O4}, \lambda, \Delta t)$  donné dans (4.95b)-(4.80) et pour le problème de Sod (4.89) au temps final  $t = 0,2$ .

## 4.6 Synthèse et conclusion

À partir d'une correction du schéma HLL standard [118], des schémas du second ordre spatial sont tout d'abord présentés dans ce chapitre. Ces schémas n'utilisent pas de limiteurs de pente [219] et ils sont définis par la donnée d'un triplet  $((\Theta_i^{O2})_{i \in \mathbb{Z}}, \lambda, \Delta t)$  dans  $\mathcal{M}_{m,m}(\mathbb{R}) \times (\mathbb{R}_+^*)^2$ . En restreignant chaque élément de ce triplet et en supposant que la séquence  $(w_i^n)_{i \in \mathbb{Z}}$  n'est pas dans un *voisinage* immédiat du bord de  $\Omega$ , ces schémas vérifient une stabilité entropique discrète globale pour une entropie strictement convexe de classe  $C^2(\Omega, \mathbb{R})$  donnée.

Les restrictions proposées pour chaque élément du triplet  $((\Theta_i^{O2})_{i \in \mathbb{Z}}, \lambda, \Delta t)$  sont respectivement données par les inégalités (4.34), (4.39) et (4.40). Des séquences  $(\Theta_i^{O2})_{i \in \mathbb{Z}}$  et des réels  $\lambda > 0$  satisfaisants les conditions requises sont explicitement donnés. En revanche, la condition CFL (4.40) n'est explicite que dans le cas de l'entropie quadratique et par conséquent, elle nécessite en général la résolution d'une équation non linéaire. L'étude de cette condition CFL dans le cas du transport linéaire scalaire montre un comportement formel de type *parabolique*. Une correction vers un comportement *hyperbolique* est toutefois possible avec une montée en ordre temporel de type *Runge-Kutta Strong Stability Preserving* [107, 109, 110]. D'après [76, 230], ces montées en ordre temporel préservent le résultat de stabilité entropique.

Sur le plan numérique, et comme les schémas proposés n'utilisent pas de technique de limitation, les ordres théoriques sont exactement retrouvés dans les simulations de solutions régulières. Dans le cas de solutions discontinues, les résultats sont conformes aux standards de la littérature y compris pour une fonction de flux  $f$  non convexe. Il est cependant clair que les matrices  $(\Theta_i^{O2})_{i \in \mathbb{Z}}$  ne sont pas toutes équivalentes. Dans le cas scalaire de l'équation de Burgers, certains choix semblent assurer le principe du maximum alors que d'autres produisent des oscillations parasites. Des différences de résultats similaires sont également observées dans le cas des équations d'Euler.

Cette famille de schémas numériques du second ordre est l'objet de la publication [16] et en généralisant sa définition, des schémas non limités du troisième et du quatrième ordre sont également proposés dans ce chapitre. Pour une entropie strictement convexe de classe  $C^2(\Omega, \mathbb{R})$  donnée, l'existence de triplets  $((\Theta_i^{O4})_{(i,k) \in \mathbb{Z} \times \{3,4\}}, \lambda, \Delta t)$  assurant une inégalité d'entropie discrète globale est démontrée.

Les résultats numériques de ces schémas d'ordre trois et quatre sont satisfaisants. Dans le cas de solutions régulières, les ordres théoriques attendus sont exactement mesurés. Dans le cas de solutions discontinues, les résultats sont conformes aux solutions de références mais ils présentent parfois localement des oscillations parasites. Comme pour les schémas du second ordre, l'amplitude de ces oscillations est variable selon le choix des matrices  $(\Theta_i^{O_k})_{(i,k) \in \mathbb{Z} \times \{3,4\}}$ .

## Chapitre 5

# Schémas d'ordre élevé, non limités et à stabilité entropique discrète globale sur maillages non structurés en dimension deux

Ce chapitre propose une extension des schémas du chapitre 4 au cadre de deux dimensions d'espace. Pour mener cette extension, il est d'abord nécessaire de présenter les systèmes hyperboliques ainsi que leurs approximations numériques dans un cadre bi-dimensionnel.

### 5.1 Approximations numériques des systèmes hyperboliques de lois de conservation en dimension deux

#### 5.1.1 Les inégalités d'entropie des systèmes hyperboliques de lois de conservation homogènes en dimension deux

Dans ce chapitre, et en notant  $\cdot^\top$  l'opération de transposition usuelle de  $\mathbb{R}^2$ ,  $x = (x_1, x_2)^\top$  désigne la variable d'espace dans  $\mathbb{R}^2$  et  $t$  est la variable de temps dans  $\mathbb{R}^+$ . Pour un entier naturel non nul  $m$ , considérons un système hyperbolique de  $m$  équations de conservations en deux dimensions d'espace, d'inconnue  $w : \mathbb{R}^2 \times \mathbb{R}^+ \rightarrow \mathbb{R}^m$  et écrit sous la forme

$$\partial_t w + \partial_{x_1} f_1(w) + \partial_{x_2} f_2(w) = 0. \quad (5.1)$$

L'inconnue  $w$  prend ses valeurs dans un ensemble convexe non vide  $\Omega$  inclus dans  $\mathbb{R}^m$ ,  $f_1, f_2 : \Omega \rightarrow \mathbb{R}^m$  désignent deux fonctions de flux données. Ces deux fonctions sont supposées être de classe  $C^2(\Omega, \mathbb{R}^m)$  et leurs matrices jacobiniennes sont respectivement notées  $\nabla f_1(w)$  et  $\nabla f_2(w)$  dans  $\mathcal{M}_{m,m}(\mathbb{R})$ . On suppose que le système d'équations (5.1) est hyperbolique ce qui signifie que pour tous vecteurs unitaires  $n = (n_1, n_2)^\top$  de  $\mathbb{R}^2$  et pour tous  $w$  de  $\Omega$ , la matrice  $\nabla f_1(w)n_1 + \nabla f_2(w)n_2$  est diagonalisable dans  $\mathbb{R}$ .

Le système hyperbolique (5.1) est associé à une donnée initiale notée  $w_0$  qui désigne une fonction mesurable de  $L^\infty(\mathbb{R}^2)$ . En notant  $F = (f_1, f_2) : \Omega \rightarrow (\mathbb{R}^m)^2$  et  $\operatorname{div}_x$  l'opérateur de divergence usuel, un système de lois de conservation bidimensionnel (5.1) s'écrit sous la forme du problème de Cauchy suivant :

$$\begin{cases} \partial_t w + \operatorname{div}_x F(w) = 0, & t > 0, \quad x \in \mathbb{R}^2, \\ w(x, t = 0) = w_0(x). \end{cases} \quad (5.2)$$

L'existence et l'étude des solutions des systèmes hyperboliques donnés ci-dessus sont l'objet de plusieurs travaux [37, 104, 189]. Dans ce mémoire de thèse, l'existence des solutions du problème de Cauchy (5.2) est admise et on admet également que ces solutions vérifient une inégalité d'entropie.

En conservant la notation  $\cdot^\top$  pour la transposition usuelle de  $\mathbb{R}^m$ , une fonction convexe  $\eta : \Omega \rightarrow \mathbb{R}$  est appelée *entropie* s'il existe deux fonctions de flux d'entropie  $g_1, g_2 : \Omega \rightarrow \mathbb{R}$  telles que

$$\nabla g_\mu^\top(w) = \nabla \eta^\top(w) \nabla f_\mu(w), \quad \forall (w, \mu) \in \Omega \times \{1, 2\}.$$

Pour une entropie  $\eta$  donnée, la notation  $G = (g_1, g_2) : \Omega \rightarrow \mathbb{R}^2$  est adoptée et on admet que les solutions du système hyperbolique (5.2) vérifient une *inégalité d'entropie locale* qui s'écrit

$$\partial_t \eta(w) + \operatorname{div}_x G(w) \leq 0.$$

Par conséquent, l'intégration de l'inégalité ci-dessus selon la variable  $x$  puis l'application du Théorème de Green donnent l'*inégalité d'entropie globale* suivante :

$$\int_{\mathbb{R}^2} \partial_t \eta(w) dx \leq 0, \quad \forall t > 0. \quad (5.3)$$

L'objectif de ce chapitre est d'étendre au cas bi-dimensionnel les schémas numériques introduits dans le chapitre précédent. Ainsi, il s'agit de présenter des méthodes numériques valides sur les maillages non structurés, formellement d'ordre élevé et qui vérifient un analogue discret de l'inégalité d'entropie globale (5.3). D'après la bibliographie générale présentée en Section 1.3, il demeure difficile d'obtenir une propriété de stabilité entropique discrète à partir d'un schéma d'ordre élevé. L'ajout de maillages non structurés bi-dimensionnels complexifie encore l'articulation de ces deux propriétés. La section suivante en présente les aspects et les illustre à l'aide d'une bibliographie dédiée.

### 5.1.2 Quelques généralités sur les approximations numériques en dimension deux

Les solutions des systèmes hyperboliques (5.2) sont approchées sur un maillage du temps associé à un maillage de l'espace non structuré. Le maillage temporel est défini par une séquence  $(t^n)_{n \in \mathbb{N}}$  dans  $\mathbb{R}^+$  de pas constant  $\Delta t > 0$ . Le maillage de l'espace est noté  $\mathcal{T}$  dans  $\mathbb{R}^2$  est il paramétré par une taille caractéristique  $h > 0$ . Ce maillage est formé de cellules polygonales simples notées  $(C_i)_{i \in \mathcal{T}}$ . L'ensemble des cellules voisines de  $C_i$  est noté  $\mathcal{N}(i)$  et chaque cellule  $C_i$  de surface  $|C_i| > 0$  est définie par un ensemble d'arêtes  $(\Gamma_{ij})_{j \in \mathcal{N}(i)}$ . Chaque arête est de longueur  $|\Gamma_{ij}| > 0$  et elle est associée à un vecteur normal noté  $n_{ij}$  dans  $\mathbb{R}^2$ . Les vecteurs normaux  $(n_{ij})_{j \in \mathcal{N}(i)}$  sont tous orientés vers l'extérieur de la cellule  $C_i$ .

À la date  $t^n$  et sur chaque cellule  $C_i$ , la quantité  $w(\cdot, t^n)$  est approchée par un objet  $w_i^n$  généralement défini par une constante de  $\mathbb{R}^m$  [27, 174] ou bien par un polynôme à valeurs dans  $\mathbb{R}^m$  [55, 56, 91]. La mise à jour de  $w_i^n$  à la date  $t^n + \Delta t$  est notée  $w_i^{n+1}$  et elle est dépend d'une contribution de chacune des cellules voisines à  $C_i$ . Pour une arête  $\Gamma_{ij}$  donnée, cette contribution s'effectue à l'aide d'un *flux numérique* noté  $\mathcal{F}_{ij}$  dans  $\mathbb{R}^m$  et qui vérifie formellement

$$|\Gamma_{ij}| \mathcal{F}_{ij} \approx \int_{\Gamma_{ij}} f_1(w(x(\xi), t^n)) n_{ij_1} + f_2(w(x(\xi), t^n)) n_{ij_2} d\xi.$$

Ainsi, à partir de la donnée de  $\mathcal{F}_{ij}$ , un *schéma numérique* approchant les solutions de (5.2) s'écrit

$$\frac{w_i^{n+1} - w_i^n}{\Delta t} + \frac{1}{|C_i|} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \mathcal{F}_{ij} = 0, \quad \forall i \in \mathcal{T}. \quad (5.4)$$

Une caractéristique principale d'un tel schéma formel est sa précision usuellement quantifiée par la notion d'*ordre*. Plusieurs définitions de cette notion ont par exemple été proposées dans [18, 32, 55, 144, 219] mais les recherches bibliographiques menées ne montrent pas l'existence d'une définition générale et unificatrice. Néanmoins, il demeure possible de donner une description formelle de l'ordre d'un schéma numérique.

En supposant une régularité suffisante de  $w$ , un schéma numérique (5.4) est dit *d'ordre  $k$  en espace (resp. en temps)* si une quantification de l'erreur d'approximation donnée s'écrit formellement  $\mathcal{O}_{h \rightarrow 0}(h^k)$  (resp.  $\mathcal{O}_{\Delta t \rightarrow 0}(\Delta t^k)$ ). Cette description formelle est l'analogue de celle proposée en une dimension d'espace (voir Chapitre 1, Sections 1.3.1 et 1.4.1) et son écriture au cas bi-dimensionnel sera détaillée en Section 5.1.5. D'une manière générale, la recherche d'un schéma numérique d'ordre élevé est justifiée par deux aspects.

D'abord et d'après [148], il est admis que les montées en ordre corrigent au moins partiellement un défaut des schémas du premier ordre qui, par nature, régularisent artificiellement les solutions



numériques calculées. La seconde justification des montées en ordre est que travailler sur des maillages grossiers avec un schéma d'ordre élevé est plus économique en ressource et en temps de calcul que de travailler sur des maillages fins avec un schéma du premier ordre [168, 193, 210].

Toutefois, sans précaution particulière, les schémas d'ordre élevé sont oscillants voire instables dans les zones de discontinuités [149, 193]. Par conséquent, et en vertu de la bibliographie introductive présentée en Section 1.3.1, il est admis que la stabilité des schémas bi-dimensionnels (5.4) repose sur des inégalités d'entropie discrètes. Pour une arête  $\Gamma_{ij}$  donnée, ces inégalités discrètes utilisent un *flux numérique d'entropie* noté  $\mathcal{G}_{ij}$  dans  $\mathbb{R}$  et qui satisfait formellement

$$|\Gamma_{ij}|\mathcal{G}_{ij} \approx \int_{\Gamma_{ij}} g_1(w(x(\xi), t^n))n_{ij_1} + g_2(w(x(\xi), t^n))n_{ij_2} d\xi.$$

À partir de cette notion, un schéma numérique (5.4) est dit à stabilité entropique locale s'il existe un flux numérique d'entropie  $\mathcal{G}_{ij}$  tel que

$$\frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} + \frac{1}{|C_i|} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}|\mathcal{G}_{ij} \leq 0, \quad \forall i \in \mathcal{T}. \quad (5.5)$$

D'après son expression, un schéma numérique (5.4) s'appuie nécessairement sur le voisinage de la cellule  $C_i$ . Sur les maillages non structurés, le positionnement relatif de  $C_j$  par rapport à  $C_i$  est a priori quelconque. Cette propriété constitue donc une différence importante avec le cas d'une dimension d'espace pour lequel la relation d'ordre de  $\mathbb{R}$  peut être utilisée. Néanmoins, l'étude monodimensionnelle n'est pas complètement décorrélée du cas de deux dimensions d'espace. En effet, en introduisant le périmètre  $|P_i| > 0$  de la cellule  $C_i$ , le schéma numérique (5.4) se réécrit de manière équivalente sous la forme

$$w_i^{n+1} = \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{|P_i|} \left( w_i^n - \frac{\Delta t |P_i|}{|C_i|} \mathcal{F}_{ij} \right), \quad \forall i \in \mathcal{T}.$$

D'après cette écriture, un schéma numérique (5.4) est une combinaison convexe de schémas monodimensionnels chacun définis dans la direction de  $n_{ij}$ . Par conséquent, et en utilisant la convexité des entropies  $\eta$ , il suit nécessairement

$$\eta(w_i^{n+1}) \leq \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{|P_i|} \eta \left( w_i^n - \frac{\Delta t |P_i|}{|C_i|} \mathcal{F}_{ij} \right), \quad \forall i \in \mathcal{T}.$$

Ainsi, en raisonnant direction par direction, il est possible de transférer les stabilités entropiques et l'ordre d'un schéma défini en une dimension d'espace vers un analogue défini sur une grille cartésienne de  $\mathbb{R}^2$  [44, 45, 52, 63, 92, 94, 156, 179, 193, 232]. En revanche, un transfert direct sur les maillages non structurés est impossible car la géométrie des cellules fausse les approximations usuelles des dérivées [201]. Comme de telles approximations sont en général indispensables aux schémas numériques précis, l'introduction de maillages non structurés requiert donc d'abord des méthodes adaptées d'approximation des gradients.

### 5.1.3 L'approximation des gradients en dimension deux

Les approximations des dérivées reposent en général sur des développements de Taylor. En notant  $x_i$  le centre de masse d'une cellule  $C_i$  et en supposant une régularité suffisante, le développement de Taylor de  $w$  sur un maillage non structuré s'écrit

$$w(x_j, t^n) = w(x_i, t^n) + \sum_{\mu=1}^2 (x_j - x_i)_\mu \partial_{x_\mu} w(x_i, t^n) + \frac{1}{2} \sum_{\mu, \nu=1}^2 (x_j - x_i)_\mu \partial_{x_\mu x_\nu} w(x_i, t^n) (x_j - x_i)_\nu + \mathcal{O}(h^2). \quad (5.6)$$

En une dimension d'espace et aussi par extension directe sur grilles cartésiennes, l'association de plusieurs développements de Taylor permet d'obtenir des approximations du gradient et de la matrice

hessienne de  $w$ . En revanche, comme le montre le développement (5.6), les maillages quelconques couplent nécessairement les deux directions de l'espace. Dans ce cas, les développements de Taylor donnent seulement accès à une projection des dérivées et par conséquent, le calcul de leurs approximations devient problématique. Plusieurs résultats ont été proposés pour surmonter cette difficulté et les paragraphes ci-après en donnent une présentation organisée en quatre catégories réparties selon les méthodes employées.

La première catégorie décrite dans [89] consiste à contraindre le maillage pour faire coïncider des directions de projections avec des quantités connues et ainsi obtenir des approximations du gradient au premier ordre. En notant  $d_{i,j}$  la distance euclidienne entre  $x_i$  et  $x_j$ , la contrainte de géométrie usuelle est  $d_{i,j}n_{ij} = x_j - x_i$  et elle permet d'écrire

$$\partial_{x_1}w(x_i, t^n)n_{ij_1} + \partial_{x_2}w(x_i, t^n)n_{ij_2} = \frac{w_j^n - w_i^n}{d_{i,j}} + \mathcal{O}(h).$$

La seconde catégorie de méthodes utilise le théorème de Green-Gauss pour transférer un gradient à la cellule vers des quantités définies sur les arêtes. Ces quantités aux arêtes sont ensuite calculées avec une procédure de point fixe [70, 201, 202] ou par la résolution d'un système linéaire couplant toutes les inconnues du maillage [28, 169]. Bien que des approximations d'ordre deux sur des maillages non structurés aient pu être proposées dans [28], les approximations du gradient obtenues sont en général du premier ordre et d'ordre deux seulement sur les grilles cartésiennes.

La troisième catégorie de méthodes s'appuie sur un maillage dual pour calculer des valeurs complémentaires de l'inconnue  $w$ . Avec ces valeurs additionnelles, les auteurs de [67, 81, 154] proposent des approximations du gradient au premier ordre sur des maillages quelconques. De leurs côtés, Herbin *et al* introduisent dans [121] des gradients discrets du second ordre sur les maillages contraints définis dans [89],

La quatrième catégorie s'articule autour de problèmes d'optimisation obtenus en pondérant plusieurs développements de Taylor de la forme (5.6). En notant  $\omega_{ij}$  des réels positifs et  $p$  un entier, ces problèmes d'optimisation s'écrivent formellement

$$\min_{(X,Y)} \sum_{ij} \omega_{ij} \left| w_j^n - w_i^n - \sum_{\mu=1}^2 (x_j - x_i)_\mu X_\mu - \frac{1}{2} \sum_{\mu,\nu=1}^2 (x_j - x_i)_\mu Y_{\mu,\nu} (x_j - x_i)_\nu \right|^p. \quad (5.7)$$

Les approximations des gradients et des matrices hessiennes de  $w$  sont ensuite construites à partir du couple  $(X^*, Y^*)$  solution du problème ci-dessus. Dans cette optique, des gradients discrets du premier ordre sont décrits dans [128] et du second ordre sur les grilles cartésiennes dans [222]. Néanmoins, sur maillages quelconques, ces méthodes d'optimisation n'assurent pas systématiquement le second ordre des approximations des gradients. D'après [43, 218] et [80, 163], cet ordre est largement influencé par les développements de Taylor et par les pondérations choisies dans la définition du problème (5.7). Ces problèmes de minimisations (5.7) sont parfois hybridés avec les méthodes détaillées dans les paragraphes ci-avant [68, 76, 167, 192] mais leurs résultats demeurent sujets aux oscillations parasites et ils nécessitent donc l'emploi de limiteurs de pente.

Ainsi, plusieurs méthodes d'approximations des gradients sur les maillages non structurés sont désormais établies. Comme le montre la recherche bibliographique exposée ci-dessus, elles ne sont pas toutes équivalentes et d'après les comparatifs proposés dans [66, 201], le choix d'une méthode n'est pas simple. Néanmoins, à partir d'un gradient discret, il est possible de reconstruire une approximation du second ordre de  $w$  le long de l'arête  $\Gamma_{ij}$ . L'introduction de cette approximation dans un flux numérique formellement mono dimensionnel dans la direction de  $n_{ij}$  donne un schéma numérique (5.4) d'ordre élevé. Toutefois, comme dans le cadre d'une dimension d'espace, cette propriété de précision doit s'articuler à une propriété de stabilité.

#### 5.1.4 La stabilité des schémas d'ordre élevé en dimension deux

À partir du schéma MUSCL usuel [219], plusieurs schémas du second ordre ont été proposés sur des maillages non structurés. Avec des limiteurs adaptés, ces schémas préservent le domaine  $\Omega$  [27, 174]

mais ils vérifient des stabilités entropiques incomplètes [22, 77]. En notant  $\bar{\eta}_i$  une approximation de  $\eta(w_i^n)$ , ces versions incomplètes de (5.5) s'écrivent

$$\frac{\eta(w_i^{n+1}) - \bar{\eta}_i}{\Delta t} + \frac{1}{|C_i|} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \mathcal{G}_{ij} \leq 0, \quad \forall i \in \mathcal{T}.$$

Des schémas d'ordre deux à quatre de type (W)ENO sur maillages non structurés sont par ailleurs établis dans [17, 90, 97, 130, 155, 157, 217] mais aucune preuve de stabilité entropique n'est proposée. D'une manière générale, et comme le détaillent par exemple les auteurs de [157, 217], la difficulté des méthodes (W)ENO sur les maillages non structurés réside dans les reconstructions polynomiales d'ordre élevé. D'après la bibliographie de la section 5.1.3, plusieurs méthodes de reconstruction sont possibles mais leurs formulations rendent extrêmement difficile voire impossible les preuves de stabilité. Ray *et al* proposent de contourner cet obstacle dans [180] par l'utilisation d'un maillage dual et une version des variables entropiques. Il en résulte un schéma numérique du second ordre vérifiant une version semi-discrète de (5.5) qui s'écrit

$$\frac{d}{dt} \eta(w_i(t)) + \frac{1}{|C_i|} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \mathcal{G}_{ij}(t) \leq 0, \quad \forall i \in \mathcal{T}. \quad (5.8)$$

En ce qui concerne les méthodes DG, le formalisme SPB a été étendu avec succès au cas des maillages non structurés [55, 56, 91]. Par conséquent, les auteurs de [50, 55, 56, 91, 188, 220, 228] proposent des schémas d'ordre quatre à cinq et satisfaisant des inégalités d'entropie semi-discrètes globales de la forme

$$\frac{d}{dt} \sum_{i \in \mathcal{T}} \eta(w_i(t)) |C_i| \leq 0, \quad \forall t > 0. \quad (5.9)$$

Un résultat remarquable est [137] dans lequel Kuzmin *et al* proposent un schéma DG explicite en temps, défini sur des maillages non structurés, du second ordre et satisfaisant une inégalité d'entropie discrète locale (5.5). Malheureusement, ce schéma aux éléments finis contient un terme de la forme  $\mathcal{O}(\Delta t)$  ce qui, avec la condition CFL, le réduit à un schéma du premier ordre.

Plus généralement, la méthode des éléments finis associée aux schémas RD [2, 69] lève deux difficultés quant à l'approximation sur les maillages non structurés. D'abord, la formulation intégrale des résidus distribués est largement indépendante du maillage [3, 4, 5, 6, 9]. Ensuite, cette formulation intégrale permet l'usage du Théorème de Green. Par conséquent, un schéma RD peut s'affranchir de l'approximation des gradients des fonctions de flux [221, Section 1.6.2] ce qui contourne les difficultés énoncées en Section 5.1.3. Ainsi, en complétant ces considérations par les résidus de stabilité présentés dans [3, 8, 9], les auteurs de [5] proposent un schéma RD pour les équations d'Euler, défini sur des maillages non structurés, d'ordre quelconque et qui satisfait

$$\sum_{i \in \mathcal{T}} \eta(w_i^N) |C_i| \leq \sum_{i \in \mathcal{T}} \eta(w_i^0) |C_i|,$$

où  $t^0$  et  $t^N$  désignent respectivement la date initiale et la date finale de simulation.

D'après les paragraphes ci-dessus, les schémas bidimensionnels d'ordre élevé vérifient très souvent des stabilités entropiques semi-discrètes (5.8)-(5.9) dont le transfert en temps discret est détaillé en Section 1.3.4. Ensuite, l'ensemble des méthodes décrites dans cette section utilisent des limiteurs de pente dont le choix, déjà difficile en une dimension d'espace, est encore complexifié par les effets de géométries des maillages non structurés [27, 63, 77].

Par conséquent, l'étude bibliographique ne montre pas l'existence de schémas numériques du second et du troisième ordre, sans limiteurs de pente, définis sur des maillages non structurés, explicites en temps et qui vérifient une inégalité d'entropie discrète globale donnée par

$$\sum_{i \in \mathcal{T}} \frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} |C_i| \leq 0, \quad \forall t > 0.$$

Ce chapitre s'intéresse à l'existence de ces schémas dans le cadre de la méthode des volumes finis en deux dimensions d'espace qui est maintenant rappelée.

### 5.1.5 La méthode des volumes finis en deux dimensions d'espace

La méthode des volumes finis nécessite tout d'abord un maillage temporel dans  $\mathbb{R}^+$ . Un tel maillage est donné par une séquence  $(t^n)_{n \in \mathbb{N}}$  de pas uniforme  $\Delta t > 0$  et qui vérifie

$$t^0 = 0, \quad t^{n+1} = t^n + \Delta t, \quad \forall n \in \mathbb{N}.$$

Ensuite, pour  $N$  dans  $\mathbb{N}^*$ , un *maillage non structuré* en dimension deux est un pavage de  $\mathbb{R}^2$  réalisé avec  $N$  polygones simples notés  $C_i$  et appelés *cellules*. Un tel maillage est noté  $\mathcal{T}$  et son nombre de cellules  $N$  est désigné par  $\text{Card}(\mathcal{T})$ . Ainsi, un maillage non structuré de  $\mathbb{R}^2$  s'écrit

$$\mathcal{T} = \bigcup_{i=1}^{\text{Card}(\mathcal{T})} C_i.$$

En considérant une bijection quelconque entre l'intervalle d'entier  $\{1, \dots, \text{Card}(\mathcal{T})\}$  et l'ensemble des cellules de  $\mathcal{T}$ , chaque cellule  $C_i$  est assimilée à son indice  $i$ . Cette convention permet la notation abusive mais synthétique  $i \in \mathcal{T}$ . De plus, la géométrie de  $C_i$  est caractérisée par  $\mathcal{N}(i)$  côtés définis par un ensemble d'arêtes  $(\Gamma_{ij})_{j \in \{1, \dots, \mathcal{N}(i)\}}$ . Sa surface est notée  $|C_i| > 0$  et  $x_i$  désigne son centre de masse dans  $\mathbb{R}^2$ . La taille de chaque cellule  $C_i$  est uniformément contrôlée par un réel  $h > 0$ . Dans ce mémoire de thèse, et en notant  $d : (\mathbb{R}^2)^2 \rightarrow \mathbb{R}^+$  la fonction de distance euclidienne usuelle de  $\mathbb{R}^2$ , le paramètre  $h$  est défini par  $h = \max_{(i,j) \in \mathcal{T}^2} d(x_i, x_j)$ .

Par abus de notations, et pour une cellule  $C_i$  donnée,  $\mathcal{N}(i)$  désigne également l'ensemble des cellules voisines de  $C_i$  et l'arête  $\Gamma_{ij}$ , de longueur  $|\Gamma_{ij}| > 0$ , est l'*interface orientée* de  $C_i$  vers  $C_j$ . Le vecteur normal de  $\Gamma_{ij}$  est noté  $n_{ij} = (n_{ij1}, n_{ij2})^\top$  dans  $\mathbb{R}^2$  et par conséquent,  $n_{ij} = -n_{ji}$ . L'ensemble des interfaces non orientées de  $\mathcal{T}$  est noté  $\mathcal{E}$  et il s'écrit formellement

$$\mathcal{E} = \bigcup_{ij} \Gamma_{ij}.$$

La Figure 5.1 illustre les notations utilisées autour d'une interface orientée  $\Gamma_{ij}$ .

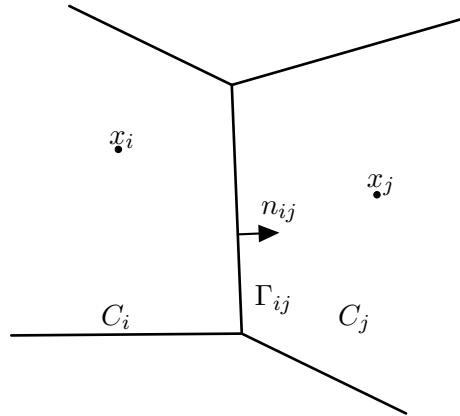


FIGURE 5.1 – Interface orientée  $\Gamma_{ij}$  localisée entre deux cellules  $C_i$  et  $C_j$  de centre de masse respectif  $x_i$  et  $x_j$ .

Avec la donnée de maillages temporels et spatiaux ainsi définis, il est alors possible de construire une approximation des solutions d'un système hyperbolique de  $m$  équations de conservations de la forme

$$\begin{cases} \partial_t w + \text{div}_x F(w) = 0, & t > 0, \quad x \in \mathbb{R}^2, \\ w(x, t = 0) = w_0(x). \end{cases} \quad (5.10)$$

Tout d'abord, la donnée initiale  $w_0$  est discrétisée à l'aide d'une séquence de  $\mathbb{R}^m$  notée  $(w_i^0)_{i \in \mathcal{T}}$  et définie par

$$w_i^0 = \frac{1}{|C_i|} \int_{C_i} w_0(x) \, dx, \quad \forall i \in \mathcal{T}.$$

Ensuite, pour  $(w_i^n)_{i \in \mathcal{T}}$  une suite d'éléments de  $\mathbb{R}^m$  pour tout  $n$  dans  $\mathbb{N}$ , une fonction constante par morceaux notée  $w_\Delta : \mathbb{R}^2 \times \mathbb{N} \rightarrow \mathbb{R}^m$  est considérée. Cette fonction s'écrit

$$w_\Delta(x, t^n) = \sum_{i \in \mathcal{T}} w_i^n \chi_{C_i}(x). \quad (5.11)$$

Par construction,  $w_\Delta(\cdot, t^0)$  est une approximation constante par morceaux de la fonction  $w_0$  qui s'identifie à  $w(\cdot, t^0)$ . Ainsi, pour obtenir une approximation de  $(w(\cdot, t^n))_{n \in \mathbb{N}}$ , il est suffisant de définir un schéma numérique qui permet le calcul de  $(w_i^{n+1})_{i \in \mathcal{Z}}$  à partir de  $(w_i^n)_{i \in \mathcal{Z}}$ .

**Définition 5.1.5.1** (Schéma numérique volumes finis conservatif en dimension deux). *Soit  $\mathcal{F} : (\mathbb{R}^m)^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^m$  une fonction de flux numérique continue. Soit  $u : \mathbb{R}^2 \rightarrow \mathbb{R}^m$  une fonction régulière et notons*

$$u_i = \frac{1}{|C_i|} \int_{C_i} u(x) dx, \quad \forall i \in \mathcal{T}.$$

Un schéma numérique volumes finis en dimension deux qui approche les solutions de (1.30) s'écrit sous la forme

$$w_i^{n+1} = w_i^n - \frac{\Delta t}{|C_i|} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \mathcal{F}(w_i^n, w_j^n, n_{ij}), \quad \forall (i, j) \in \mathcal{T} \times \mathcal{N}(i). \quad (5.12)$$

Un tel schéma est dit

i) *conservatif si le flux numérique  $\mathcal{F}$  satisfait  $\mathcal{F}(w_i^n, w_j^n, n_{ij}) = -\mathcal{F}(w_j^n, w_i^n, n_{ji})$ . Dans ce cas, aux contributions des termes de bord près, la séquence  $(w_i^{n+1})_{i \in \mathcal{T}}$  vérifie*

$$\sum_{i \in \mathcal{T}} w_i^{n+1} |C_i| = \sum_{i \in \mathcal{T}} w_i^n |C_i|. \quad (5.13)$$

ii) *consistant si le flux numérique  $\mathcal{F}$  vérifie*

$$\mathcal{F}(w, w, n_{ij}) = f_1(w) n_{ij1} + f_2(w) n_{ij2}, \quad \forall (w, n_{ij}) \in \Omega \times \mathbb{R}^2.$$

iii) *d'ordre  $k$  si, pour une interface orientée  $\Gamma_{ij}$  fixée, le flux numérique  $\mathcal{F}$  vérifie également*

$$\mathcal{F}(u_i, u_j, n_{ij}) = \frac{1}{|\Gamma_{ij}|} \int_{\Gamma_{ij}} (f_1(u(x(\xi))) n_{ij1} + f_2(u(x(\xi))) n_{ij2}) d\xi + \mathcal{O}(h^k).$$

Avec la définition ci-dessus, et comme il n'y a pas de confusion possible, la notation  $\mathcal{F}(w_i^n, w_j^n, n_{ij}) = \mathcal{F}_{ij}$  est désormais adoptée. Par conséquent un schéma numérique conservatif (5.12) en deux dimensions d'espace se lit

$$w_i^{n+1} = w_i^n - \frac{\Delta t}{|C_i|} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \mathcal{F}_{ij}, \quad \forall (i, j) \in \mathcal{T} \times \mathcal{N}(i). \quad (5.14)$$

D'après la Définition 5.1.5.1, l'égalité de conservation (5.13) est conditionnée par une propriété sur le flux numérique qui s'écrit  $\mathcal{F}_{ij} = -\mathcal{F}_{ji}$ . D'une manière générale, une quantité de  $\mathbb{R}^m$  qui vérifie une telle égalité est qualifiée de *conservative*. Cette propriété permet l'utilisation d'intégrations par parties discrètes.

**Lemme 5.1.5.1** (Intégrations par parties discrètes). *Considérons une séquence  $(X_i)_{i \in \mathcal{T}}$  à support compact et à valeurs dans  $\mathbb{R}^m$ . Considérons ensuite  $Y_{ij}$  dans  $\mathbb{R}^m$  une quantité conservative qui vérifie pour toute interface  $\Gamma_{ij}$  l'égalité  $Y_{ij} = -Y_{ji}$ . En notant  $\cdot$  le produit scalaire usuel de  $\mathbb{R}^m$ , l'égalité suivante est satisfaite :*

$$-\sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| X_i \cdot Y_{ij} = \sum_{ij \in \mathcal{E}} |\Gamma_{ij}| Y_{ij} \cdot (X_j - X_i). \quad (5.15)$$

L'égalité ci-dessus est appelée *intégration par parties discrète*.

*Démonstration.* L'égalité (5.15) est obtenue par une réorganisation de son terme de gauche. Par définition de ce terme, chaque interface  $\Gamma_{ij}$  reçoit deux contributions : la première de  $i$  vers  $j$  et la seconde de  $j$  vers  $i$ . Par conséquent, en utilisant la propriété  $Y_{ij} = -Y_{ji}$ , il suit

$$\begin{aligned} \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| X_i \cdot Y_{ij} &= \sum_{ij \in \mathcal{E}} \left( \underbrace{|\Gamma_{ij}| X_i \cdot Y_{ij}}_{i \rightarrow j} + \underbrace{|\Gamma_{ji}| X_j \cdot Y_{ji}}_{j \rightarrow i} \right), \\ &= \sum_{ij \in \mathcal{E}} \left( |\Gamma_{ij}| X_i \cdot Y_{ij} - |\Gamma_{ij}| X_j \cdot Y_{ij} \right), \\ &= - \sum_{ij \in \mathcal{E}} |\Gamma_{ij}| (X_j - X_i) \cdot Y_{ij}. \end{aligned}$$

Au signe près, l'égalité ci-dessus correspond au résultat (5.15), et par conséquent ceci conclut la démonstration.  $\square$

Le calcul ci-dessus fait clairement apparaître la nécessité de la propriété  $\mathcal{F}_{ij} = -\mathcal{F}_{ji}$  quant à l'obtention de l'égalité de conservation discrète (5.13). Cette propriété de conservation est également nécessaire à la définition des inégalités d'entropie discrètes bi-dimensionnelles qui sont maintenant détaillées.

**Définition 5.1.5.2** (Stabilité entropique discrète en dimension deux). *Soit un système hyperbolique de la forme (5.10) équipé d'un couple entropie, flux d'entropie  $(\eta, G)$  et dont les solutions vérifient une inégalité d'entropie de la forme*

$$\partial_t \eta(w) + \operatorname{div}_x G(w) \leq 0.$$

i) *Soit  $\mathcal{G} : (\mathbb{R}^m)^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$  une fonction continue. La fonction  $\mathcal{G}$  est dite consistante avec  $G$  si*

$$\mathcal{G}(w, w, n_{ij}) = g_1(w) n_{ij1} + g_2(w) n_{ij2}, \quad \forall (w, n_{ij}) \in \Omega \times \mathbb{R}^2.$$

ii) *Un schéma numérique volumes finis conservatif (5.12) approchant les solutions de (5.10) est dit à stabilité entropique locale pour le couple  $(\eta, G)$  s'il existe une fonction  $\mathcal{G}$  consistante avec  $G$ , qui vérifie  $\mathcal{G}(w_i^n, w_j^n, n_{ij}) = -\mathcal{G}(w_j^n, w_i^n, n_{ji})$  et telle que*

$$\frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} + \frac{\Delta t}{|C_i|} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \mathcal{G}(w_i^n, w_j^n, n_{ij}) \leq 0, \quad \forall (i, j) \in \mathcal{T} \times \mathcal{N}(i). \quad (5.16)$$

*Dans ce cas, la fonction  $\mathcal{G}$  est appelée flux numérique d'entropie.*

iii) *Un schéma numérique volumes finis conservatif (5.12) approchant les solutions de (5.10) est dit à stabilité entropique globale pour le couple  $(\eta, G)$  si*

$$\sum_{i \in \mathbb{Z}} \frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} |C_i| \leq 0.$$

À partir des définitions ci-dessus, un schéma numérique conservatif (5.14) est entièrement défini par la donnée de son flux numérique  $\mathcal{F}_{ij}$ . Ainsi, l'objectif de ce chapitre est de définir  $\mathcal{F}_{ij}$  comme une extension des flux numériques non limités mono-dimensionnels du chapitre 4. Puisqu'une extension des ordres deux et trois est considérée, il est maintenant nécessaire de proposer des approximations des gradients et des matrices hessiennes consistantes, non limités et définies sur des maillages non structurés.

## 5.2 Gradients et matrices hessiennes discrets non limités

### 5.2.1 Définition de maillage admissible

Pour proposer des approximations des gradients et des matrices hessiennes, une notion de maillage admissible est introduite dans cette section. La définition de cette notion nécessite des notations complémentaires qui sont maintenant détaillées.

Tout d'abord, la fonction de distance euclidienne usuelle de  $\mathbb{R}^2$  est notée  $d : (\mathbb{R}^2)^2 \rightarrow \mathbb{R}^+$  et elle s'écrit

$$d(x_a, x_b) = \sqrt{\sum_{\mu=1}^2 (x_a - x_b)_\mu^2}, \quad \forall (x_a, x_b) \in (\mathbb{R}^2)^2.$$

Comme cela ne porte pas à confusion, la quantité  $d(x_a, x_b)$  est notée  $d_{a,b}$ . Par conséquent, le paramètre  $h > 0$  qui contrôle uniformément la taille des cellules se lit désormais  $h = \max_{ij \in \mathcal{E}} d_{i,j}$ .

Ensuite, pour une interface orientée  $\Gamma_{ij}$  fixée,  $\bar{x}_{ij}$  dans  $\mathbb{R}^2$  est le milieu de cette interface. L'intersection de  $\Gamma_{ij}$  avec le segment formé par les centres de masse des cellules  $C_i$  et  $C_j$  est noté  $x_{ij}$  dans  $\mathbb{R}^2$ . Les quantités  $\bar{x}_{ij}$  et  $x_{ij}$  sont illustrées par la Figure 5.2 et la convention d'écriture de la fonction  $d$  donne

$$d_{i,j} = d(x_i, x_j) > 0, \quad d_{j,ij} = d(x_{ij}, x_j) > 0, \quad d_{i,ij} = d(x_{ij}, x_i) > 0. \quad (5.17)$$

Par conséquent, l'égalité  $d_{j,ij} + d_{i,ij} = d_{i,j}$  est immédiate. Pour  $\alpha$  dans  $\{C_i, C_j, \Gamma_{ij}\}$ , des formes linéaires  $Q_{ij}^\alpha$  et  $\mathcal{L}_{ij}$  sont aussi considérées. Ces formes linéaires sont associées à l'interface orientée  $\Gamma_{ij}$  donnée et elles s'écrivent respectivement

$$Q_{ij}^\alpha : \begin{cases} \mathcal{M}_{2,2}(\mathbb{R}) \rightarrow \mathbb{R}, \\ A \mapsto \frac{1}{|\alpha|} \int_\alpha (x - x_{ij})^\top A (x - x_{ij}) dx, \quad \forall \alpha \in \{C_i, C_j, \Gamma_{ij}\}, \end{cases} \quad (5.18)$$

$$\mathcal{L}_{ij} : \begin{cases} \mathcal{M}_{2,2}(\mathbb{R}) \rightarrow \mathbb{R}, \\ A \mapsto -\frac{d_{i,ij}}{2d_{i,j}} Q_{ij}^{C_j}(A) - \frac{d_{j,ij}}{2d_{i,j}} Q_{ij}^{C_i}(A) + (\bar{x}_{ij} - x_{ij})^\top A (x_{ij} - x_i) + \frac{1}{2} Q_{ij}^{\Gamma_{ij}}(A). \end{cases} \quad (5.19)$$

D'après l'expression ci-dessus, si la matrice  $A$  est indépendante de la variable  $x$  alors l'intégrale qui définit  $Q_{ij}^\alpha(A)$  dans (5.18) se calcule explicitement avec les coordonnées des objets  $\alpha$ . Ces formules de calcul explicites sont par exemples données dans [113]. Dans le cadre des approximations développées dans cette section,  $A$  est constante et par conséquent le calcul pratique de  $Q_{ij}^\alpha$  et de  $\mathcal{L}_{ij}$  respectivement donnés par (5.18)-(5.19) ne pose aucune difficulté.

Enfin, pour  $r$  dans  $\mathbb{N}^*$  et pour  $(Y_{11}, Y_{12}, Y_{21}, Y_{22})$  dans  $(\mathbb{R}^r)^4$ , il est possible de définir  $Y$  une matrice par blocs dans  $\mathcal{M}_{2r,2}(\mathbb{R})$  et une séquence  $(Y_l)_{l \in \{1, \dots, r\}}$  dans  $\mathcal{M}_{2,2}(\mathbb{R})$  telles que

$$Y = \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix}, \quad Y_l = \begin{pmatrix} (Y_{11})_l & (Y_{12})_l \\ (Y_{21})_l & (Y_{22})_l \end{pmatrix}.$$

Avec les définitions ci-dessus et par abus de notations, les formes linéaires  $Q_{ij}^\alpha$  et  $\mathcal{L}$  respectivement définies dans (5.18) et dans (5.19) sont étendues dans  $\mathbb{R}^r$  de la manière suivante :

$$Q_{ij}^\alpha(Y) = (Q_{ij}^\alpha(Y_1), \dots, Q_{ij}^\alpha(Y_l), \dots, Q_{ij}^\alpha(Y_r))^\top, \quad (5.20a)$$

$$\mathcal{L}_{ij}(Y) = (\mathcal{L}_{ij}(Y_1), \dots, \mathcal{L}_{ij}(Y_l), \dots, \mathcal{L}_{ij}(Y_r))^\top. \quad (5.20b)$$

Dans le cas  $r = 1$ , les notations ci-dessus dégènèrent respectivement vers les définitions (5.18) et (5.19).

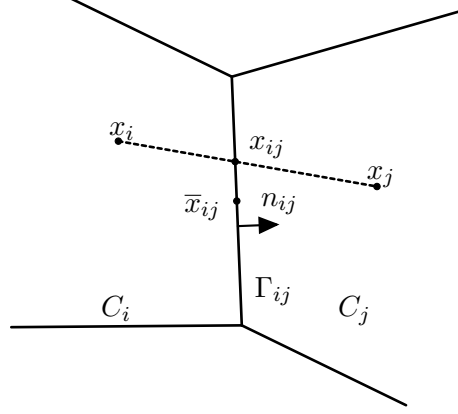


FIGURE 5.2 – Interface orientée  $\Gamma_{ij}$  localisée entre deux cellules  $C_i$  et  $C_j$  d'un maillage non structuré  $\mathcal{T}$  dans lequel  $\bar{x}_{ij}$  est le milieu de  $\Gamma_{ij}$  et avec  $x_{ij} = \Gamma_{ij} \cap [x_i, x_j]$ .

À partir des notations définies ci-avant, des hypothèses sur la géométrie des cellules du maillage sont maintenant réunies sous la notion de maillage admissible.

**Définition 5.2.1** (Maillage non structuré admissible). *Soient  $\mathcal{T}$  un maillage dans  $\mathbb{R}^2$  tel que défini en Section 5.1.5,  $I_2$  la matrice identité de taille deux et une séquence  $(v_i)_{i \in \mathcal{T}}$  dans  $\mathbb{R}$ . Pour toute cellule  $C_i$  dans  $\mathcal{T}$ , considérons les matrices  $M_i$  dans  $\mathcal{M}_{2,2}(\mathbb{R})$  telles que*

$$M_i = I_2 - \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{|C_i|} n_{ij} (\bar{x}_{ij} - x_{ij})^\top. \quad (5.21)$$

En notant  $\text{GL}_2(\mathbb{R})$  l'ensemble des matrices carrées inversibles de tailles deux et avec la convention  $\mathcal{N}(\alpha, \beta) = \mathcal{N}(\alpha) \times \mathcal{N}(\beta)$ , un maillage  $\mathcal{T}$  dans  $\mathbb{R}^2$  est dit admissible si les conditions suivantes sont satisfaites.

i) Pour toute cellule  $C_i$  dans  $\mathcal{T}$ ,  $M_i$  vérifie

$$M_i \in \text{GL}_2(\mathbb{R}) \quad \text{et} \quad M_i^{-1} = \mathcal{O}(1). \quad (5.22)$$

Dans ce cas, pour une cellule  $C_i$  donnée, il est possible de définir une séquence  $(b_{ij})_{j \in \mathcal{N}(i)}$  dans  $\mathbb{R}^2$  telle que

$$b_{ij} = \frac{|\Gamma_{ij}|}{|C_i|} M_i^{-1} n_{ij}, \quad \forall j \in \mathcal{N}(i), \quad (5.23a)$$

$$\sum_{j \in \mathcal{N}(i)} b_{ij} = 0. \quad (5.23b)$$

ii) Pour toute cellule  $C_i$  dans  $\mathcal{T}$ , il existe une unique matrice  $\mathbf{H}_i$  dans  $\mathcal{M}_{2,2}(\mathbb{R})$  telle que

$$\begin{aligned} \mathbf{H}_i - \sum_{(j,r) \in \mathcal{N}(i,i)} \frac{d_{j,ij}}{d_{i,j}} \mathcal{L}_{ir}(\mathbf{H}_i) b_{ir} b_{ij}^\top - \sum_{(j,p) \in \mathcal{N}(i,j)} \frac{d_{i,ij}}{d_{i,j}} \mathcal{L}_{jp}(\mathbf{H}_i) b_{jp} b_{ij}^\top \\ = \sum_{(j,r) \in \mathcal{N}(i,i)} \frac{d_{j,ij}}{d_{i,j}} \frac{d_{l,ir} v_i + d_{i,ir} v_r}{d_{i,r}} b_{ir} b_{ij}^\top + \sum_{(j,p) \in \mathcal{N}(i,j)} \frac{d_{i,ij}}{d_{i,j}} \frac{d_{p,jp} v_j + d_{j,jp} v_p}{d_{j,p}} b_{jp} b_{ij}^\top, \quad \forall (v_i)_{i \in \mathcal{T}} \subset \mathbb{R}. \end{aligned} \quad (5.24)$$

Les conditions (5.22) sont des contraintes qui restreignent l'anisotropie des cellules. L'ensemble des maillages qui vérifie ces conditions est non vide puisque les grilles cartésiennes conviennent. Dans ce cas,  $M_i = I_2$  et par conséquent, les conditions (5.22) sont trivialement satisfaites. Les maillages obtenus par déformation des maillages cartésiens satisfont la contrainte d'inversibilité (5.22) au moins tant que



$M_i$  reste à diagonale strictement dominante. Dans la pratique, il semble que les maillages constitués de triangles ou de quadrangles convexes non structurés vérifient (5.22).

La propriété (5.23b) est une conséquence directe du Théorème de Green-Gauss appliqué à une fonction constante. En effet, et comme détaillé dans [28, 201] par exemple, ce théorème permet d'écrire

$$\sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| n_{ij} = 0.$$

Ainsi, et d'après la définition de  $b_{ij}$  donnée par (5.23a), la multiplication à gauche de l'égalité ci-dessus par  $M_i^{-1}/|C_i|$  permet de déduire la propriété (5.23b).

L'assertion *ii)* est également une contrainte sur l'anisotropie des cellules qui se traduit par l'unicité des solutions de l'équation (5.24) posée dans  $\mathcal{M}_{2,2}(\mathbb{R})$ . La notation  $\mathcal{N}(\alpha, \beta)$  désigne le produit cartésien  $\mathcal{N}(\alpha) \times \mathcal{N}(\beta)$  et par conséquent, les sommes utilisées dans (5.24) se lisent formellement

$$\sum_{(j,r) \in \mathcal{N}(i,i)} = \sum_{j \in \mathcal{N}(i)} \sum_{r \in \mathcal{N}(i)} \quad \text{et} \quad \sum_{(j,p) \in \mathcal{N}(i,j)} = \sum_{j \in \mathcal{N}(i)} \sum_{p \in \mathcal{N}(j)}. \quad (5.25)$$

Avec la convention d'écriture des sommes doubles (5.25) et en s'appuyant sur les symétries d'une grille cartésienne uniforme de taille  $h > 0$ , il est possible de déduire

$$\sum_{(j,r) \in \mathcal{N}(i,i)} \frac{d_{j,ij}}{d_{i,j}} \mathcal{L}_{ir}(\cdot) b_{ir} b_{ij}^\top = \frac{1}{2h^2} \sum_{(j,r) \in \mathcal{N}(i,i)} \mathcal{L}_{ir}(\cdot) n_{ir} n_{ij}^\top = 0. \quad (5.26)$$

Par conséquent, en notant  $(l, k)$  dans  $(\mathbb{N}^*)^2$  le couple d'entier désignant la cellule  $C_i$  d'une grille cartésienne uniforme, l'équation (5.24) s'écrit

$$\mathbf{H}_{k,l} = \frac{1}{4h^2} \begin{pmatrix} v_{k+2,l} - 2v_{k,l} + v_{k-2,l} & v_{k+1,l+1} - v_{k+1,l-1} - v_{k-1,l+1} + v_{k-1,l-1} \\ v_{k+1,l+1} - v_{k+1,l-1} - v_{k-1,l+1} + v_{k-1,l-1} & v_{k,l+2} - 2v_{k,l} + v_{k,l-2} \end{pmatrix}. \quad (5.27)$$

Ainsi, d'après l'égalité ci-dessus, les grilles cartésiennes uniformes satisfont la Définition 5.2.1-*ii)* et comme le montrera la section 5.2.3, d'autres maillages vérifient également cette condition. Numériquement, il semble une nouvelle fois que les maillages non structurés constitués de triangles ou de quadrangles convexes vérifient la Définition 5.2.1-*ii)*.

Avec la notion de maillage admissible donnée dans la Définition 5.2.1, il est désormais possible d'établir une approximation des gradients et des matrices hessiennes.

## 5.2.2 Gradients discrets du premier et du second ordre

Pour une fonction régulière  $u : \mathbb{R}^2 \rightarrow \mathbb{R}^m$  donnée, des gradients discrets de  $u$  au premier et au second ordre sont proposés dans cette section. Ces gradients discrets sont chacun désignés par un exposant  $Ok$  pour  $k$  dans  $\{1, 2\}$  et ils s'écrivent dans  $\mathcal{M}_{2m,1}(\mathbb{R})$  de la manière suivante :

$$\overline{\nabla_x u_i}^{Ok} = \left( \overline{\partial_{x_1} u_i}^{Ok} \quad \overline{\partial_{x_2} u_i}^{Ok} \right)^\top, \quad \forall i \in \mathcal{T}. \quad (5.28)$$

La définition des gradients du second ordre repose sur des matrices hessiennes discrètes de  $u$  qui sont notées  $\overline{H_{x,u_i}}$  dans  $\mathcal{M}_{2m,2}(\mathbb{R})$  et telles que

$$\overline{H_{x,u_i}} = \begin{pmatrix} \overline{\partial_{x_1 x_1} u_i} & \overline{\partial_{x_1 x_2} u_i} \\ \overline{\partial_{x_2 x_1} u_i} & \overline{\partial_{x_2 x_2} u_i} \end{pmatrix}, \quad \forall i \in \mathcal{T}. \quad (5.29)$$

Comme ces hessiennes discrètes sont des matrices par bloc de  $\mathcal{M}_{2m,2}(\mathbb{R})$ , les conventions d'écritures de  $Q(\overline{H_{x,u_i}})$  et de  $\mathcal{L}(\overline{H_{x,u_i}})$  données par (5.20) sont maintenant utilisées et elles seront conservées dans la suite de ce chapitre. De plus, et par simplicité, l'ensemble  $\mathcal{M}_{2,1}(\mathbb{R})$  sera désormais assimilé à  $\mathbb{R}^2$ .

Avec les définitions ci-dessus, des approximations consistantes des gradients peuvent maintenant être énoncées. Ces approximations dépendent des matrices hessiennes discrètes  $(\overline{H_{x,u_i}})_{i \in \mathcal{T}}$  qui seront définies dans la section suivante.

**Lemme 5.2.2.1** (Gradients discrets). *Soit  $\mathcal{T}$  un maillage admissible au sens de la Définition 5.2.1. Pour  $u : \mathbb{R}^2 \rightarrow \mathbb{R}^m$  une fonction régulière donnée, considérons la séquence  $(u_i)_{i \in \mathcal{T}}$  dans  $\mathbb{R}^m$  définie par*

$$u_i = \frac{1}{|C_i|} \int_{C_i} u(x) \, dx. \quad (5.30)$$

*Considérons ensuite  $(\overline{H_{x,u_i}})_{i \in \mathcal{T}}$  une approximation consistante au premier ordre de la matrice hessienne de  $u$  dans  $\mathcal{M}_{2m,2}(\mathbb{R})$ . Considérons enfin les distances  $d$ , les vecteurs  $b$  et les formes linéaires  $\mathcal{L}$  respectivement définis dans (5.17), (5.23a) et (5.19). En utilisant la convention d'écriture de  $\mathcal{L}_{ij}(\overline{H_{x,u_i}})$  énoncée dans (5.20b), des approximations consistantes des gradients de  $u$  sont données par*

$$\overline{\nabla_x u_i}^{O1} = \left( \overline{\partial_{x_1} u_i}^{O1} \quad \overline{\partial_{x_2} u_i}^{O1} \right)^\top = \sum_{j \in \mathcal{N}(i)} \frac{d_{j,ij} u_i + d_{i,ij} u_j}{d_{i,j}} \left( b_{ij1} \quad b_{ij2} \right)^\top, \quad (5.31a)$$

$$\overline{\nabla_x u_i}^{O2} = \left( \overline{\partial_{x_1} u_i}^{O2} \quad \overline{\partial_{x_2} u_i}^{O2} \right)^\top = \sum_{j \in \mathcal{N}(i)} \left( \frac{d_{j,ij} u_i + d_{i,ij} u_j}{d_{i,j}} + \mathcal{L}_{ij}(\overline{H_{x,u_i}}) \right) \left( b_{ij1} \quad b_{ij2} \right)^\top. \quad (5.31b)$$

*Démonstration.* Pour  $x$  dans  $\mathbb{R}^2$ , et en notant  $u^{(l)}(x)$  la  $l$ -ième composante du vecteur  $u(x)$ , le gradient et la matrice hessienne de  $u^{(l)}$  respectivement dans  $\mathbb{R}^2$  et dans  $\mathcal{M}_{2,2}(\mathbb{R})$  sont donnés par

$$\nabla_x u^{(l)}(x) = \left( \partial_{x_1} u^{(l)}(x) \quad \partial_{x_2} u^{(l)}(x) \right)^\top \in \mathbb{R}^2, \quad (5.32a)$$

$$H_{x,u^{(l)}(x)} = \begin{pmatrix} \partial_{x_1 x_1} u^{(l)}(x) & \partial_{x_1 x_2} u^{(l)}(x) \\ \partial_{x_2 x_1} u^{(l)}(x) & \partial_{x_2 x_2} u^{(l)}(x) \end{pmatrix} \in \mathcal{M}_{2,2}(\mathbb{R}). \quad (5.32b)$$

Avec ces notations, l'application du Théorème de Green-Gauss au gradient (5.32a) donne les égalités suivantes dans  $\mathbb{R}^2$  :

$$\begin{aligned} \frac{1}{|C_i|} \int_{C_i} \nabla_x u^{(l)} \, dx &= \frac{1}{|C_i|} \int_{C_i} \left( \partial_{x_1} u^{(l)} \quad \partial_{x_2} u^{(l)} \right)^\top \, dx, \\ &= \left( \frac{1}{|C_i|} \int_{C_i} \partial_{x_1} u^{(l)} \, dx \quad \frac{1}{|C_i|} \int_{C_i} \partial_{x_2} u^{(l)} \, dx \right)^\top, \\ &= \left( \frac{1}{|C_i|} \sum_{j \in \mathcal{N}(i)} \int_{\Gamma_{ij}} u^{(l)}(x(\xi)) n_{ij1} \, d\xi, \quad \frac{1}{|C_i|} \sum_{j \in \mathcal{N}(i)} \int_{\Gamma_{ij}} u^{(l)}(x(\xi)) n_{ij2} \, d\xi \right)^\top, \\ &= \frac{1}{|C_i|} \sum_{j \in \mathcal{N}(i)} \left( \int_{\Gamma_{ij}} u^{(l)}(x(\xi)) \, d\xi \right) n_{ij}. \end{aligned} \quad (5.33)$$

En utilisant la définition de  $Q_{ij}^\alpha$  explicitée dans (5.18), un développement de Taylor au voisinage de  $x_{ij}$  permet d'écrire

$$\begin{aligned} \int_{\Gamma_{ij}} u^{(l)}(x(\xi)) n_{ij} \, d\xi &= \int_{\Gamma_{ij}} \left( u^{(l)}(x_{ij}) + (x(\xi) - x_{ij})^\top \nabla_x u^{(l)}(x_{ij}) \right) n_{ij} \, d\xi \\ &\quad + \frac{1}{2} \int_{\Gamma_{ij}} (x(\xi) - x_{ij})^\top H_{x,u^{(l)}(x_{ij})} (x(\xi) - x_{ij}) \, d\xi n_{ij} + \mathcal{O}(h^4), \\ &= |\Gamma_{ij}| \left( u^{(l)}(x_{ij}) + (\bar{x}_{ij} - x_{ij})^\top \nabla_x u^{(l)}(x_{ij}) \right) n_{ij} \\ &\quad + \frac{|\Gamma_{ij}|}{2} Q_{ij}^{\Gamma_{ij}}(H_{x,u^{(l)}(x_{ij})}) n_{ij} + \mathcal{O}(h^4). \end{aligned} \quad (5.34)$$

L'objectif est maintenant de réécrire les quantités du membre de droite de l'équation ci-dessus avec des approximations consistantes définies à l'aide de la séquence  $(u_i)_{i \in \mathcal{T}}$ . Cette réécriture s'effectue en considérant successivement les quantités réelles  $u^{(l)}(x_{ij})$  puis  $(\bar{x}_{ij} - x_{ij})^\top \nabla_x u^{(l)}(x_{ij})$  et enfin  $Q_{ij}^{\Gamma_{ij}}(H_{x,u^{(l)}(x_{ij})})$ .

Tout d'abord, pour approcher  $u^{(l)}(x_{ij})$ , des développements de Taylor usuels ainsi que les définitions de  $Q_{ij}^\alpha$  et de  $u_i$  respectivement données dans (5.18) et dans (5.30) permettent d'écrire

$$u_i^{(l)} = u^{(l)}(x_{ij}) + (x_i - x_{ij})^\top \nabla_x u^{(l)}(x_{ij}) + \frac{1}{2} Q_{ij}^{C_i}(H_{x,u^{(l)}}(x)) + \mathcal{O}(h^3), \quad (5.35a)$$

$$u_j^{(l)} = u^{(l)}(x_{ij}) + (x_j - x_{ij})^\top \nabla_x u^{(l)}(x_{ij}) + \frac{1}{2} Q_{ij}^{C_j}(H_{x,u^{(l)}}(x)) + \mathcal{O}(h^3). \quad (5.35b)$$

En multipliant l'égalité (5.35a) (*resp.* (5.35b)) par  $d_{j,ij}/d_{i,j}$  (*resp.* par  $d_{i,ij}/d_{i,j}$ ) puis en sommant les résultats, le terme multiplicatif de  $\nabla_x u^{(l)}(x_{ij})$  s'annule et il suit

$$\frac{d_{j,ij}u_i^{(l)} + d_{i,ij}u_j^{(l)}}{d_{i,j}} = u^{(l)}(x_{ij}) + \frac{d_{i,ij}}{2d_{i,j}} Q_{ij}^{C_j}(H_{x,u^{(l)}}(x)) + \frac{d_{j,ij}}{2d_{i,j}} Q_{ij}^{C_i}(H_{x,u^{(l)}}(x)) + \mathcal{O}(h^3).$$

D'après sa définition (5.18),  $Q_{ij}^\alpha$  est d'ordre  $\mathcal{O}(h^2)$  et par hypothèse,  $\overline{H_{x,u_i}}$  est une approximation consistante au premier ordre de la matrice hessienne de  $u$ . Par conséquent, l'équation ci-dessus peut se récrire

$$u^{(l)}(x_{ij}) = \frac{d_{j,ij}u_i^{(l)} + d_{i,ij}u_j^{(l)}}{d_{i,j}} - \frac{d_{i,ij}}{2d_{i,j}} Q_{ij}^{C_j}(\overline{H_{x,u_{ij}^{(l)}}}) - \frac{d_{j,ij}}{2d_{i,j}} Q_{ij}^{C_i}(\overline{H_{x,u_i^{(l)}}}) + \mathcal{O}(h^3). \quad (5.36)$$

Cette dernière égalité donne la réécriture de  $u^{(l)}(x_{ij})$  souhaitée.

Ensuite, pour approcher les quantités  $(\bar{x}_{ij} - x_{ij})^\top \nabla_x u^{(l)}(x_{ij})$  et  $Q_{ij}^{\Gamma_{ij}}(H_{x,u^{(l)}}(x_{ij}))$ , des développements de Taylor, des arguments de consistances comparables et la formule du point milieu, permettent de déduire

$$\begin{aligned} & (\bar{x}_{ij} - x_{ij})^\top \nabla_x u^{(l)}(x_{ij}) \\ &= (\bar{x}_{ij} - x_{ij})^\top \nabla_x u^{(l)}(x_i) + (\bar{x}_{ij} - x_{ij})^\top H_{x,u^{(l)}}(x_i) (x_{ij} - x_i) + \mathcal{O}(h^3), \\ &= (\bar{x}_{ij} - x_{ij})^\top \left( \frac{1}{|C_i|} \int_{C_i} \nabla_x u^{(l)} dx \right) + (\bar{x}_{ij} - x_{ij})^\top \overline{H_{x,u_i^{(l)}}} (x_{ij} - x_i) + \mathcal{O}(h^3), \end{aligned} \quad (5.37)$$

$$Q_{ij}^{\Gamma_{ij}}(H_{x,u^{(l)}}(x_{ij})) = Q_{ij}^{\Gamma_{ij}}(\overline{H_{x,u_i^{(l)}}}) + \mathcal{O}(h^3). \quad (5.38)$$

L'utilisation des estimations (5.36), (5.37) et (5.38) dans le membre de droite de (5.34) permet de déduire la réécriture de l'égalité (5.34) attendue. En s'appuyant sur l'expression de la forme linéaire  $\mathcal{L}_{ij}$  définie par (5.19), et comme  $|\Gamma_{ij}| = \mathcal{O}(h)$ , cette réécriture est donnée par

$$\begin{aligned} \int_{\Gamma_{ij}} u^{(l)}(x(\xi)) n_{ij} d\xi &= |\Gamma_{ij}| (\bar{x}_{ij} - x_{ij})^\top \left( \frac{1}{|C_i|} \int_{C_i} \nabla_x u^{(l)} dx \right) n_{ij} \\ &+ |\Gamma_{ij}| \left( \frac{d_{j,ij}u_i^{(l)} + d_{i,ij}u_j^{(l)}}{d_{i,j}} + \mathcal{L}_{ij}(\overline{H_{x,u_i^{(l)}}}) \right) n_{ij} + \mathcal{O}(h^4). \end{aligned}$$

Puisque l'égalité  $(a^\top b)c = (ca^\top)b$  est satisfaite pour tout vecteur  $(a, b, c)$  de  $(\mathbb{R}^2)^3$ , l'estimation ci-dessus se réécrit également sous la forme

$$\begin{aligned} \int_{\Gamma_{ij}} u^{(l)}(x(\xi)) n_{ij} d\xi &= |\Gamma_{ij}| n_{ij} (\bar{x}_{ij} - x_{ij})^\top \left( \frac{1}{|C_i|} \int_{C_i} \nabla_x u^{(l)} dx \right) \\ &+ |\Gamma_{ij}| \left( \frac{d_{j,ij}u_i^{(l)} + d_{i,ij}u_j^{(l)}}{d_{i,j}} + \mathcal{L}_{ij}(\overline{H_{x,u_i^{(l)}}}) \right) n_{ij} + \mathcal{O}(h^4). \end{aligned}$$

En introduisant le résultat ci-dessus dans l'égalité (5.33) et en utilisant la définition des matrices  $M_i$  données par (5.21), il suit

$$M_i \left( \frac{1}{|C_i|} \int_{C_i} \nabla_x u^{(l)} dx \right) = \sum_{j \in \mathcal{N}(i)} \left( \frac{d_{j,ij} u_i^{(l)} + d_{i,ij} u_j^{(l)}}{d_{i,j}} + \mathcal{L}_{ij}(\overline{H_{x,u_i}}) \right) \left( \frac{|\Gamma_{ij}|}{|C_i|} n_{ij} \right) + \mathcal{O}(h^2). \quad (5.39)$$

Par hypothèse du maillage admissible,  $M_i^{-1}$  existe et elle satisfait  $M_i^{-1} = \mathcal{O}(1)$ . Par conséquent, en multipliant à gauche l'équation (5.39) par  $M_i^{-1}$ , on en déduit le résultat (5.31b) relatif aux gradients discrets du second ordre.

Concernant les gradients discrets du premier ordre (5.31a), d'après les définitions (5.19) et (5.23a), il est clair que  $\mathcal{L}_{ij}$  et  $b_{ij}$  vérifient respectivement

$$\mathcal{L}_{ij} = \mathcal{O}(h^2), \quad b_{ij} = \mathcal{O}(h^{-1}).$$

L'utilisation des estimations ci-dessus dans l'expression (5.31b) donne le résultat attendu quant aux gradients discrets du premier ordre (5.31a) ce qui conclut la preuve.  $\square$

Pour une séquence  $(u_i)_{i \in \mathcal{T}}$  donnée, le lemme ci-dessus propose dans (5.31a) des gradients discrets locaux, linéaires, explicites, sans technique de limitation additionnelle et qui sont consistants au premier ordre.

Les gradients discrets (5.31a) sont des versions dégradées des approximations du second ordre données par (5.31b). Ces approximations du second ordre demandent la connaissance des matrices hessiennes discrètes  $(\overline{H_{x,u_i}})_{i \in \mathcal{T}}$  et la définition de ces objets est maintenant détaillée.

### 5.2.3 Matrices hessiennes discrètes du premier ordre

Cette section s'intéresse à l'approximation de la matrice hessienne d'une fonction régulière  $u$  donnée. Le premier résultat est une équation linéaire vérifiée par la quantité  $\overline{H_{x,u_i}}$  qui désigne une approximation consistante d'ordre un.

**Lemme 5.2.3.1.** *Soit  $\mathcal{T}$  un maillage admissible qui vérifie la Définition 5.2.1. Pour une fonction régulière  $u : \mathbb{R}^2 \rightarrow \mathbb{R}^m$  donnée, notons  $\overline{H_{x,u_i}}$  dans  $\mathcal{M}_{2m,2}(\mathbb{R})$  une approximation consistante au premier ordre de la matrice hessienne de  $u$  à la cellule  $C_i$ . Considérons ensuite la séquence  $(u_i)_{i \in \mathcal{T}}$ , les distances  $d$ , les vecteurs  $b$  et les formes linéaires  $\mathcal{L}$  respectivement définis dans (5.30), (5.17), (5.23a) et (5.19). En utilisant les conventions d'écritures de  $\mathcal{L}_{ij}(\overline{H_{x,u_i}})$  et des sommes doubles respectivement énoncées dans (5.20b) et dans (5.25),  $\overline{H_{x,u_i}}$  vérifie*

$$\begin{aligned} \overline{H_{x,u_i}} - \sum_{(j,r) \in \mathcal{N}(i,i)} \frac{d_{j,ij}}{d_{i,j}} \mathcal{L}_{ir}(\overline{H_{x,u_i}}) b_{ir} b_{ij}^\top - \sum_{(j,p) \in \mathcal{N}(i,j)} \frac{d_{i,ij}}{d_{i,j}} \mathcal{L}_{jp}(\overline{H_{x,u_i}}) b_{jp} b_{ij}^\top \\ = \sum_{(j,r) \in \mathcal{N}(i,i)} \frac{d_{j,ij}}{d_{i,j}} \frac{d_{r,ir} u_i + d_{i,ir} u_r}{d_{i,r}} b_{ir} b_{ij}^\top + \sum_{(j,p) \in \mathcal{N}(i,j)} \frac{d_{i,ij}}{d_{i,j}} \frac{d_{p,jp} u_j + d_{j,jp} u_p}{d_{j,p}} b_{jp} b_{ij}^\top, \quad \forall i \in \mathcal{T}. \end{aligned} \quad (5.40)$$

*Démonstration.* Pour  $x$  dans  $\mathbb{R}^2$ , et en notant  $u^{(l)}(x)$  la  $l$ -ième composante du vecteur  $u(x)$ , l'application directe du Théorème de Green-Gauss donne

$$\frac{1}{|C_i|} \int_{C_i} \partial_{x_\mu} \partial_{x_\nu} u^{(l)}(x) dx = \frac{1}{|C_i|} \sum_{j \in \mathcal{N}(i)} \int_{\Gamma_{ij}} \partial_{x_\nu} u^{(l)}(x(\xi)) d\xi n_{ij\mu}, \quad \forall (\mu, \nu) \in \{1, 2\}^2.$$

En utilisant la notation de la matrice hessienne de  $u^{(l)}$  définie dans (5.32b), les quatre égalités ci-dessus se récrivent dans  $\mathcal{M}_{2,2}(\mathbb{R})$  de la manière suivante :

$$\frac{1}{|C_i|} \int_{C_i} H_{x,u^{(l)}(x)} dx = \frac{1}{|C_i|} \sum_{j \in \mathcal{N}(i)} \left( \int_{\Gamma_{ij}} \nabla_x u^{(l)}(x(\xi)) d\xi \right) n_{ij}^\top. \quad (5.41)$$

Comme  $|\Gamma_{ij}|d(\bar{x}_{ij}, x_{ij}) = \mathcal{O}(h^2)$ , un développement de Taylor autour de  $x_{ij}$  permet d'écrire

$$\begin{aligned} \int_{\Gamma_{ij}} \nabla_x u^{(l)}(x(\xi)) \, d\xi &= |\Gamma_{ij}| \left( \nabla_x u^{(l)}(x_{ij}) + H_{x,u^{(l)}(x_{ij})}(\bar{x}_{ij} - x_{ij}) \right) + \mathcal{O}(h^3). \\ &= |\Gamma_{ij}| \nabla_x u^{(l)}(x_{ij}) + |\Gamma_{ij}| \left( \frac{1}{|C_i|} \int_{C_i} H_{x,u^{(l)}(x)} \, dx \right) (\bar{x}_{ij} - x_{ij}) + \mathcal{O}(h^3). \end{aligned}$$

En introduisant l'égalité ci-dessus dans l'égalité (5.41), il suit dans  $\mathcal{M}_{2,2}(\mathbb{R})$

$$\left( \frac{1}{|C_i|} \int_{C_i} H_{x,u^{(l)}(x)} \, dx \right) \left( I_2 - \frac{1}{|C_i|} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| (\bar{x}_{ij} - x_{ij}) n_{ij}^\top \right) = \frac{1}{|C_i|} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \nabla_x u^{(l)}(x_{ij}) n_{ij}^\top + \mathcal{O}(h).$$

Comme le maillage  $\mathcal{T}$  vérifie la Définition 5.2.1-*i*), il est possible d'utiliser les matrices  $M_i$  et les vecteurs  $(b_{ij})_{j \in \mathcal{N}(i)}$  respectivement donnés dans (5.21) et dans (5.23a). Avec ces quantités, l'équation ci-dessus posée dans  $\mathcal{M}_{2,2}(\mathbb{R})$  se reformule

$$\left( \frac{1}{|C_i|} \int_{C_i} H_{x,u^{(l)}(x)} \, dx \right) M_i^\top = \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \nabla_x u^{(l)}(x_{ij}) (M_i b_{ij})^\top + \mathcal{O}(h).$$

Par hypothèse (5.22),  $M_i^{-1}$  existe et d'après (5.22), elle vérifie  $M_i^{-1} = \mathcal{O}(1)$ . Par conséquent, en multipliant à droite l'équation ci-dessus par  $M_i^{-\top}$ , il suit

$$\frac{1}{|C_i|} \int_{C_i} H_{x,u^{(l)}(x)} \, dx = \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \nabla_x u^{(l)}(x_{ij}) b_{ij}^\top + \mathcal{O}(h). \quad (5.42)$$

Il est désormais nécessaire de récrire  $\nabla_x u^{(l)}(x_{ij})$  avec des approximations consistantes définies à l'aide de la séquence  $(u_i)_{i \in \mathcal{T}}$ . En considérant les gradients discrets du second ordre donnés dans (5.31b), des développements de Taylor usuels permettent de déduire

$$\nabla_x u^{(l)}(x_{ij}) = \frac{\overline{d_{j,ij} \nabla_x u_j^{(l)}}^{O2} + \overline{d_{j,ij} \nabla_x u_i^{(l)}}^{O2}}{d_{i,j}} + \mathcal{O}(h^2).$$

Ainsi, en utilisant la convention d'écriture des sommes doubles énoncée dans (5.25) et puisque  $b_{ij}^\top = \mathcal{O}(h^{-1})$ , l'équation (5.42) se lit

$$\begin{aligned} \frac{1}{|C_i|} \int_{C_i} H_{x,u^{(l)}(x)} \, dx &= \sum_{j \in \mathcal{N}(i)} \frac{\overline{d_{j,ij} \nabla_x u_i^{(l)}}^{O2} + \overline{d_{i,ij} \nabla_x u_j^{(l)}}^{O2}}{d_{i,j}} b_{ij}^\top + \mathcal{O}(h), \\ &= \sum_{(j,r) \in \mathcal{N}(i,i)} \frac{d_{j,ij}}{d_{i,j}} \left( \frac{d_{r,ir} u_i^{(l)} + d_{i,ir} u_r^{(l)}}{d_{i,r}} + \mathcal{L}_{ir}(\overline{H_{x,u_i^{(l)}}}) \right) b_{ir} b_{ij}^\top \\ &\quad + \sum_{(j,p) \in \mathcal{N}(i,j)} \frac{d_{i,ij}}{d_{i,j}} \left( \frac{d_{p,jp} u_j^{(l)} + d_{j,jp} u_p^{(l)}}{d_{j,p}} + \mathcal{L}_{jp}(\overline{H_{x,u_j^{(l)}}}) \right) b_{jp} b_{ij}^\top + \mathcal{O}(h). \end{aligned} \quad (5.43)$$

D'après l'hypothèse du Lemme 5.2.2.1, les quantités  $(\overline{H_{x,u_\alpha^{(l)}}})_{\alpha \in \{i,j\}}$  dans l'équation ci-dessus sont consistantes au premier ordre avec la matrice hessienne de  $u$ . De plus, il est clair que la forme linéaire  $\mathcal{L}$  définie par (5.19) est d'ordre  $\mathcal{O}(h^2)$ . En conséquence, il suit

$$\mathcal{L}(\overline{H_{x,u_\alpha^{(l)}}}) = \mathcal{L} \left( \frac{1}{|C_i|} \int_{C_i} H_{x,u^{(l)}(x)} \, dx \right) + \mathcal{O}(h^3), \quad \forall \alpha \in \{i,j\}.$$

Puisque l'introduction de l'estimation ci-dessus dans l'équation (5.43) permet de déduire le résultat (5.40), ceci termine la preuve.  $\square$

Avec la convention d'écriture de  $\mathcal{L}_{ij}(\overline{H_{x,u_i}})$  donnée dans (5.20b), l'équation (5.40) s'identifie à l'équation (5.24) qui par hypothèse de maillage admissible admet une unique solution. Par conséquent, d'après le lemme ci-dessus, l'équation (5.40) définit une approximation de la matrice hessienne de  $u$ . Cette approximation est consistante au premier ordre sur maillage admissible quelconque et d'après l'expression (5.27), elle dégénère vers une approximation du second ordre sur grilles cartésiennes.

En s'appuyant sur la linéarité de l'équation (5.40), il est maintenant possible de proposer une expression des matrices hessiennes discrètes  $(\overline{H_{x,u_i}})_{i \in \mathcal{T}}$ . Cette expression ainsi qu'une propriété additionnelle sont maintenant établies.

**Lemme 5.2.3.2** (Matrices hessiennes discrètes). *Soit  $\mathcal{T}$  un maillage admissible qui vérifie la Définition 5.2.1. Pour une fonction régulière  $u : \mathbb{R}^2 \rightarrow \mathbb{R}^m$  donnée, considérons la séquence  $(\overline{H_{x,u_i}})_{i \in \mathcal{T}}$  dans  $\mathcal{M}_{2m,2}(\mathbb{R})$  définie par l'équation (5.40). Pour chaque cellule  $C_i$  dans  $\mathcal{T}$  définie par ses interfaces orientées  $(\Gamma_{ij})_{j \in \mathcal{N}(i)}$ , les assertions suivantes sont vérifiées.*

i) *Pour chaque couple  $(C_i, \Gamma_{ij})$  fixé, il existe deux séquences de matrices  $(B_{ir,ij})_{r \in \mathcal{N}(i)}$  et  $(B_{jp,ij})_{p \in \mathcal{N}(j)}$  toutes deux dans  $\mathcal{M}_{2,2}(\mathbb{R})$  et telles que*

$$\overline{H_{x,u_i}} = \sum_{(j,r) \in \mathcal{N}(i,i)} \frac{d_{r,ir}u_i + d_{i,ir}u_r}{d_{i,r}} B_{ir,ij} + \sum_{(j,p) \in \mathcal{N}(i,j)} \frac{d_{p,jp}u_j + d_{j,jp}u_p}{d_{j,p}} B_{jp,ij}, \quad \forall i \in \mathcal{T}, \quad (5.44)$$

ii) *Pour chaque couple  $(C_i, \Gamma_{ij})$  fixé, les séquences de matrices  $(B_{ir,ij})_{r \in \mathcal{N}(i)}$  et  $(B_{jp,ij})_{p \in \mathcal{N}(j)}$  ne dépendent que de la géométrie du maillage  $\mathcal{T}$  et elles vérifient de plus*

$$\sum_{r \in \mathcal{N}(\alpha)} B_{\alpha r,ij} = 0, \quad \forall \alpha \in \{i, j\}. \quad (5.45)$$

iii) *Pour  $Q : \mathcal{M}_{2,2}(\mathbb{R}) \rightarrow \mathbb{R}$  une forme linéaire quelconque, la convention d'écriture de  $Q(\overline{H_{x,u_i}})$  définie par (5.20a) et appliquée à l'expression (5.44) se lit dans  $\mathbb{R}^m$*

$$Q(\overline{H_{x,u_i}}) = \sum_{(j,r) \in \mathcal{N}(i,i)} \frac{d_{r,ir}u_i + d_{i,ir}u_r}{d_{i,r}} Q(B_{ir,ij}) + \sum_{(j,p) \in \mathcal{N}(i,j)} \frac{d_{p,jp}u_j + d_{j,jp}u_p}{d_{j,p}} Q(B_{jp,ij}). \quad (5.46)$$

*Démonstration.* Pour montrer l'assertion i), il est d'abord nécessaire de reformuler l'équation (5.40). Pour une cellule  $C_i$  donnée et d'après cette équation, la matrice hessienne discrète de la  $l$ -ième composante de  $u$  satisfait dans  $\mathcal{M}_{2,2}(\mathbb{R})$

$$\begin{aligned} \overline{H_{x,u_i}^{(l)}} - \sum_{(j,r) \in \mathcal{N}(i,i)} \frac{d_{j,ij}}{d_{i,j}} \mathcal{L}_{ir}(\overline{H_{x,u_i}^{(l)}}) b_{ir} b_{ij}^\top - \sum_{(j,p) \in \mathcal{N}(i,j)} \frac{d_{i,ij}}{d_{i,j}} \mathcal{L}_{jp}(\overline{H_{x,u_i}^{(l)}}) b_{jp} b_{ij}^\top \\ = \sum_{(j,r) \in \mathcal{N}(i,i)} \frac{d_{j,ij}}{d_{i,j}} \frac{d_{r,ir}u_i^{(l)} + d_{i,ir}u_r^{(l)}}{d_{i,r}} b_{ir} b_{ij}^\top + \sum_{(j,p) \in \mathcal{N}(i,j)} \frac{d_{i,ij}}{d_{i,j}} \frac{d_{p,jp}u_j^{(l)} + d_{j,jp}u_p^{(l)}}{d_{j,p}} b_{jp} b_{ij}^\top. \end{aligned} \quad (5.47)$$

Comme  $\mathcal{L}_{ij}$  défini dans (5.19) est une forme linéaire, il existe des réels  $(L_{ij,\mu\nu})_{\mu,\nu \in \{1,2\}^2}$  tels que

$$\mathcal{L}_{ij}(A) = \sum_{(\mu,\nu) \in \{1,2\}^2} L_{ij,\mu\nu} A_{\mu,\nu}, \quad \forall A \in \mathcal{M}_{2,2}(\mathbb{R}).$$

Par conséquent, l'équation (5.47) d'inconnus réels  $(\overline{\partial_{x_{\mu\nu}} u_i^{(l)}})_{(\mu,\nu) \in \{1,2\}^2}$  se réécrit

$$\overline{H_{x,u_i}^{(l)}} - \sum_{\mu,\nu \in \{1,2\}^2} \left( \sum_{(j,r) \in \mathcal{N}(i,i)} \frac{d_{j,ij}}{d_{i,j}} L_{ir,\mu\nu} b_{ir} b_{ij}^\top + \sum_{(j,p) \in \mathcal{N}(i,j)} \frac{d_{i,ij}}{d_{i,j}} L_{jp,\mu\nu} b_{jp} b_{ij}^\top \right) \overline{\partial_{x_{\mu\nu}} u_i^{(l)}}$$

$$= \sum_{(j,r) \in \mathcal{N}(i,i)} \frac{d_{j,ij}}{d_{i,j}} \frac{d_{r,ir}u_i^{(l)} + d_{i,ir}u_r^{(l)}}{d_{i,r}} b_{ir}b_{ij}^\top + \sum_{(j,p) \in \mathcal{N}(i,j)} \frac{d_{i,ij}}{d_{i,j}} \frac{d_{p,jp}u_j^{(l)} + d_{j,jp}u_p^{(l)}}{d_{j,p}} b_{jp}b_{ij}^\top. \quad (5.48)$$

L'équation ci-dessus est linéaire et elle impose une égalité dans  $\mathcal{M}_{2,2}(\mathbb{R})$ . Par conséquent, cette équation est équivalente à la résolution d'un système linéaire de taille quatre et l'objectif est maintenant d'exhiber ce système. Pour ce faire, plusieurs notations et définitions sont introduites.

La première définition concerne une application  $\Phi : \mathcal{M}_{2,2}(\mathbb{R}) \rightarrow \mathbb{R}^4$  telle que

$$\Phi : \left| \begin{array}{l} \mathcal{M}_{2,2}(\mathbb{R}) \rightarrow \mathbb{R}^4, \\ \left( \begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \end{array} \right) \mapsto (a_{11}, a_{12}, a_{21}, a_{22})^\top. \end{array} \right. \quad (5.49)$$

D'après son expression ci-dessus, il est clair que  $\Phi$  est linéaire, bijective et elle permet la notation  $X_i^{(l)} = \Phi(\overline{H_{x,u_i^{(l)}}})$  dans  $\mathbb{R}^4$ . En s'appuyant sur la bijectivité de  $\Phi$  et par abus de notation sur les indices  $(\mu, \nu)$  dans  $\{1, 2\}^2$ , la convention d'écriture suivante est adoptée :

$$(l_1, l_2, l_3, l_4) = (11, 12, 21, 22). \quad (5.50)$$

Enfin, des matrices  $A_{i,\mu\nu}$  dans  $\mathcal{M}_{2,2}(\mathbb{R})$  sont considérées et elles sont données par

$$A_{i,\mu\nu} = \sum_{(j,r) \in \mathcal{N}(i,i)} \frac{d_{j,ij}}{d_{i,j}} L_{ir,\mu\nu} b_{ir}b_{ij}^\top + \sum_{(j,p) \in \mathcal{N}(i,j)} \frac{d_{i,ij}}{d_{i,j}} L_{jp,\mu\nu} b_{jp}b_{ij}^\top.$$

Avec les matrices ci-dessus et les notations  $(l_k)_{k \in \{1, \dots, 4\}}$ , l'application de la fonction  $\Phi$  définie par (5.49) à l'équation (5.48) donne dans  $\mathbb{R}^4$

$$\begin{aligned} & \sum_{(j,r) \in \mathcal{N}(i,i)} \frac{d_{j,ij}}{d_{i,j}} \frac{d_{r,ir}u_i^{(l)} + d_{i,ir}u_r^{(l)}}{d_{i,r}} \Phi(b_{ir}b_{ij}^\top) + \sum_{(j,p) \in \mathcal{N}(i,j)} \frac{d_{i,ij}}{d_{i,j}} \frac{d_{p,jp}u_j^{(l)} + d_{j,jp}u_p^{(l)}}{d_{j,p}} \Phi(b_{jp}b_{ij}^\top) \\ &= X_i^{(l)} - \sum_{\mu, \nu \in \{1, 2\}^2} \Phi(A_{i,\mu\nu}) \overline{\partial_{x_\mu x_\nu} u_i^{(l)}}, \\ &= X_i^{(l)} - \sum_{k=1}^4 \Phi(A_{i,l_k}) (X_i^{(l)})_k. \end{aligned}$$

En définissant une matrice  $A_i$  dans  $\mathcal{M}_{4,4}(\mathbb{R})$  telle que sa  $k$ -ième colonne est donnée par  $\Phi(A_{i,l_k})$ , et en notant  $I_4$  la matrice identité de taille quatre, on en déduit que l'équation (5.48) est équivalente à

$$(I_4 - A_i) X_i^{(l)} = \sum_{(j,m) \in \mathcal{N}(i,i)} \frac{d_{j,ij}}{d_{i,j}} \frac{d_{m,im}u_i^{(l)} + d_{i,im}u_m^{(l)}}{d_{i,m}} \Phi(b_{im}b_{ij}^\top) + \sum_{(j,p) \in \mathcal{N}(i,j)} \frac{d_{i,ij}}{d_{i,j}} \frac{d_{p,jp}u_j^{(l)} + d_{j,jp}u_p^{(l)}}{d_{j,p}} \Phi(b_{jp}b_{ij}^\top). \quad (5.51)$$

Comme l'équation ci-dessus est une réécriture équivalente de l'équation (5.24) qui admet, par hypothèse de maillage admissible, une unique solution, la matrice  $I_4 - A_i$  est nécessairement inversible. Par conséquent il suit

$$\begin{aligned} X_i^{(l)} &= \sum_{(j,r) \in \mathcal{N}(i,i)} \frac{d_{j,ij}}{d_{i,j}} \frac{d_{r,ir}u_i^{(l)} + d_{i,ir}u_r^{(l)}}{d_{i,r}} (I_4 - A_i)^{-1} \Phi(b_{ir}b_{ij}^\top) \\ &+ \sum_{(j,p) \in \mathcal{N}(i,j)} \frac{d_{i,ij}}{d_{i,j}} \frac{d_{p,jp}u_j^{(l)} + d_{j,jp}u_p^{(l)}}{d_{j,p}} (I_4 - A_i)^{-1} \Phi(b_{jp}b_{ij}^\top). \end{aligned} \quad (5.52)$$

Pour montrer l'assertion *i*) avec cette dernière expression, et comme  $\Phi$  est bijective, il est suffisant de définir

$$B_{ir,ij} = \Phi^{-1}((I_4 - \mathbf{A}_i)^{-1}\Phi(b_{ir}b_{ij}^\top)), \quad B_{jp,ij} = \Phi^{-1}((I_4 - \mathbf{A}_i)^{-1}\Phi(b_{jp}b_{ij}^\top)). \quad (5.53)$$

Concernant l'assertion *ii*) et d'après la définition ci-dessus, il est clair que les matrices  $B_{\cdot,ij}$  ne dépendent que la géométrie du maillage. Pour la propriété (5.45), par linéarité de  $\Phi^{-1}$  puis de  $\Phi$ , il suit immédiatement

$$\begin{aligned} \sum_{r \in \mathcal{N}(\alpha)} B_{\alpha r,ij} &= \sum_{r \in \mathcal{N}(\alpha)} \Phi^{-1}((I_4 - \mathbf{A}_i)^{-1}\Phi(b_{\alpha r}b_{ij}^\top)), \\ &= \Phi^{-1}\left(\sum_{r \in \mathcal{N}(\alpha)} (I_4 - \mathbf{A}_i)^{-1}\Phi(b_{\alpha r}b_{ij}^\top)\right), \\ &= \Phi^{-1}((I_4 - \mathbf{A}_i)^{-1}\Phi\left(\sum_{r \in \mathcal{N}(\alpha)} b_{\alpha r}b_{ij}^\top\right)), \quad \forall \alpha \in \{i, j\}. \end{aligned} \quad (5.54)$$

D'après la définition de maillages admissible donnée en Section 5.2.1, Définition 5.2.1-*i*), les vecteurs  $(b_{\alpha r})_{r \in \mathcal{N}(\alpha)}$  vérifient

$$\sum_{r \in \mathcal{N}(\alpha)} b_{\alpha r} = 0, \quad \forall \alpha \in \{i, j\}. \quad (5.55)$$

L'utilisation de cette dernière égalité dans l'équation (5.54) permet de déduire le résultat (5.45).

L'assertion *iii*) est immédiate en appliquant la convention d'écriture (5.20a) à  $Q(\overline{H_{x,u_i}})$  avec  $\overline{H_{x,u_i}}$  donné par (5.44), puis en utilisant la linéarité de  $Q$ . Ces derniers arguments concluent la démonstration.  $\square$

D'après la démonstration ci-dessus, la condition d'unicité de la Définition 5.2.1-*ii*) est équivalente à l'existence de la matrice  $(I_4 - \mathbf{A}_i)^{-1}$  définie par l'équation (5.51). Par construction, la matrice  $\mathbf{A}_i$  ne dépend que des géométries voisines à  $C_i$  et les égalités (5.26) permettent de déduire que cette matrice est nulle dans le cas d'une grille cartésienne uniforme. Par conséquent, la matrice  $\mathbf{A}_i$  est une quantification de l'anisotropie des cellules. L'ensemble des maillages  $\mathcal{T}$  pour lesquels  $I_4 - \mathbf{A}_i$  est inversible pour toute cellule  $C_i$  est visiblement plus vaste que l'ensemble des grilles cartésienne de  $\mathbb{R}^2$ .

Lorsque l'existence de  $(I_4 - \mathbf{A}_i)^{-1}$  est vérifiée, le Lemme 5.2.3.2 donne dans (5.44) une expression des matrices hessiennes discrètes. Pour être mise en œuvre, cette expression nécessite néanmoins la connaissance des matrices  $B_{\cdot,ij}$  dans  $\mathcal{M}_{2,2}(\mathbb{R})$  et définies par (5.53). Seule une expression symbolique de ces matrices est donnée dans (5.53) mais la preuve du Lemme 5.2.3.2 est une stratégie de programmation possible.

Les gradients discrets (5.31) et les matrices hessiennes discrètes (5.44) sont linéaires, locaux, consistants et ils ne nécessitent pas de technique de limitation additionnelle. Ces quantités discrètes sont obtenues à l'aide du Théorème de Green-Gauss puis par des reformulations consistantes des développements de Taylor. De tels stratégies sont par exemple évoquées dans [28, 201] mais l'originalité des méthodes proposées tient à leurs formulations locales.

Ces formulations sont définies à l'aide des vecteurs  $(b_{ir})_{r \in \mathcal{N}(i)}$  et des matrices  $(B_{\alpha r,ij})_{(\alpha,r) \in \{i,j\} \times \mathcal{N}(\alpha)}$  et leur intérêt est double. D'abord,  $(b_{ir})_{r \in \mathcal{N}(i)}$  et  $(B_{\alpha r,ij})_{(\alpha,r) \in \{i,j\} \times \mathcal{N}(\alpha)}$  vérifient respectivement les propriétés (5.55) et (5.45) qui sont utilisables pour des démonstrations théoriques. Ensuite, sur un plan pratique, une méthode locale est en général plus simple à programmer que des méthodes globales qui couplent toutes les quantités  $u_i$ .

La première application directe des résultats (5.31) et (5.44) sont les approximations des dérivées des fonctions de flux  $f_1$  et  $f_2$ .

## 5.2.4 Gradients et matrices hessiennes discrets des fonctions de flux

En notant  $r$  dans  $\{1, 2\}$  et en considérant une fonction de flux régulière  $f_r : \mathbb{R}^m \rightarrow \mathbb{R}^m$ , cette section présente des approximations consistantes des dérivées de la fonction composée  $f_r(u) : \mathbb{R}^2 \rightarrow \mathbb{R}^m$ .



D'abord, la matrice jacobienne de  $u \mapsto f_r(u)$  est notée  $\nabla f_r(u)$  dans  $\mathcal{M}_{m,m}(\mathbb{R})$ . La  $l$ -ième composante de  $f_r$  est notée  $f_r^{(l)}$  et la matrice hessienne de  $u \mapsto f_r^{(l)}(u)$  est notée  $H_{f_r^{(l)}}$  dans  $\mathcal{M}_{m,m}(\mathbb{R})$ . Cette matrice hessienne est définie par ses coefficients  $(H_{f_r^{(l)}})_{\mu,\nu}$  qui s'écrivent

$$\left(H_{f_r^{(l)}}\right)_{\mu,\nu} = \frac{\partial^2 f_r^{(l)}}{\partial u_\mu \partial u_\nu}, \quad \forall (r, l, \mu, \nu) \in \{1, 2\} \times \{1, \dots, m\}^3. \quad (5.56)$$

Ensuite, pour  $x$  dans  $\mathbb{R}^2$ , les gradients discrets du premier et du second ordre de la fonction  $x \mapsto f_r(u(x))$  sont chacun désignés par un exposant  $Ok$  pour  $k$  dans  $\{1, 2\}$ . Ils s'écrivent dans  $\mathcal{M}_{2m,1}(\mathbb{R})$  de la manière suivante :

$$\overline{\nabla_x f_r(u)}_i^{Ok} = \left( \overline{\partial_{x_1} f_r(u)}_i^{Ok} \quad \overline{\partial_{x_2} f_r(u)}_i^{Ok} \right)^\top, \quad \forall i \in \mathcal{T}. \quad (5.57)$$

Les matrices hessiennes discrètes du premier ordre de  $x \mapsto f_r(u(x))$  sont notées  $\overline{H_{x,f_r}}_i$  dans  $\mathcal{M}_{2m,2}(\mathbb{R})$  et sont telles que

$$\overline{H_{x,f_r(u)}}_i = \begin{pmatrix} \overline{\partial_{x_1 x_1} f_r(u)}_i & \overline{\partial_{x_1 x_2} f_r(u)}_i \\ \overline{\partial_{x_2 x_1} f_r(u)}_i & \overline{\partial_{x_2 x_2} f_r(u)}_i \end{pmatrix}, \quad \forall i \in \mathcal{T}. \quad (5.58)$$

Enfin, les notations des gradients discrets et des matrices hessiennes discrètes de  $u$  définies dans la section 5.2.2 sont conservées et par conséquent ces objets se lisent respectivement

$$\overline{\nabla_x u}_i^{Ok} = \left( \overline{\partial_{x_1} u}_i^{Ok} \quad \overline{\partial_{x_2} u}_i^{Ok} \right)^\top \in \mathcal{M}_{2m,1}(\mathbb{R}), \quad (5.59a)$$

$$\overline{H_{x,u}}_i = \begin{pmatrix} \overline{\partial_{x_1 x_1} u}_i & \overline{\partial_{x_1 x_2} u}_i \\ \overline{\partial_{x_2 x_1} u}_i & \overline{\partial_{x_2 x_2} u}_i \end{pmatrix} \in \mathcal{M}_{2m,2}(\mathbb{R}). \quad (5.59b)$$

En considérant que les quantités  $\overline{\nabla_x u}_i^{Ok}$  et  $\overline{H_{x,u}}_i$  écrites ci-dessus sont respectivement données par (5.31) et par (5.44), des expressions de  $\overline{\nabla_x f_r(u)}_i^{Ok}$  et de  $\overline{H_{x,f_r(u)}}_i$  sont maintenant établies.

**Lemme 5.2.4.1** (Gradient et matrice hessienne discrets des fonctions de flux). *Soit  $\mathcal{T}$  un maillage admissible qui vérifie la Définition 5.2.1. Pour une fonction régulière  $u : \mathbb{R}^2 \rightarrow \mathbb{R}^m$  donnée, considérons les trois séquences  $(u_i)_{i \in \mathcal{T}}$ ,  $(\overline{\nabla_x u}_i^{Ok})_{i \in \mathcal{T}}$  et  $(\overline{H_{x,u}}_i)_{i \in \mathcal{T}}$  respectivement dans  $\mathbb{R}^m$ ,  $\mathcal{M}_{2m,1}(\mathbb{R})$  et dans  $\mathcal{M}_{2m,2}(\mathbb{R})$  et respectivement définies par (5.30), (5.31) et par (5.44). Pour deux fonctions de flux régulières  $f_1, f_2 : \mathbb{R}^m \rightarrow \mathbb{R}^m$  données et pour chaque cellule  $C_i$  dans  $\mathcal{T}$ , les assertions suivantes sont vérifiées.*

i) *Un gradient discret de  $x \mapsto f_r(u(x))$  d'ordre  $k$  est donné par*

$$\overline{\partial_{x_\mu} f_r(u)}_i^{Ok} = \nabla f_r(u_i) \overline{\partial_{x_\mu} u}_i^{Ok} \quad \forall (i, r, \mu, k) \in \mathcal{T} \times \{1, 2\}^3, \quad (5.60)$$

ii) *En considérant la séquence de matrices  $(H_{f_r^{(l)}})_{l \in \{1, \dots, m\}}$  dans  $\mathcal{M}_{m,m}(\mathbb{R})$  et définie par (5.56), une matrice hessienne discrète d'ordre un de  $x \mapsto f_r(u(x))$  s'écrit*

$$\begin{aligned} \overline{\partial_{x_\mu x_\nu} f_r(u)}_i &= \left( (\overline{\partial_{x_\mu} u}_i^{O1})^\top H_{f_r^{(1)}}(u_i) \overline{\partial_{x_\nu} u}_i^{O1}, \dots, (\overline{\partial_{x_\mu} u}_i^{O1})^\top H_{f_r^{(m)}}(u_i) \overline{\partial_{x_\nu} u}_i^{O1} \right)^\top \\ &+ \nabla f_r(u_i) \overline{\partial_{x_\mu x_\nu} u}_i, \quad \forall (i, r, \mu, \nu) \in \mathcal{T} \times \{1, 2\}^3. \end{aligned} \quad (5.61)$$

*Démonstration.* Concernant l'assertion i), la règle des dérivations composées et la formule du point milieu permettent d'écrire pour tout  $\mu$  dans  $\{1, 2\}$

$$\begin{aligned} \frac{1}{|C_i|} \int_{C_i} \partial_{x_\mu} f_r(u(x)) \, dx &= \frac{1}{|C_i|} \int_{C_i} \nabla f_r(u(x)) \partial_{x_\mu} u(x) \, dx, \\ &= \nabla f_r(u(x_i)) \left( \frac{1}{|C_i|} \int_{C_i} \partial_{x_\mu} u(x) \, dx \right) + \mathcal{O}(h^2), \\ &= \nabla f_r(u_i) \left( \frac{1}{|C_i|} \int_{C_i} \partial_{x_\mu} u(x) \, dx \right) + \mathcal{O}(h^2). \end{aligned}$$

Comme les gradients discrets  $\overline{\partial_{x_\mu} u_i}^{Ok}$  donnés par (5.31) sont consistants à l'ordre  $k$ , le résultat (5.60) se déduit de l'égalité ci-dessus.

Pour montrer l'assertion  $ii$ ), en utilisant la règle des dérivations composées, il suit pour tout  $(\mu, \nu)$  dans  $\{1, 2\}^2$

$$\begin{aligned} \frac{1}{|C_i|} \int_{C_i} \partial_{x_\mu x_\nu} f_r(u) \, dx &= \frac{1}{|C_i|} \int_{C_i} \partial_{x_\mu} (\nabla f_r(u) \partial_{x_\nu} u) \, dx, \\ &= \frac{1}{|C_i|} \int_{C_i} \left( \partial_{x_\mu} u^\top H_{f_r^{(1)}}(u) \partial_{x_\nu} u, \dots, \partial_{x_\mu} u^\top H_{f_r^{(m)}}(u) \partial_{x_\nu} u \right)^\top dx \\ &\quad + \frac{1}{|C_i|} \int_{C_i} \nabla f_r(u) \partial_{x_\mu x_\nu} u \, dx. \end{aligned} \tag{5.62}$$

Puisque les matrices hessiennes discrètes  $\overline{H_{x, u_i}}$  définies dans (5.44) sont consistantes au premier ordre, il est possible d'écrire

$$\begin{aligned} \frac{1}{|C_i|} \int_{C_i} \nabla f_r(u(x)) \partial_{x_\mu x_\nu} u(x) \, dx &= \nabla f_r(u(x_i)) \left( \frac{1}{|C_i|} \int_{C_i} \partial_{x_\mu x_\nu} u(x) \, dx \right) + \mathcal{O}(h^2), \\ &= \nabla f_r(u_i) \overline{\partial_{x_\mu x_\nu} u_i} + \mathcal{O}(h). \end{aligned} \tag{5.63}$$

De plus, la consistance d'ordre un des gradients discrets  $\overline{\nabla_x u_i}^{O1}$  définis dans (5.31a) donne pour tout  $l$  dans  $\{1, \dots, m\}$

$$\begin{aligned} \frac{1}{|C_i|} \int_{C_i} \partial_{x_\mu} u^\top H_{f_r^{(l)}}(u) \partial_{x_\nu} u \, dx &= (\overline{\partial_{x_\mu} u_i}^{O1})^\top \left( \frac{1}{|C_i|} \int_{C_i} H_{f_r^{(l)}}(u) \, dx \right) \overline{\partial_{x_\nu} u_i}^{O1} + \mathcal{O}(h), \\ &= (\overline{\partial_{x_\mu} u_i}^{O1})^\top H_{f_r^{(l)}}(u_i) \overline{\partial_{x_\nu} u_i}^{O1} + \mathcal{O}(h). \end{aligned}$$

L'introduction de l'estimation ci-dessus et de l'estimation (5.63) dans l'égalité (5.62) permet de déduire le résultat (5.61) ce qui conclut la preuve.  $\square$

Avec les approximations des dérivées établies dans cette section ainsi que dans les sections 5.2.2 et 5.2.3, il est désormais possible de proposer des flux numériques qui définissent des schémas du second et du troisième ordre sur maillages non structurés bi-dimensionnels.

## 5.3 Schémas non limités d'ordre deux et trois

### 5.3.1 Définitions et notations

Cette section propose des schémas numériques du second et du troisième ordre définis sur maillage non structurés bi-dimensionnels. D'après la Définition 5.1.5.1 qui introduit la méthode des volumes finis, de tels schémas sont entièrement définis par la donnée de deux flux numériques distincts. Ces flux numériques sont respectivement notés  $\mathcal{F}_{ij}^{O2}$  et  $\mathcal{F}_{ij}^{O3}$  dans  $\mathbb{R}^m$  et les schémas associés s'écrivent

$$w_i^{n+1} = w_i^n - \frac{\Delta t}{|C_i|} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \mathcal{F}_{ij}^{Ok}, \quad \forall (i, k) \in \mathcal{T} \times \{2, 3\}. \tag{5.64}$$

Pour  $k$  dans  $\{2, 3\}$  et pour un doublet  $(\mathfrak{F}_{ij}^{Ok}, A_{ij}^{Ok})$  dans  $(\mathbb{R}^m)^2$  complété pour un réel  $\lambda > 0$  donné, les flux numériques d'ordre élevé considérés sont de la forme

$$\mathcal{F}_{ij}^{Ok} = \mathfrak{F}_{ij}^{Ok} - \frac{\lambda}{2} (w_j^n - w_i^n) + A_{ij}^{Ok} d_{i,j}. \tag{5.65}$$

Ces flux numériques généralisent au cas des maillages non structurés les flux numériques (4.55) proposés dans le cadre d'une dimension d'espace. Ils se composent tout d'abord de la quantité  $\mathfrak{F}_{ij}^{Ok}$  qui doit

assurer une discrétisation d'ordre  $k$  des intégrales des fonctions de flux  $f_1$  et  $f_2$ . Ensuite, et pour ne pas dégrader cet ordre, la quantité  $-\frac{\lambda}{2}(w_j^n - w_i^n) + A_{ij}^{Ok} d_{i,j}$  doit être un résidu de la forme  $\mathcal{O}(h^k)$  avec  $h = \max_{ij \in \mathcal{E}} d_{i,j}$ . Enfin, les choix retenus pour le doublet  $(\mathfrak{F}_{ij}^{Ok}, A_{ij}^{Ok})$  doivent assurer le caractère conservatif du flux numérique (5.65) ce qui s'écrit

$$\mathcal{F}_{ij}^{Ok} = -\mathcal{F}_{ji}^{Ok}. \quad (5.66)$$

D'après la Définition 5.1.5.1-*i*), l'expression ci-dessus est nécessaire à la propriété de conservation.

Ainsi, et à partir de la définition du flux numérique  $\mathcal{F}_{ij}^{Ok}$  donnée par (5.65), l'objectif est désormais de déterminer les quantités  $\mathfrak{F}_{ij}^{Ok}$  et  $A_{ij}^{Ok}$  selon les conditions

$$\mathfrak{F}_{ij}^{Ok} = -\mathfrak{F}_{ji}^{Ok}, \quad A_{ij}^{Ok} = -A_{ji}^{Ok}, \quad (5.67a)$$

$$\mathfrak{F}_{ij}^{Ok} = \frac{1}{|\Gamma_{ij}|} \int_{\Gamma_{ij}} \left( f_1(w(x(\xi)), t^n) n_{ij1} + f_2(w(x(\xi)), t^n) n_{ij2} \right) d\xi + \mathcal{O}(h^k), \quad (5.67b)$$

$$A_{ij}^{Ok} d_{i,j} = \frac{\lambda}{2} (w_j^n - w_i^n) + \mathcal{O}(h^k). \quad (5.67c)$$

Les égalités (5.67a) sont imposées car elles assurent le caractère conservatif (5.66). Les estimations (5.67b)-(5.67c) sont retenues car elles garantissent l'ordre formel des schémas tel que donné dans la Définition 5.1.5.1-*iii*). Ainsi, un flux numérique (5.65) qui satisfait les conditions (5.67) donne un schéma numérique (5.64) conservatif et formellement d'ordre  $k$ .

La prochaine section concerne une définition de  $\mathfrak{F}_{ij}^{Ok}$  qui satisfait les conditions (5.67a)-(5.67b).

### 5.3.2 Discrétisation des intégrales des fonctions de flux

Cette section est consacrée à une expression de la quantités  $\mathfrak{F}_{ij}^{Ok}$  dans  $\mathbb{R}^m$  utilisée pour les flux numérique  $\mathcal{F}_{ij}^{Ok}$  définis dans (5.65). Pour assurer l'ordre des schémas, cette quantité  $\mathfrak{F}_{ij}^{Ok}$  doit satisfaire l'estimation (5.67b). Une formulation possible est maintenant proposée.

**Lemme 5.3.2.1** (Discrétisation des intégrales de flux de  $\mathfrak{F}_{ij}^{Ok}$ ). *Soient  $\mathcal{T}$  un maillage admissible qui vérifie la Définition 5.2.1, une fonction régulière  $u : \mathbb{R}^2 \rightarrow \mathbb{R}^m$  et deux fonctions de flux régulières  $f_1, f_2 : \mathbb{R}^m \rightarrow \mathbb{R}^m$ . Considérons les séquences  $(u_i)_{i \in \mathcal{T}}$ ,  $(\overline{\nabla_x f_r(u)})_{(i,r) \in \mathcal{T} \times \{1,2\}}^{Ok}$  et  $(\overline{H_{x,f_r(u)}})_{(i,r) \in \mathcal{T} \times \{1,2\}}$  respectivement dans  $\mathbb{R}^m$ ,  $\mathcal{M}_{2m,1}(\mathbb{R})$  et dans  $\mathcal{M}_{2m,2}(\mathbb{R})$  et respectivement définies par (5.30), (5.60) et par (5.61).*

*Pour une interface orientée  $\Gamma_{ij}$  fixée, considérons ensuite les distances  $d$ , les formes linéaires  $(Q_{ij}^\alpha)_{\alpha \in \{C_i, C_j, \Gamma_{ij}\}}$  respectivement données par (5.17), (5.18) ainsi que  $\overline{H_{x,f_r(u)}}$  dans  $\mathcal{M}_{2m,2}(\mathbb{R})$  défini de la manière suivante :*

$$\overline{H_{x,f_r(u)_{ij}}} = \frac{d_{j,ij} \overline{H_{x,f_r(u)_i}} + d_{i,ij} \overline{H_{x,f_r(u)_j}}}{d_{i,j}}, \quad \forall r \in \{1,2\}. \quad (5.68)$$

*Pour  $(\mathfrak{F}_{ij,r}^{Ok})_{r \in \{1,2\}}$  dans  $\mathbb{R}^m$ , considérons enfin  $\mathfrak{F}_{ij}^{Ok}$  sous la forme*

$$\mathfrak{F}_{ij}^{Ok} = \mathfrak{F}_{ij,1}^{Ok} n_{ij1} + \mathfrak{F}_{ij,2}^{Ok} n_{ij2}. \quad (5.69)$$

*En notant  $(\mathfrak{F}_{ij,r}^{Ok})^{(l)}$  la  $l$ -ième composante du vecteur  $\mathfrak{F}_{ij,r}^{Ok}$ , les trois assertions suivantes sont vérifiées.*

*i) Si pour tout  $(l,r)$  dans  $\{1, \dots, m\} \times \{1,2\}$*

$$(\mathfrak{F}_{ij,r}^{O2})^{(l)} = \frac{d_{j,ij} f_r^{(l)}(u_i) + d_{i,ij} f_r^{(l)}(u_j)}{d_{i,j}} + (\bar{x}_{ij} - x_{ij})^\top \frac{\overline{d_{j,ij} \nabla_x f_r^{(l)}(u)_i}^{O1} + d_{i,ij} \overline{\nabla_x f_r^{(l)}(u)_j}^{O1}}{d_{i,j}}, \quad (5.70)$$

*alors*

$$\mathfrak{F}_{ij}^{O2} = \frac{1}{|\Gamma_{ij}|} \int_{\Gamma_{ij}} \left( f_1(u(x(\xi))) n_{ij1} + f_2(u(x(\xi))) n_{ij2} \right) d\xi + \mathcal{O}(h^2).$$

ii) Si pour tout  $(l, r)$  dans  $\{1, \dots, m\} \times \{1, 2\}$

$$\begin{aligned} (\mathfrak{F}_{ij,r}^{O3})^{(l)} &= \frac{d_{j,ij} f_r^{(l)}(u_i) + d_{i,ij} f_r^{(l)}(u_j)}{d_{i,j}} + (\bar{x}_{ij} - x_{ij})^\top \frac{\overline{d_{j,ij} \nabla_x f_r^{(l)}(u)_i}^{O2} + d_{i,ij} \overline{\nabla_x f_r^{(l)}(u)_j}^{O2}}{d_{i,j}} \\ &\quad - \frac{d_{i,ij}}{2d_{i,j}} Q_{ij}^{C_j}(\overline{H_{x,f_r^{(l)}(u)}_{ij}}) - \frac{d_{j,ij}}{2d_{i,j}} Q_{ij}^{C_i}(\overline{H_{x,f_r^{(l)}(u)}_{ij}}) + \frac{1}{2} Q_{ij}^{\Gamma_{ij}}(\overline{H_{x,f_r^{(l)}(u)}_{ij}}), \end{aligned} \quad (5.71)$$

alors

$$\mathfrak{F}_{ij}^{O3} = \frac{1}{|\Gamma_{ij}|} \int_{\Gamma_{ij}} (f_1(u(x(\xi))) n_{ij_1} + f_2(u(x(\xi))) n_{ij_2}) d\xi + \mathcal{O}(h^3).$$

iii) Si  $\mathfrak{F}_{ij,r}^{Ok}$  est donné par l'expression (5.71) ou par (5.70), alors  $\mathfrak{F}_{ij}^{Ok}$  est conservatif ce qui s'écrit

$$\mathfrak{F}_{ij}^{Ok} = -\mathfrak{F}_{ji}^{Ok}.$$

*Démonstration.* Pour  $x$  dans  $\mathbb{R}^2$ , et en notant  $f_r^{(l)}(u(x))$  la  $l$ -ième composante du vecteur  $f_r(u(x))$ , le gradient et la matrice hessienne de  $x \mapsto f_r^{(l)}(u(x))$  respectivement dans  $\mathbb{R}^2$  et dans  $\mathcal{M}_{2,2}(\mathbb{R})$  sont donnés par

$$\begin{aligned} \nabla_x (f_r^{(l)}(u))(x) &= \left( \partial_{x_1} (f_r^{(l)}(u))(x) \quad \partial_{x_2} (f_r^{(l)}(u))(x) \right)^\top \in \mathbb{R}^2, \\ H_{x,f_r^{(l)}(u)}(x) &= \begin{pmatrix} \partial_{x_1 x_1} (f_r^{(l)}(u))(x) & \partial_{x_1 x_2} (f_r^{(l)}(u))(x) \\ \partial_{x_2 x_1} (f_r^{(l)}(u))(x) & \partial_{x_2 x_2} (f_r^{(l)}(u))(x) \end{pmatrix} \in \mathcal{M}_{2,2}(\mathbb{R}). \end{aligned}$$

Pour montrer l'assertion *i*), en utilisant la définition des formes linéaires  $(Q_{ij}^\alpha)_{\alpha \in \{\Gamma_{ij}, C_i, C_j\}}$  donnée dans (5.18), des développements de Taylor de  $x \mapsto f_r^{(l)}(u(x))$  au voisinage de  $x_{ij}$  permettent d'écrire

$$\begin{aligned} &\frac{1}{|\Gamma_{ij}|} \int_{\Gamma_{ij}} f_r^{(l)}(u(x(\xi))) d\xi \\ &= \frac{1}{|\Gamma_{ij}|} \int_{\Gamma_{ij}} (f_r^{(l)}(u(x_{ij})) + (x(\xi) - x_{ij})^\top \nabla_x (f_r^{(l)}(u))(x_{ij})) d\xi + \frac{1}{2} Q_{ij}^{\Gamma_{ij}}(H_{x,f_r^{(l)}(u)}(x_{ij})) + \mathcal{O}(h^3), \\ &= f_r^{(l)}(u(x_{ij})) + (\bar{x}_{ij} - x_{ij})^\top \nabla_x (f_r^{(l)}(u))(x_{ij}) + \frac{1}{2} Q_{ij}^{\Gamma_{ij}}(H_{x,f_r^{(l)}(u)}(x_{ij})) + \mathcal{O}(h^3). \end{aligned}$$

Tout d'abord, en considérant les approximations consistantes de  $H_{x,f_r^{(l)}(u)}(x_{ij})$  données dans (5.68) et comme  $Q_{ij}^{\Gamma_{ij}}$  est d'ordre  $\mathcal{O}(h^2)$ , l'estimation ci-dessus se récrit

$$\frac{1}{|\Gamma_{ij}|} \int_{\Gamma_{ij}} f_r^{(l)}(u(x(\xi))) d\xi = f_r^{(l)}(u(x_{ij})) + (\bar{x}_{ij} - x_{ij})^\top \nabla_x (f_r^{(l)}(u))(x_{ij}) + \frac{1}{2} Q_{ij}^{\Gamma_{ij}}(\overline{H_{x,f_r^{(l)}(u)}_{ij}}) + \mathcal{O}(h^3). \quad (5.73)$$

Ensuite, en utilisant une nouvelle fois les approximations consistantes  $\overline{H_{x,f_r^{(l)}(u)}_{ij}}$  énoncées dans (5.68) ainsi que la séquence  $(u_i)$  définie par (5.30), deux autres développements de Taylor donnent

$$f_r^{(l)}(u_i) = f_r^{(l)}(u(x_{ij})) + (x_i - x_{ij})^\top \nabla_x (f_r^{(l)}(u))(x_{ij}) + \frac{1}{2} Q_{ij}^{C_i}(\overline{H_{x,f_r^{(l)}(u)}_{ij}}) + \mathcal{O}(h^3), \quad (5.74a)$$

$$f_r^{(l)}(u_j) = f_r^{(l)}(u(x_{ij})) + (x_j - x_{ij})^\top \nabla_x (f_r^{(l)}(u))(x_{ij}) + \frac{1}{2} Q_{ij}^{C_j}(\overline{H_{x,f_r^{(l)}(u)}_{ij}}) + \mathcal{O}(h^3). \quad (5.74b)$$

En multipliant l'égalité (5.74a) (*resp.* (5.74b)) par  $d_{j,ij}/d_{i,j}$  (*resp.* par  $d_{i,ij}/d_{i,j}$ ) puis en sommant les résultats, le terme multiplicatif de  $\nabla_x (f_r^{(l)}(u))(x_{ij})$  s'annule et il suit

$$f_r^{(l)}(u(x_{ij})) = \frac{d_{j,ij} f_r^{(l)}(u_i)_i + d_{i,ij} f_r^{(l)}(u_j)_j}{d_{i,j}} - \frac{d_{i,ij}}{2d_{i,j}} Q_{ij}^{C_j}(\overline{H_{x,f_r^{(l)}(u)}_{ij}}) - \frac{d_{j,ij}}{2d_{i,j}} Q_{ij}^{C_i}(\overline{H_{x,f_r^{(l)}(u)}_{ij}}) + \mathcal{O}(h^3). \quad (5.75)$$

Enfin, en utilisant une démarche analogue qui s'appuie sur les approximations du second ordre de  $\nabla_x(f_r^{(l)}(u))(x_{ij})$  définies par (5.60) et comme  $d(\bar{x}_{ij}, x_{ij}) = \mathcal{O}(h)$ , on montre

$$(\bar{x}_{ij} - x_{ij})^\top \nabla_x(f_r^{(l)}(u))(x_{ij}) = (\bar{x}_{ij} - x_{ij})^\top \frac{\overline{d_{j,ij} \nabla_x f_r^{(l)}(u)_i}^{O2} + \overline{d_{i,ij} \nabla_x f_r^{(l)}(u)_j}^{O2}}{d_{i,j}} + \mathcal{O}(h^3).$$

En insérant l'expression ci-dessus ainsi que le résultat intermédiaire (5.75) dans l'estimation (5.73), il suit

$$\begin{aligned} \frac{1}{|\Gamma_{ij}|} \int_{\Gamma_{ij}} f_r^{(l)}(u(x(\xi))) \, d\xi &= \frac{d_{j,ij} f_r^{(l)}(u_i) + d_{i,ij} f_r^{(l)}(u_j)}{d_{i,j}} \\ &+ (\bar{x}_{ij} - x_{ij})^\top \frac{\overline{d_{j,ij} \nabla_x f_r^{(l)}(u)_i}^{O2} + \overline{d_{i,ij} \nabla_x f_r^{(l)}(u)_j}^{O2}}{d_{i,j}} \\ &- \frac{d_{i,ij}}{2d_{i,j}} Q_{ij}^{C_j}(\overline{H_{x,f_r^{(l)}(u)_{ij}}}) - \frac{d_{j,ij}}{2d_{i,j}} Q_{ij}^{C_i}(\overline{H_{x,f_r^{(l)}(u)_{ij}}}) \\ &+ \frac{1}{2} Q_{ij}^{\Gamma_{ij}}(\overline{H_{x,f_r^{(l)}(u)_{ij}}}) + \mathcal{O}(h^3). \end{aligned}$$

La multiplication de l'égalité ci-dessus par  $n_{ij,r}$  puis la sommation sur l'indice  $r$ , permet de déduire l'assertion *i*).

Pour l'assertion *ii*), une démarche analogue qui néglige les termes d'ordre  $\mathcal{O}(h^3)$  au profit des termes d'ordre  $\mathcal{O}(h^2)$  permet de déduire le résultat.

En ce qui concerne l'assertion *iii*), et comme  $n_{ij} = -n_{ji}$ , il est suffisant de montrer  $\mathfrak{F}_{ji,r}^{Ok} = \mathfrak{F}_{ij,r}^{Ok}$  pour tout  $r$  dans  $\{1, 2\}$ . D'après les définitions des quantités  $x_{ij}$ ,  $\bar{x}_{ij}$  et des distances  $d$  données en Section 5.2.1, il vient immédiatement

$$x_{ij} = x_{ji}, \quad \bar{x}_{ij} = \bar{x}_{ji}, \quad d_{i,j} = d_{j,i}, \quad d_{\alpha,ij} = d_{\alpha,ji}, \quad \forall \alpha \in \{i, j\}.$$

De plus, et d'après la définition donnée par (5.68),  $\overline{H_{x,f_r(u)_{ij}}} = \overline{H_{x,f_r(u)_{ji}}}$ . Par conséquent, en utilisant l'expression de  $\mathfrak{F}_{ij,r}^{O3}$  énoncée dans (5.71), il suit pour tout  $r$  dans  $\{1, 2\}$

$$\begin{aligned} &(\mathfrak{F}_{ji}^{O3})^{(l)} \\ &= \frac{d_{i,ji} f_r^{(l)}(u_j) + d_{j,ji} f_r^{(l)}(u_i)}{d_{j,i}} + (\bar{x}_{ji} - x_{ji})^\top \frac{\overline{d_{i,ji} \nabla_x f_r^{(l)}(u)_j}^{O2} + \overline{d_{j,ji} \nabla_x f_r^{(l)}(u)_i}^{O2}}{d_{j,i}} \\ &- \frac{d_{j,ji}}{2d_{j,i}} Q_{ji}^{C_i}(\overline{H_{x,f_r^{(l)}(u)_{ji}}}) - \frac{d_{i,ji}}{2d_{j,i}} Q_{ji}^{C_j}(\overline{H_{x,f_r^{(l)}(u)_{ji}}}) + \frac{1}{2} Q_{ji}^{\Gamma_{ji}}(\overline{H_{x,f_r^{(l)}(u)_{ji}}}), \\ &= \frac{d_{i,ij} f_r^{(l)}(u_j) + d_{j,ij} f_r^{(l)}(u_i)}{d_{i,j}} + (\bar{x}_{ij} - x_{ij})^\top \frac{\overline{d_{i,ij} \nabla_x f_r^{(l)}(u)_j}^{O2} + \overline{d_{j,ij} \nabla_x f_r^{(l)}(u)_i}^{O2}}{d_{i,j}} \\ &- \frac{d_{j,ij}}{2d_{i,j}} Q_{ij}^{C_i}(\overline{H_{x,f_r^{(l)}(u)_{ij}}}) - \frac{d_{i,ij}}{2d_{i,j}} Q_{ij}^{C_j}(\overline{H_{x,f_r^{(l)}(u)_{ij}}}) + \frac{1}{2} Q_{ij}^{\Gamma_{ij}}(\overline{H_{x,f_r^{(l)}(u)_{ij}}}), \\ &= (\mathfrak{F}_{ij,r}^{O3})^{(l)}. \end{aligned}$$

Comme un calcul similaire peut être réalisé avec l'expression de  $\mathfrak{F}_{ij,r}^{O2}$  définie par (5.70) ceci conclut la preuve.  $\square$

Les deux premières assertions du lemme ci-dessus donnent dans (5.70)-(5.71) des approximations respectivement consistantes au ordre deux et trois des intégrales des fonctions de flux. D'après l'assertion *iii*), ces approximations sont de plus conservatives. Ainsi, les expressions de  $\mathfrak{F}_{ij}^{Ok}$  proposées dans (5.70) et dans (5.71) satisfont les critères (5.67a)-(5.67b).

Pour terminer la définition des flux numériques de la forme (5.65), il est maintenant nécessaire de proposer une expression des quantités  $A_{ij}^{Ok}$ .

### 5.3.3 Définitions des corrections dissipatives

Pour une interface orientée  $\Gamma_{ij}$  donnée, et en considérant un flux numérique de la forme (5.65), cette section concerne la définition des quantités  $A_{ij}^{O2}$  et  $A_{ij}^{O3}$  toutes deux dans  $\mathbb{R}^m$ . Pour  $k$  dans  $\{2, 3\}$  et d'après la section 5.3.1 qui définit les notations générales, ces quantités doivent être conservatives ce qui s'écrit  $A_{ij}^{Ok} = -A_{ji}^{Ok}$ . Elles doivent également préserver l'ordre du schéma ce qui se traduit par l'estimation

$$A_{ij}^{Ok} d_{i,j} = \frac{\lambda}{2} (w_j^n - w_i^n) + \mathcal{O}(h^k). \quad (5.76)$$

L'estimation ci-dessus permet de déduire qu'une définition de  $A_{ij}^{Ok}$  s'appuie nécessairement sur des gradients discrets et sur des matrices hessiennes discrètes. D'après la section 5.2.2. et pour  $u : \mathbb{R}^2 \rightarrow \mathbb{R}^m$  une fonction régulière donnée, ces objets se lisent respectivement

$$\overline{\nabla_x u_i}^{Ok} = \begin{pmatrix} \overline{\partial_{x_1} u_i}^{Ok} & \overline{\partial_{x_2} u_i}^{Ok} \end{pmatrix}^T \in \mathcal{M}_{2m,1}(\mathbb{R}), \quad (5.77a)$$

$$\overline{H_{x,u_i}} = \begin{pmatrix} \overline{\partial_{x_1 x_1} u_i} & \overline{\partial_{x_1 x_2} u_i} \\ \overline{\partial_{x_2 x_1} u_i} & \overline{\partial_{x_2 x_2} u_i} \end{pmatrix} \in \mathcal{M}_{2m,2}(\mathbb{R}). \quad (5.77b)$$

Le lemme ci-dessous donne une définition possible de  $A_{ij}^{O2}$  qui utilise les gradients discrets du premier ordre établis dans le Lemme 5.2.2.1.

**Lemme 5.3.3.1** (Définition et propriétés de  $A_{ij}^{O2}$ ). *Soit  $\mathcal{T}$  un maillage admissible qui vérifie la Définition 5.2.1. Pour une fonction régulière  $u : \mathbb{R}^2 \rightarrow \mathbb{R}^m$  donnée, considérons les deux séquences  $(u_i)_{i \in \mathcal{T}}$  et  $(\overline{\nabla_x u_i}^{O1})_{i \in \mathcal{T}}$  respectivement dans  $\mathbb{R}^m$ ,  $\mathcal{M}_{2m,1}(\mathbb{R})$  et respectivement définies par (5.30), (5.31a).*

*Pour une interface orientée  $\Gamma_{ij}$  donnée, considérons  $\Theta_{ij}^{O2}$  une matrice de  $\mathcal{M}_{m,m}(\mathbb{R})$  bornée lorsque  $h \rightarrow 0$  ainsi que des gradients discrets  ${}^* \overline{\nabla_x u_{i,ij}}^{O1}$ ,  ${}^* \overline{\nabla_x u_{j,ji}}^{O1}$  dans  $(\mathcal{M}_{2m,1}(\mathbb{R}))^2$ , respectivement notés*

$${}^* \overline{\nabla_x u_{i,ij}}^{O1} = \left( {}^* \overline{\partial_{x_1} u_{i,ij}}^{O1} \quad {}^* \overline{\partial_{x_2} u_{i,ij}}^{O1} \right)^T \quad \text{et} \quad {}^* \overline{\nabla_x u_{j,ji}}^{O1} = \left( {}^* \overline{\partial_{x_1} u_{j,ji}}^{O1} \quad {}^* \overline{\partial_{x_2} u_{j,ji}}^{O1} \right)^T,$$

et chacun définis par

$${}^* \overline{\partial_{x_\mu} u_{i,ij}}^{O1} = \Theta_{ij}^{O2} \overline{\partial_{x_\mu} u_i}^{O1} + (I_m - \Theta_{ij}^{O2}) \overline{\partial_{x_\mu} u_j}^{O1}, \quad (5.78a)$$

$${}^* \overline{\partial_{x_\mu} u_{j,ji}}^{O1} = \Theta_{ji}^{O2} \overline{\partial_{x_\mu} u_j}^{O1} + (I_m - \Theta_{ji}^{O2}) \overline{\partial_{x_\mu} u_i}^{O1}, \quad \forall \mu \in \{1, 2\}. \quad (5.78b)$$

Si la quantité  $A_{ij}^{O2}$  dans  $\mathbb{R}^m$  s'écrit

$$d_{i,j} A_{ij}^{O2} = \frac{\lambda}{2} \sum_{\mu=1}^2 (x_j - x_i)_\mu \frac{d_{j,ij} {}^* \overline{\partial_{x_\mu} u_{i,ij}}^{O1} + d_{i,ij} {}^* \overline{\partial_{x_\mu} u_{j,ji}}^{O1}}{d_{i,j}}, \quad (5.79)$$

alors

i)  $A_{ij}^{O2}$  est conservatif, i.e  $A_{ij}^{O2} = -A_{ji}^{O2}$ .

ii) l'estimation suivante est vérifiée :  $-\frac{\lambda}{2} (u_j - u_i) + A_{ij}^{O2} d_{i,j} = \mathcal{O}(h^2)$ .

*Démonstration.* La preuve de l'assertion i) s'obtient immédiatement en permutant les indices  $i$  et  $j$  dans la définition de  $A_{ij}^{O2}$  donnée par (5.79).

Concernant l'assertion ii), d'après le Lemme 5.2.2.1, les gradients discrets  $(\overline{\nabla_x u_i}^{O1})_{i \in \mathcal{T}}$  donnés par (5.31a) sont consistants au premier ordre. Comme la matrice  $\Theta_{ij}^{O2}$  est bornée, il suit

$$\begin{aligned} {}^* \overline{\partial_{x_\mu} u_{i,ij}}^{O1} &= \Theta_{ij}^{O2} (\partial_{x_\mu} u(x_{ij}) + \mathcal{O}(h)) + (I_m - \Theta_{ij}^{O2}) (\partial_{x_\mu} u(x_{ij}) + \mathcal{O}(h)), \\ &= \partial_{x_\mu} u(x_{ij}) + \mathcal{O}(h), \quad \forall \mu \in \{1, 2\}. \end{aligned} \quad (5.80)$$

Les mêmes arguments permettent également d'écrire

$$\overline{\partial_{x_\mu} u_{j,ji}}^{O1} = \partial_{x_\mu} u(x_{ij}) + \mathcal{O}(h), \quad \forall \mu \in \{1, 2\}.$$

Par conséquent, et puisque  $d(x_i, x_j) = \mathcal{O}(h)$ , on en déduit que  $A_{ij}^{O2}$  donné par (5.79) vérifie

$$d_{i,j} A_{ij}^{O2} = \frac{\lambda}{2} \sum_{\mu=1}^2 (x_j - x_i)_\mu \partial_{x_\mu} u(x_{ij}) + \mathcal{O}(h^2).$$

En utilisant l'estimation ci-dessus et un développement de Taylor de  $u_i$  et de  $u_j$  autour de  $x_{ij}$ , il est alors possible d'écrire

$$\begin{aligned} -\frac{\lambda}{2} (u_j - u_i) + A_{ij}^{O2} d_{i,j} &= -\frac{\lambda}{2} \sum_{\mu=1}^2 (x_j - x_i)_\mu \partial_{x_\mu} u(x_{ij}) + A_{ij}^{O2} d_{i,j} + \mathcal{O}(h^2), \\ &= \frac{\lambda}{2} \sum_{\mu=1}^2 (x_j - x_i)_\mu (-\partial_{x_\mu} u(x_{ij}) + \partial_{x_\mu} u(x_{ij})) + \mathcal{O}(h^2), \\ &= \mathcal{O}(h^2). \end{aligned}$$

Comme cette dernière estimation correspond au résultat attendu, ceci conclut la preuve.  $\square$

La quantité  $A_{ij}^{O2}$  proposée dans (5.79) est l'analogue de  $(\alpha_i^n + \alpha_{i+1}^n)/2$  défini dans (4.14) pour le cas mono-dimensionnel. D'après les définitions (5.78)-(5.79),  $A_{ij}^{O2}$  s'appuie sur des matrices  $(\Theta_{ij}^{O2}, \Theta_{ji}^{O2})$  dans  $(\mathcal{M}_{m,m}(\mathbb{R}))^2$  définies relativement à une interface. Ce choix est retenu car il garanti la propriété de conservation énoncée dans le Lemme 5.3.3.1-*i*).

Ainsi, pour une interface orientée  $\Gamma_{ij}$  donnée, il est seulement nécessaire de définir  $\Theta_{ij}^{O2}$  puisque la matrice  $\Theta_{ji}^{O2}$  se déduit par permutation des indices  $i$  et  $j$ . Ces deux matrices sont ensuite utilisées dans (5.78) pour définir les gradients discrets  $\overline{\nabla_x u_{i,ij}}^{O1}$  et  $\overline{\nabla_x u_{j,ji}}^{O1}$  qui sont tous deux des approximations consistantes au premier ordre de  $\nabla_x u(x_{ij})$ . D'après la preuve du Lemme 5.3.3.1-*ii*), ces gradients assurent l'estimation (5.76) à l'ordre  $\mathcal{O}(h^2)$ .

Le lemme suivant propose une définition de  $A_{ij}^{O3}$  qui étend ces considérations à l'ordre trois.

**Lemme 5.3.3.2** (Définition et propriétés de  $A_{ij}^{O3}$ ). *Soit  $\mathcal{T}$  un maillage admissible qui vérifie la Définition 5.2.1. Pour une fonction régulière  $u : \mathbb{R}^2 \rightarrow \mathbb{R}^m$  donnée, considérons les séquences  $(u_i)_{i \in \mathcal{T}}$ ,  $(\overline{\nabla_x u_i})_{i \in \mathcal{T}}$  et  $(\overline{H_{x,u_i}})_{i \in \mathcal{T}}$  respectivement dans  $\mathbb{R}^m$ ,  $\mathcal{M}_{2m,1}(\mathbb{R})$  et dans  $\mathcal{M}_{2m,2}(\mathbb{R})$  et respectivement définies par (5.30), (5.31b) et par (5.44).*

*Pour une interface orientée  $\Gamma_{ij}$  fixée, et en considérant les distances  $d$  données par (5.17), notons dans  $\mathbb{R}^m$*

$$\overline{\partial_{x_\mu x_\nu} u_{ij}} = \frac{d_{j,ij} \overline{\partial_{x_\mu x_\nu} u_i} + d_{i,ij} \overline{\partial_{x_\mu x_\nu} u_j}}{d_{i,j}}, \quad \forall (\mu, \nu) \in \{1, 2\}^2. \quad (5.81)$$

*Considérons ensuite  $\Theta_{ij}^{O3}$  une matrice de  $\mathcal{M}_{m,m}(\mathbb{R})$  bornée lorsque  $h \rightarrow 0$  ainsi que des gradients discrets  $\overline{\nabla_x u_{i,ij}}^{O2}$ ,  $\overline{\nabla_x u_{j,ji}}^{O2}$  dans  $(\mathcal{M}_{2m,1}(\mathbb{R}))^2$ , respectivement notés*

$$\overline{\nabla_x u_{i,ij}}^{O2} = \begin{pmatrix} \overline{\partial_{x_1} u_{i,ij}}^{O2} & \overline{\partial_{x_2} u_{i,ij}}^{O2} \end{pmatrix}^T \quad \text{et} \quad \overline{\nabla_x u_{j,ji}}^{O2} = \begin{pmatrix} \overline{\partial_{x_1} u_{j,ji}}^{O2} & \overline{\partial_{x_2} u_{j,ji}}^{O2} \end{pmatrix}^T,$$

et chacun définis par

$$\begin{aligned} {}^*\overline{\partial_{x_\mu} u_{i,ij}}^{O2} &= \Theta_{ij}^{O3} \left( \overline{\partial_{x_\mu} u_i}^{O2} + \sum_{\nu=1}^2 \overline{\partial_{x_\mu x_\nu} u_{i,j}} (x_{ij} - x_i)_\nu \right) \\ &+ (I_m - \Theta_{ij}^{O3}) \left( \overline{\partial_{x_\mu} u_j}^{O2} + \sum_{\nu=1}^2 \overline{\partial_{x_\mu x_\nu} u_{i,j}} (x_{ij} - x_j)_\nu \right), \end{aligned} \quad (5.82a)$$

$$\begin{aligned} {}^*\overline{\partial_{x_\mu} u_{j,ji}}^{O2} &= \Theta_{ji}^{O3} \left( \overline{\partial_{x_\mu} u_j}^{O2} + \sum_{\nu=1}^2 \overline{\partial_{x_\mu x_\nu} u_{i,j}} (x_{ij} - x_j)_\nu \right) \\ &+ (I_m - \Theta_{ji}^{O3}) \left( \overline{\partial_{x_\mu} u_i}^{O2} + \sum_{\nu=1}^2 \overline{\partial_{x_\mu x_\nu} u_{i,j}} (x_{ij} - x_i)_\nu \right), \quad \forall \mu \in \{1, 2\}. \end{aligned} \quad (5.82b)$$

Considérons de plus des matrices hessiennes discrètes  ${}^*\overline{H_{x,u_{i,ij}}}$ ,  ${}^*\overline{H_{x,u_{j,ji}}}$  dans  $(\mathcal{M}_{2m,2}(\mathbb{R}))^2$ , respectivement notées

$${}^*\overline{H_{x,u_{i,ij}}} = \begin{pmatrix} {}^*\overline{\partial_{x_1 x_1} u_{i,ij}} & {}^*\overline{\partial_{x_1 x_2} u_{i,ij}} \\ {}^*\overline{\partial_{x_2 x_1} u_{i,ij}} & {}^*\overline{\partial_{x_2 x_2} u_{i,ij}} \end{pmatrix} \quad \text{et} \quad {}^*\overline{H_{x,u_{j,ji}}} = \begin{pmatrix} {}^*\overline{\partial_{x_1 x_1} u_{j,ji}} & {}^*\overline{\partial_{x_1 x_2} u_{j,ji}} \\ {}^*\overline{\partial_{x_2 x_1} u_{j,ji}} & {}^*\overline{\partial_{x_2 x_2} u_{j,ji}} \end{pmatrix},$$

et chacune définies par

$${}^*\overline{\partial_{x_\mu x_\nu} u_{i,ij}} = \Theta_{ij}^{O3} \overline{\partial_{x_\mu x_\nu} u_i} + (I_m - \Theta_{ij}^{O3}) \overline{\partial_{x_\mu x_\nu} u_j}, \quad (5.83a)$$

$${}^*\overline{\partial_{x_\mu x_\nu} u_{j,ji}} = \Theta_{ji}^{O3} \overline{\partial_{x_\mu x_\nu} u_j} + (I_m - \Theta_{ji}^{O3}) \overline{\partial_{x_\mu x_\nu} u_i}, \quad \forall (\mu, \nu) \in \{1, 2\}^2. \quad (5.83b)$$

Avec la convention d'écriture des formes linéaires  $(Q_{ij}^\alpha)_{\alpha \in \{C_i, C_j\}}$  donnée dans (5.20a) (voir aussi le Lemme 5.2.3.2-iii)), si la quantité  $A_{ij}^{O3}$  dans  $\mathbb{R}^m$  s'écrit

$$\begin{aligned} d_{i,j} A_{ij}^{O3} &= \frac{\lambda}{2} \sum_{\mu=1}^2 (x_j - x_i)_\mu \frac{d_{j,ij} {}^*\overline{\partial_{x_\mu} u_{i,ij}}^{O2} + d_{i,ij} {}^*\overline{\partial_{x_\mu} u_{j,ji}}^{O2}}{d_{i,j}} \\ &+ \frac{\lambda}{4} Q_{ij}^{C_j} \left( \frac{d_{j,ij} {}^*\overline{H_{x,u_{i,ij}}} + d_{i,ij} {}^*\overline{H_{x,u_{j,ji}}}}{d_{i,j}} \right) - \frac{\lambda}{4} Q_{ij}^{C_i} \left( \frac{d_{j,ij} {}^*\overline{H_{x,u_{i,ij}}} + d_{i,ij} {}^*\overline{H_{x,u_{j,ji}}}}{d_{i,j}} \right), \end{aligned} \quad (5.84)$$

alors

i)  $A_{ij}^{O3}$  est conservatif, i.e  $A_{ij}^{O3} = -A_{ji}^{O3}$ .

ii) l'estimation suivante est vérifiée :  $-\frac{\lambda}{2} (u_j - u_i) + A_{ij}^{O3} d_{i,j} = \mathcal{O}(h^3)$ .

*Démonstration.* Pour l'assertion i), le résultat est immédiat en permutant les indices  $i$  et  $j$  puis en remarquant que la quantité  $\overline{\partial_{x_\mu x_\nu} u_{i,j}}$  donnée dans (5.81) vérifie  $\overline{\partial_{x_\mu x_\nu} u_{i,j}} = \overline{\partial_{x_\mu x_\nu} u_{j,i}}$ .

En ce qui concerne l'assertion ii), la séquence  $(u_i)_{i \in \mathcal{T}}$  donnée par (5.30) est pour rappel définie comme suit :

$$u_i = \frac{1}{|C_i|} \int_{C_i} u(x) dx, \quad \forall i \in \mathcal{T}.$$

Par conséquent, et en utilisant la convention d'écriture des formes linéaires  $(Q_{ij}^\alpha)_{\alpha \in \{C_i, C_j\}}$  donnée dans (5.20a), un développement de Taylor usuel autour de  $x_{ij}$  permet d'écrire

$$-\frac{\lambda}{2} (u_j - u_i) = -\frac{\lambda}{2} \sum_{\mu=1}^2 (x_j - x_i)_\mu \partial_{x_\mu} u(x_{ij}) - \frac{\lambda}{4} Q_{ij}^{C_j} (H_{x,u}(x_{ij})) + \frac{\lambda}{4} Q_{ij}^{C_i} (H_{x,u}(x_{ij})) + \mathcal{O}(h^3). \quad (5.85)$$

Ainsi, l'objectif de la preuve est de montrer que  $d_{i,j} A_{ij}^{O3}$  compense la quantité ci dessus à un terme d'erreur près de la forme  $\mathcal{O}(h^3)$ .



La première étape est de montrer la consistance des quantités  ${}^*\overline{\nabla_x u_{i,j}}^{O2}$  et  ${}^*\overline{\nabla_x u_{j,i}}^{O2}$  définies dans (5.78). En utilisant la consistance des gradients discrets  $(\overline{\nabla_x u_i}^{O2})_{i \in \mathcal{T}}$  et des hessiennes discrètes  $(\overline{H_{x,u_i}})_{i \in \mathcal{T}}$  respectivement donnés par (5.31b) et par (5.44), il est possible d'écrire

$$\begin{aligned} & {}^*\overline{\partial_{x_\mu} u_{i,j}}^{O2} \\ &= \Theta_{ij}^{O3} \left( \overline{\partial_{x_\mu} u_i}^{O2} + \sum_{\nu=1}^2 \overline{\partial_{x_\mu x_\nu} u_{ij}}(x_{ij} - x_i)_\nu \right) + (I_m - \Theta_{ij}^{O3}) \left( \overline{\partial_{x_\mu} u_j}^{O2} + \sum_{\nu=1}^2 \overline{\partial_{x_\mu x_\nu} u_{ij}}(x_{ij} - x_j)_\nu \right), \\ &= \Theta_{ij}^{O3} \left( (\partial_{x_\mu} u(x_i) + \mathcal{O}(h^2)) + \sum_{\nu=1}^2 (\partial_{x_\mu x_\nu} u(x_{ij}) + \mathcal{O}(h))(x_{ij} - x_i)_\nu \right) \\ &\quad + (I_m - \Theta_{ij}^{O3}) \left( (\partial_{x_\mu} u(x_j) + \mathcal{O}(h^2)) + \sum_{\nu=1}^2 (\partial_{x_\mu x_\nu} u(x_{ij}) + \mathcal{O}(h))(x_{ij} - x_j)_\nu \right). \end{aligned}$$

Puisque les matrices  $\Theta_{ij}^{O3}$  sont bornées et comme  $d(x_\alpha, x_{ij}) = \mathcal{O}(h)$  pour tout  $\alpha$  dans  $\{i, j\}$ , l'estimation ci-dessus se reformule

$$\begin{aligned} {}^*\overline{\partial_{x_\mu} u_{i,j}}^{O2} &= \Theta_{ij}^{O3} \left( \partial_{x_\mu} u(x_i) + \sum_{\nu=1}^2 \partial_{x_\mu x_\nu} u(x_{ij})(x_{ij} - x_i)_\nu \right) \\ &\quad + (I_m - \Theta_{ij}^{O3}) \left( \partial_{x_\mu} u(x_j) + \sum_{\nu=1}^2 \partial_{x_\mu x_\nu} u(x_{ij})(x_{ij} - x_j)_\nu \right) + \mathcal{O}(h^2), \\ &= \Theta_{ij}^{O3} \partial_{x_\mu} u(x_{ij}) + (I_m - \Theta_{ij}^{O3}) \partial_{x_\mu} u(x_{ij}) + \mathcal{O}(h^2), \\ &= \partial_{x_\mu} u(x_{ij}) + \mathcal{O}(h^2). \end{aligned}$$

Comme un calcul similaire montre que  ${}^*\overline{\partial_{x_\mu} u_{j,i}}^{O2} = \partial_{x_\mu} u(x_{ij}) + \mathcal{O}(h^2)$  et que par ailleurs  $d(x_i, x_j) = \mathcal{O}(h)$ , on en déduit une première estimation partielle de  $d_{i,j} A_{ij}^{O3}$  qui s'écrit

$$\begin{aligned} d_{i,j} A_{ij}^{O3} &= \frac{\lambda}{2} \sum_{\mu=1}^2 (x_j - x_i)_\mu \partial_{x_\mu} u(x_{ij}) \\ &\quad + \frac{\lambda}{4} Q_{ij}^{C_j} \left( \frac{d_{j,ij} \overline{{}^*H_{x,u_{i,j}}} + d_{i,ij} \overline{{}^*H_{x,u_{j,i}}}}{d_{i,j}} \right) - \frac{\lambda}{4} Q_{ij}^{C_i} \left( \frac{d_{j,ij} \overline{{}^*H_{x,u_{i,j}}} + d_{i,ij} \overline{{}^*H_{x,u_{j,i}}}}{d_{i,j}} \right) + \mathcal{O}(h^3). \end{aligned} \quad (5.86)$$

L'étape suivante consiste récrire les termes qui contiennent des formes linéaires  $Q_{ij}^\alpha$  dans l'expression (5.86). D'après les définitions de  $\overline{{}^*H_{x,u_{i,j}}}$  et  $\overline{{}^*H_{x,u_{j,i}}}$  données dans (5.83) et comme les matrices  $\Theta_{ij}^{O3}$  sont bornées, un calcul analogue à (5.80) (voir le Lemme 5.3.3.1) permet d'écrire

$$\begin{aligned} \overline{\partial_{x_\mu x_\nu} u_{i,j}} &= \partial_{x_\mu x_\nu} u(x_{ij}) + \mathcal{O}(h), \\ \overline{\partial_{x_\mu x_\nu} u_{j,i}} &= \partial_{x_\mu x_\nu} u(x_{ij}) + \mathcal{O}(h), \quad \forall (\mu, \nu) \in \{1, 2\}^2. \end{aligned}$$

En utilisant les estimations ci-dessus, et comme  $Q_{ij}^\alpha$  est d'ordre  $\mathcal{O}(h^2)$ , il suit

$$Q_{ij}^\alpha \left( \frac{d_{j,ij} \overline{{}^*H_{x,u_{i,j}}} + d_{i,ij} \overline{{}^*H_{x,u_{j,i}}}}{d_{i,j}} \right) = Q_{ij}^\alpha (H_{x,u}(x_{ij})) + \mathcal{O}(h^3), \quad \forall \alpha \in \{C_i, C_j\}.$$

En introduisant cette dernière estimation dans le résultat intermédiaire (5.86), on en déduit que  $A_{ij}^{O3}$  vérifie

$$d_{i,j} A_{ij}^{O3} = \frac{\lambda}{2} \sum_{\mu=1}^2 (x_j - x_i)_\mu \partial_{x_\mu} u(x_{ij}) + \frac{\lambda}{4} Q_{ij}^{C_j} (H_{x,u}(x_{ij})) - \frac{\lambda}{4} Q_{ij}^{C_i} (H_{x,u}(x_{ij})) + \mathcal{O}(h^3).$$

En sommant cette expression à l'estimation de  $-\lambda(u_j - u_i)/2$  donnée par (5.85), on en déduit le résultat annoncé, ce qui conclut la preuve.  $\square$

Pour  $k$  dans  $\{2, 3\}$ , les Lemmes 5.3.3.1 et 5.3.3.2 définissent des expressions de  $A_{ij}^{Ok}$  conservatives et qui satisfont la propriété de correction à l'ordre  $k$  rappelée dans (5.76). L'association de ces deux lemmes aux résultats du Lemme 5.3.2.1 qui donnent dans (5.71) et dans (5.70), les expressions de  $\mathfrak{F}_{ij}^{Ok}$  termine de définir les flux numériques  $\mathcal{F}_{ij}^{Ok}$  de la forme (5.65). Par construction, les schémas numériques obtenus sont sans limiteur de pente, respectivement du second et du troisième ordre et il convient maintenant de détailler leurs propriétés de stabilité en spécifiant les matrices  $\Theta_{ij}^{Ok}$ .

## 5.4 Stabilité entropique discrète

Cette section est consacrée aux propriétés de stabilité entropique des schémas d'ordre deux et trois définis dans les sections précédentes. L'objectif est de montrer que ces schémas satisfont une inégalité d'entropie discrète globale de la forme

$$\sum_{i \in \mathcal{T}} \frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} |C_i| \leq 0.$$

Pour établir ce résultat, il est nécessaire de montrer que les corrections  $A_{ij}^{O2}$  et  $A_{ij}^{O3}$  respectivement définies dans (5.79) et dans (5.84) sont dissipatives au sens de

$$\int_0^1 \sum_{ij \in \mathcal{E}} \nabla^2 \eta(w_i^n + s(w_j^n - w_i^n))(w_j^n - w_i^n) \cdot A_{ij}^{Ok} d_{i,j} ds < 0, \quad \forall k \in \{2, 3\}. \quad (5.87)$$

L'inégalité ci-dessus est la version bi-dimensionnelle de l'inégalité (4.53) proposée dans le cas d'une dimension d'espace. Par conséquent, une séquence  $(A_{ij}^{Ok})_{ij \in \mathcal{E}}$  qui satisfait l'estimation

$$A_{ij}^{Ok} d_{i,j} = \frac{\lambda}{2} (w_j^n - w_i^n) + \mathcal{O}(h^k),$$

ainsi que l'inégalité (5.87) est une correction dissipative d'ordre  $k$  sur maillages non structurés. Cette estimation et cette inégalité étendent la Définition 4.4.1 au cas de deux dimensions d'espace.

Puisque l'inégalité de dissipation (5.87) est globale, il faut maintenant s'assurer de la convergence de son membre de gauche. L'approche la plus simple et qui assure également la convergence de toutes les sommes dorénavant exposées est la notion de support compact étendu.

**Définition 5.4.1** (Support compact étendu dans un ensemble). *Pour  $d : (\mathbb{R}^2)^2 \rightarrow \mathbb{R}^+$  une fonction de distance quelconque de  $\mathbb{R}^2$ , notons  $B(0, R)$  la boule ouverte relativement à  $d$ , centrée en  $(0, 0)^\top$  et de rayon  $R > 0$ . Soient  $\mathcal{A}$  un ensemble quelconque,  $a$  dans  $\mathcal{A}$  et  $(v_i)_{i \in \mathcal{T}}$  une séquence à valeurs dans  $\mathcal{A}$ . La séquence  $(v_i)_{i \in \mathcal{T}}$  est dite à support compact étendu dans  $\mathcal{A}$  s'il existe un réel  $R > 0$  fini tel que*

$$\forall C_i \in \mathcal{T} \quad \text{si} \quad C_i \notin B(0, R), \quad \text{alors} \quad v_i = a. \quad (5.88)$$

La notion de support compact étendu donnée dans la définition ci-dessus possède seulement un intérêt pratique. Elle signifie formellement qu'au delà d'une certaine distance de  $(0, 0)^\top$  dans  $\mathbb{R}^2$ , la séquence  $(w_i^n)_{i \in \mathbb{Z}}$  est constante. Par conséquent, toute somme qui contient des termes de la forme  $w_j - w_i$  est convergente ce qui est par exemple le cas du membre de gauche de l'inégalité (5.87).

De plus, et puisque les entropies  $\eta$  sont définies à une constante additive près, il est toujours possible de choisir  $a$  dans (5.88) tel que  $(\eta(w_i^n))_{i \in \mathcal{T}}$  est à support compact. Par conséquent, cette notion de support compact étendu assure également la convergence de la somme  $\sum_{i \in \mathcal{T}} \eta(w_i^n) |C_i|$ .

Ainsi, sauf si cela est nécessaire, la sommabilité des quantités proposées ne sera pas mentionnée. Avec ce cadre fonctionnel, la prochaine section montre que l'inégalité (5.87) peut être satisfaite pour le schéma du second ordre.

### 5.4.1 Inégalité de dissipation associée aux corrections du second ordre

Cette section concerne les propriétés de dissipation des corrections  $A_{ij}^{O2}$  définies par (5.79). Par conséquent, l'objectif est de montrer que l'inégalité globale (5.87) avec  $k = 2$  peut être satisfaite. Cette inégalité se lit

$$\int_0^1 \sum_{ij \in \mathcal{E}} |\Gamma_{ij}| \nabla^2 \eta(w_i^n + s(w_j^n - w_i^n))(w_j^n - w_i^n) \cdot A_{ij}^{O2} d_{i,j} ds < 0. \quad (5.89)$$

Par construction, et d'après le Lemme 5.3.3.1, les quantités  $A_{ij}^{O2}$  dépendent linéairement des matrices  $\Theta_{ij}^{O2}$  et des gradients discrets  $\overline{\nabla_x w_i^n}^{O1}$  définis dans (5.31a). La démonstration de l'inégalité (5.89) exploite cette dépendance linéaire et par conséquent, pour  $r$  dans  $\mathbb{N}^*$ , il est utile de définir la fonction

$$\mathbf{D}^{O1} : \begin{cases} (\mathbb{R}^r)^{\text{Card}(\mathcal{T})} \times \{1, \dots, \text{Card}(\mathcal{T})\} \rightarrow \mathcal{M}_{2r,1}(\mathbb{R}) \\ ((v_i)_{i \in \mathcal{T}}, \alpha) \mapsto \sum_{l \in \mathcal{N}(\alpha)} \frac{d_{l,\alpha} v_\alpha + d_{\alpha,\alpha} v_l}{d_{\alpha,l}} b_{\alpha,l}. \end{cases} \quad (5.90)$$

L'expression ci-dessus permet d'écrire les gradients discrets  $\overline{\nabla_x w_i^n}^{O1}$  définis dans (5.31a) de la manière suivante :

$$\overline{\nabla_x w_i^n}^{O1} = \mathbf{D}^{O1}((w_l^n)_{l \in \mathcal{T}}, i), \quad \forall i \in \mathcal{T}.$$

La fonction  $\mathbf{D}^{O1}$  définie par (5.90) dépend de la géométrie du maillage  $\mathcal{T}$  mais comme cela ne porte pas à confusion, cette dépendance est omise. De plus, pour un indice  $\alpha$  fixé, il est clair que l'application  $\cdot \mapsto \mathbf{D}^{O1}(\cdot, \alpha)$  est linéaire.

Avec cette fonction, le lemme qui suit établit l'existence de séquences  $(\Theta_{ij}^{O2})_{ij \in \mathcal{E}}$  telles que l'inégalité stricte globale (5.89) est satisfaite.

**Lemme 5.4.1.1** (Existence de l'inégalité (5.89)). *Soient  $\mathcal{T}$  un maillage admissible qui vérifie la Définition 5.2.1,  $\eta$  une entropie strictement convexe de classe  $C^2(\Omega, \mathbb{R})$  et  $K^n$  un compact de  $\Omega$ . Considérons une séquence  $(w_i^n)_{i \in \mathcal{T}}$  non constante, à valeurs et à support compact étendu dans  $K^n$ . Considérons ensuite les distances  $d$ , les vecteurs  $b$  et la fonction  $\mathbf{D}^{O1}$  respectivement définis dans (5.17), (5.23a) et dans (5.90) puis pour une cellule  $C_i$  fixée notons*

$$\mathbf{r}_{il}^n = \int_0^1 (1-s) \nabla^2 \eta(w_i^n + s(w_l^n - w_i^n))(w_l^n - w_i^n) \cdot (w_l^n - w_i^n) ds, \quad (5.91a)$$

$$= \eta(w_l^n) - \eta(w_i^n) - \nabla \eta(w_i^n) \cdot (w_l^n - w_i^n),$$

$$\mathbf{D}^{O1}((\mathbf{r}_{i,l}^n)_{l \in \mathcal{T}}, \alpha) = \mathbf{D}^{O1}(\mathbf{r}_{i,\alpha}^n) \in \mathbb{R}^2 \quad (5.91b)$$

$$\zeta_{i,j}^{O2} = (x_j - x_i)^\top \mathbf{D}^{O1}(\mathbf{r}_{i,i}^n) - (x_j - x_i)^\top \mathbf{D}^{O1}(\mathbf{r}_{i,j}^n). \quad (5.91c)$$

Supposons que  $(w_i^n)_{i \in \mathcal{T}}$  est telle que  $\sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| |\zeta_{i,j}^{O2} - \zeta_{j,i}^{O2}| \neq 0$ . Pour chaque interface orientée  $\Gamma_{ij}$ , considérons enfin que les matrices  $\Theta_{ij}^{O2}$  qui définissent  $A_{ij}^{O2}$  dans (5.79) s'écrivent

$$\Theta_{ij}^{O2} = \theta_{ij}^{O2} I_m, \quad \text{avec } \theta_{ij}^{O2} \in \mathbb{R}.$$

Si, les réels  $\theta_{ij}^{O2}$  sont tels que

$$\theta_{ij}^{O2} = -\theta_{ji}^{O2} = -\kappa^{O2} \text{sign}(\zeta_{i,j}^{O2} - \zeta_{j,i}^{O2}), \quad (5.92a)$$

$$\kappa^{O2} > \frac{2 \max \left( 0, \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{d_{i,j}} (x_j - x_i)^\top \mathbf{D}^{O1}(d_{j,ij} \mathbf{r}_{i,j}^n + d_{i,ij} \mathbf{r}_{i,i}^n) \right)}{\sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| |\zeta_{i,j}^{O2} - \zeta_{j,i}^{O2}|}, \quad (5.92b)$$

alors l'inégalité de dissipation associée aux corrections  $(A_{ij}^{O2})_{ij \in \mathcal{E}}$  donnée par (5.89) est satisfaite.

*Démonstration.* La preuve de l'implication annoncée repose sur une réécriture du membre de gauche de l'expression (5.89).

D'après le Lemme 5.3.3.1-*i*),  $A_{ij}^{O2}$  vérifie l'égalité de conservation  $A_{ij}^{O2} = -A_{ji}^{O2}$  et par conséquent, une intégration par parties discrète telle que décrite dans le Lemme 5.1.5.1 est possible. En notant  $\cdot$  le produit scalaire usuel de  $\mathbb{R}^m$ , cette intégration donne

$$\int_0^1 \sum_{ij \in \mathcal{E}} \nabla^2 \eta(w_i^n + s(w_j^n - w_i^n))(w_j^n - w_i^n) \cdot A_{ij}^{O2} d_{i,j} ds = - \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \nabla \eta(w_i^n) \cdot A_{ij}^{O2} d_{i,j}. \quad (5.93)$$

Il est maintenant nécessaire de ré-exprimer la quantité  $\nabla \eta(w_i^n) \cdot A_{ij}^{O2} d_{i,j}$ . D'après la définition de  $A_{ij}^{O2}$  donnée par (5.79) et puisque  $\Theta_{ij}^{O2} = \theta_{ij}^{O2} I_m$ , il suit

$$\begin{aligned} & \frac{2}{\lambda} \nabla \eta(w_i^n) \cdot A_{ij}^{O2} d_{i,j} \\ &= \sum_{\mu=1}^2 (x_j - x_i)_\mu \frac{d_{j,ij} \nabla \eta(w_i^n) \cdot \overline{\partial_{x_\mu} w_{i,ij}^n}^{O1} + d_{i,ij} \nabla \eta(w_i^n) \cdot \overline{\partial_{x_\mu} w_{j,ji}^n}^{O1}}{d_{i,j}}, \\ &= \frac{d_{j,ij}}{d_{i,j}} \sum_{\mu=1}^2 (x_j - x_i)_\mu \left( \theta_{ij}^{O2} \nabla \eta(w_i^n) \cdot \overline{\partial_{x_\mu} w_i^n}^{O1} + (1 - \theta_{ij}^{O2}) \nabla \eta(w_i^n) \cdot \overline{\partial_{x_\mu} w_j^n}^{O1} \right) \\ &+ \frac{d_{i,ij}}{d_{i,j}} \sum_{\mu=1}^2 (x_j - x_i)_\mu \left( \theta_{ji}^{O2} \nabla \eta(w_i^n) \cdot \overline{\partial_{x_\mu} w_j^n}^{O1} + (1 - \theta_{ji}^{O2}) \nabla \eta(w_i^n) \cdot \overline{\partial_{x_\mu} w_i^n}^{O1} \right). \end{aligned} \quad (5.94)$$

L'étape suivante consiste à réécrire les termes de la forme  $\nabla \eta(w_i^n) \cdot \overline{\partial_{x_\mu} w_\alpha^n}^{O1}$  pour  $\alpha$  dans  $\{i, j\}$ .

En utilisant la propriété des vecteurs  $b$  qui vérifient  $\sum_{l \in \mathcal{N}(\alpha)} b_{\alpha l} = 0$  (voir la Définition de maillage admissible 5.2.1-*i*)) ainsi que l'égalité  $d_{l,\alpha} + d_{\alpha,l} = d_{\alpha,l}$ , il suit

$$\begin{aligned} 0 &= \sum_{l \in \mathcal{N}(\alpha)} b_{\alpha l}, \\ &= \sum_{l \in \mathcal{N}(\alpha)} \frac{d_{l,\alpha} + d_{\alpha,l}}{d_{\alpha,l}} b_{\alpha l}, \\ &= \sum_{l \in \mathcal{N}(\alpha)} \frac{d_{l,\alpha} \nabla \eta(w_i^n) \cdot w_\alpha^n + d_{\alpha,l} \nabla \eta(w_i^n) \cdot w_l^n}{d_{\alpha,l}} b_{\alpha l}. \end{aligned}$$

En associant cette dernière égalité à un développement de Taylor avec reste intégral dans les expressions des gradients discrets  $(\overline{\nabla_x w_i^n}^{O1})_{i \in \mathcal{T}}$  définis par (5.31a), il vient

$$\begin{aligned} \nabla \eta(w_i^n) \cdot \overline{\partial_{x_\mu} w_\alpha^n}^{O1} &= \sum_{l \in \mathcal{N}(\alpha)} \left( \frac{d_{l,\alpha}}{d_{\alpha,l}} \nabla \eta(w_i^n) \cdot w_\alpha^n + \frac{d_{\alpha,l}}{d_{\alpha,l}} \nabla \eta(w_i^n) \cdot w_l^n \right) b_{\alpha \mu}, \\ &= \sum_{l \in \mathcal{N}(\alpha)} \left( \frac{d_{l,\alpha}}{d_{\alpha,l}} \nabla \eta(w_i^n) \cdot (w_\alpha^n - w_i^n) + \frac{d_{\alpha,l}}{d_{\alpha,l}} \nabla \eta(w_i^n) \cdot (w_l^n - w_i^n) \right) b_{\alpha \mu}, \\ &= \sum_{l \in \mathcal{N}(\alpha)} \left( \frac{d_{l,\alpha}}{d_{\alpha,l}} (\eta(w_\alpha^n) - \eta(w_i^n) - \mathbf{r}_{i\alpha}^n) + \frac{d_{\alpha,l}}{d_{\alpha,l}} (\eta(w_l^n) - \eta(w_i^n) - \mathbf{r}_{il}^n) \right) b_{\alpha \mu}. \end{aligned} \quad (5.95)$$

En notant  $D_\mu^{O1}$  les applications partielles de la fonction  $D^{O1}$  définie par (5.90) et en utilisant la convention  $D^{O1}((\eta(w_l^n))_{l \in \mathcal{T}}, \alpha) = D^{O1}(\eta(w_\alpha^n))$ , cette dernière égalité se réécrit sous la forme

$$\nabla \eta(w_i^n) \cdot \overline{\partial_{x_\mu} w_\alpha^n}^{O1} = D_\mu^{O1}(\eta(w_\alpha^n)) - D_\mu^{O1}(\mathbf{r}_{i,\alpha}^n), \quad \forall (\alpha, \mu) \in \{i, j\} \times \{1, 2\}.$$

En considérant l'égalité ci dessus dans l'expression de  $\nabla\eta(w_i^n) \cdot A_{ij}^{O2}d_{i,j}$  donnée par (5.94), on en déduit que le membre de gauche de l'inégalité de dissipation formulé par (5.93) se lit

$$\begin{aligned}
& \int_0^1 \sum_{ij \in \mathcal{E}} \nabla^2 \eta(w_i^n + s(w_j^n - w_i^n))(w_j^n - w_i^n) \cdot A_{ij}^{O2} d_{i,j} ds \\
&= -\frac{\lambda}{2} \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| (x_j - x_i)^\top \left( \theta_{ij}^{O2} \mathbf{D}^{O1}(\eta(w_i^n)) + (1 - \theta_{ij}^{O2}) \mathbf{D}^{O1}(\eta(w_j^n)) \right) \frac{d_{j,ij}}{d_{i,j}} \\
&\quad - \frac{\lambda}{2} \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| (x_j - x_i)^\top \left( \theta_{ji}^{O2} \mathbf{D}^{O1}(\eta(w_j^n)) + (1 - \theta_{ji}^{O2}) \mathbf{D}^{O1}(\eta(w_i^n)) \right) \frac{d_{i,ij}}{d_{i,j}} \\
&\quad + \frac{\lambda}{2} \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| (x_j - x_i)^\top \left( \theta_{ij}^{O2} \mathbf{D}^{O1}(\mathbf{r}_{i,i}^n) + (1 - \theta_{ij}^{O2}) \mathbf{D}^{O1}(\mathbf{r}_{i,j}^n) \right) \frac{d_{j,ij}}{d_{i,j}} \\
&\quad + \frac{\lambda}{2} \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| (x_j - x_i)^\top \left( \theta_{ji}^{O2} \mathbf{D}^{O1}(\mathbf{r}_{i,j}^n) + (1 - \theta_{ji}^{O2}) \mathbf{D}^{O1}(\mathbf{r}_{i,i}^n) \right) \frac{d_{i,ij}}{d_{i,j}}.
\end{aligned}$$

Les deux premières contributions du membre de droite de l'égalité ci-dessus sont des sommes de termes conservatifs. Par conséquent, ces contributions sont nulles et en utilisant la définition de  $\zeta_{i,j}^{O2}$  donnée par (5.91c), l'équation précédente se réécrit

$$\begin{aligned}
& \frac{2}{\lambda} \int_0^1 \sum_{ij \in \mathcal{E}} \nabla^2 \eta(w_i^n + s(w_j^n - w_i^n))(w_j^n - w_i^n) \cdot A_{ij}^{O2} d_{i,j} ds \\
&= \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{d_{i,j}} (d_{j,ij} \theta_{ij}^{O2} - d_{i,ij} \theta_{ji}^{O2}) \zeta_{i,j}^{O2} + \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{d_{i,j}} (x_j - x_i)^\top \mathbf{D}^{O1}(d_{j,ij} \mathbf{r}_{i,j}^n + d_{i,ij} \mathbf{r}_{i,i}^n).
\end{aligned} \tag{5.96}$$

En introduisant la quantité  $\zeta_{j,i}^{O2}$  dans la première somme du membre de droite de l'égalité ci-dessus, on trouve

$$\begin{aligned}
& \frac{1}{2} \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{d_{i,j}} (d_{j,ij} \theta_{ij}^{O2} - d_{i,ij} \theta_{ji}^{O2}) \zeta_{i,j}^{O2} = \\
& \underbrace{\sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{d_{i,j}} (d_{j,ij} \theta_{ij}^{O2} - d_{i,ij} \theta_{ji}^{O2}) \frac{\zeta_{i,j}^{O2} + \zeta_{j,i}^{O2}}{2}}_{=0 \text{ (somme de termes conservatifs)}} - \frac{1}{2} \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{d_{i,j}} (d_{j,ij} \theta_{ij}^{O2} - d_{i,ij} \theta_{ji}^{O2}) \zeta_{j,i}^{O2}.
\end{aligned}$$

Par conséquent, l'égalité (5.96) se reformule également de la manière suivante :

$$\begin{aligned}
& \frac{2}{\lambda} \int_0^1 \sum_{ij \in \mathcal{E}} \nabla^2 \eta(w_i^n + s(w_j^n - w_i^n))(w_j^n - w_i^n) \cdot A_{ij}^{O2} d_{i,j} ds = \\
& \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{2d_{i,j}} (d_{j,ij} \theta_{ij}^{O2} - d_{i,ij} \theta_{ji}^{O2}) (\zeta_{i,j}^{O2} - \zeta_{j,i}^{O2}) + \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{d_{i,j}} (x_j - x_i)^\top \mathbf{D}^{O1}(d_{j,ij} \mathbf{r}_{i,j}^n + d_{i,ij} \mathbf{r}_{i,i}^n).
\end{aligned}$$

En vertu de la définition de  $\theta_{ij}^{O2}$  donnée par (5.92a), cette dernière expression permet de déduire l'inégalité (5.89) ce qui conclut la preuve.  $\square$

La démonstration ci-dessus repose essentiellement sur la linéarité des gradients discrets  $(\nabla_x w_i^n)^{O1}_{i \in \mathcal{T}}$  interprétée sous la forme de l'application  $\mathbf{D}^{O1}$  définie par (5.90). L'association de cette linéarité à des développements de Taylor avec reste intégral autorise une écriture synthétique de l'inégalité (5.89). Cette écriture permet ensuite de proposer une famille de matrices  $(\Theta_{ij}^{O2})_{ij \in \mathcal{E}}$  dissipative et définie par (5.92). La prochaine section montre comment une stratégie analogue peut être déployée dans le cas du schéma du troisième ordre.

### 5.4.2 Inégalité de dissipation associée aux corrections du troisième ordre

Cette section établit l'existence de l'inégalité globale de dissipation (5.87) pour  $k = 3$  et en considérant les corrections  $(A_{ij}^{O3})_{ij \in \mathcal{E}}$  données par (5.84). Par conséquent, cette inégalité de dissipation se lit

$$\int_0^1 \sum_{ij \in \mathcal{E}} |\Gamma_{ij}| \nabla^2 \eta(w_i^n + s(w_j^n - w_i^n))(w_j^n - w_i^n) \cdot A_{ij}^{O3} d_{i,j} ds < 0. \quad (5.97)$$

D'après la définition (5.84) énoncée dans le Lemme 5.3.3.2, la quantité  $A_{ij}^{O3}$  dépend linéairement des matrices  $\Theta_{ij}^{O3}$ , des gradients discrets  $\overline{\nabla_x w_i^n}^{O2}$  définis dans (5.31b) et des matrices hessiennes discrètes  $\overline{H_{x,w_i^n}}$  données par (5.44). L'objectif est d'exploiter cette linéarité pour montrer que l'inégalité (5.97) peut être satisfaite. Par conséquent, en utilisant la convention d'écriture relative à  $\mathcal{L}(H)$  définie dans (5.20b) et pour  $r$  dans  $\mathbb{N}^*$ , les fonctions suivantes sont maintenant considérées :

$$\mathbf{H} : \left( (\mathbb{R}^r)^{\text{Card}(\mathcal{T})} \times \{1, \dots, \text{Card}(\mathcal{T})\} \rightarrow \mathcal{M}_{2r,2}(\mathbb{R}) \right) \\ ((v_i)_{i \in \mathcal{T}}, \alpha) \mapsto \sum_{(q,l) \in \mathcal{N}(\alpha,\alpha)} \frac{d_{l,\alpha} v_\alpha + d_{\alpha,\alpha} v_l}{d_{\alpha,l}} B_{\alpha l, \alpha q} + \sum_{(q,p) \in \mathcal{N}(\alpha,q)} \frac{d_{p,q} v_q + d_{q,q} v_p}{d_{q,p}} B_{qp, \alpha q}, \quad (5.98)$$

$$\mathbf{D}^{O2} : \left( (\mathbb{R}^r)^{\text{Card}(\mathcal{T})} \times \{1, \dots, \text{Card}(\mathcal{T})\} \rightarrow \mathcal{M}_{2r,1}(\mathbb{R}) \right) \\ ((v_i)_{i \in \mathcal{T}}, \alpha) \mapsto \sum_{l \in \mathcal{N}(\alpha)} \left( \frac{d_{l,\alpha} v_\alpha + d_{\alpha,\alpha} v_l}{d_{\alpha,l}} + \mathcal{L}_{\alpha l}(\mathbf{H}((v_i)_{i \in \mathcal{T}}, \alpha)) \right) b_{\alpha l}. \quad (5.99)$$

Les fonctions  $\mathbf{H}$  et  $\mathbf{D}^{O2}$  définies ci-dessus dépendent également de la géométrie du maillage  $\mathcal{T}$ . Par clarté, et puisqu'il n'y a pas de confusion possible, cette dépendance est volontairement omise.

Les deux fonctions  $\mathbf{H}$  et  $\mathbf{D}^{O2}$  permettent une réécriture équivalente des matrices hessiennes discrètes  $\overline{H_{x,w_i^n}}$  et des gradients discrets  $\overline{\nabla_x w_i^n}^{O2}$  respectivement définis dans (5.44) et dans (5.31b). Cette réécriture est donnée par

$$\overline{H_{x,w_i^n}} = \mathbf{H}((w_l^n)_{l \in \mathcal{T}}, i), \quad \overline{\nabla_x w_i^n}^{O2} = \mathbf{D}^{O2}((w_l^n)_{l \in \mathcal{T}}, i), \quad \forall i \in \mathcal{T}.$$

De plus, pour un indice  $\alpha$  fixé, il est clair que les fonctions  $\cdot \mapsto \mathbf{H}(\cdot, \alpha)$  et  $\cdot \mapsto \mathbf{D}^{O2}(\cdot, \alpha)$  sont linéaires.

Puisque la définition de  $A_{ij}^{O3}$  donnée par (5.84) contient de nombreux éléments, plusieurs lemmes intermédiaires sont maintenant présentés. En s'appuyant sur les fonctions définies dans (5.98)-(5.99), ces lemmes proposent des reformulations des quantités de la forme  $\nabla \eta(w_i^n) \cdot X_j$ . Ces reformulations sont nécessaires à l'expression de  $\nabla \eta(w_i^n) \cdot A_{i,j}^{O3} d_{i,j}$  qui est utilisée pour la preuve de l'inégalité (5.97). Le premier lemme concerne une réécriture de  $\nabla \eta(w_i^n) \cdot Q(\overline{H_{x,w_j^n}})$ .

**Lemme 5.4.2.1** (Reformulation des quantités  $\nabla \eta(w_i^n) \cdot Q(\overline{H_{x,w_j^n}})$ ). *Soient  $\mathcal{T}$  un maillage admissible qui vérifie la Définition 5.2.1 et  $\eta$  une entropie strictement convexe de classe  $C^2(\Omega, \mathbb{R})$ . Considérons une séquence  $(w_j^n)_{j \in \mathcal{T}}$  incluse dans  $\Omega$ , les matrices hessiennes discrètes  $(\overline{H_{x,w_j^n}})_{j \in \mathcal{T}}$  données par (5.44) ainsi que la fonction  $\mathbf{H}$  définie dans (5.98). Pour une cellule  $C_i$  fixée et en considérant la séquence de réel positifs  $(\mathbf{r}_{il}^n)_{l \in \mathcal{T}}$  définie par (5.91a), notons  $(\widetilde{\mathbf{H}}(\mathbf{r}_{i,j}^n))_{j \in \mathcal{T}}$  une séquence dans  $\mathcal{M}_{2,2}(\mathbb{R})$  telle que*

$$\widetilde{\mathbf{H}}(\mathbf{r}_{i,j}^n) = \mathbf{H}((\mathbf{r}_{il}^n)_{l \in \mathcal{T}}, j), \quad \forall j \in \mathcal{T}. \quad (5.100)$$

Pour  $Q : \mathcal{M}_{2,2}(\mathbb{R}) \rightarrow \mathbb{R}$  une forme linéaire quelconque et avec la convention d'écriture de  $Q(H)$  définie dans (5.20a) (voir aussi le Lemme 5.2.3.2-iii), l'égalité suivante est vérifiée :

$$\nabla \eta(w_i^n) \cdot Q(\overline{H_{x,w_j^n}}) = Q(\mathbf{H}((\eta(w_l^n))_{l \in \mathcal{T}}, j)) - Q(\widetilde{\mathbf{H}}(\mathbf{r}_{i,j}^n)). \quad (5.101)$$

*Démonstration.* En utilisant tout d'abord l'expression de la matrice hessienne discrète  $\overline{H_{x,w_j^n}}$  donnée par (5.44), il suit

$$\begin{aligned} \nabla\eta(w_i^n) \cdot Q(\overline{H_{x,w_j^n}}) &= \sum_{(q,l) \in \mathcal{N}(j,j)} \frac{d_{l,jl} \nabla\eta(w_i^n) \cdot w_j^n + d_{j,jl} \nabla\eta(w_i^n) \cdot w_l^n}{d_{j,l}} Q(B_{jl,jq}) \\ &+ \sum_{(q,p) \in \mathcal{N}(j,q)} \frac{d_{p,qp} \nabla\eta(w_i^n) \cdot w_q^n + d_{q,qp} \nabla\eta(w_i^n) \cdot w_p^n}{d_{q,p}} Q(B_{qp,jq}). \end{aligned} \quad (5.102)$$

En s'appuyant ensuite sur la propriété des matrices  $B_{\cdot,jq}$  qui satisfont  $\sum_{p \in \mathcal{N}(q)} B_{jl,jq} = 0$  (voir le Lemme 5.2.3.2-ii)), et comme  $d_{l,jl} + d_{j,jl} = d_{j,l}$ , il est possible d'écrire

$$\begin{aligned} 0 &= \sum_{(q,l) \in \mathcal{N}(j,j)} Q(B_{jl,jq}), \\ &= \sum_{(q,l) \in \mathcal{N}(j,j)} \frac{d_{l,jl} + d_{j,jl}}{d_{j,l}} Q(B_{jl,jq}), \\ &= \sum_{(q,l) \in \mathcal{N}(j,j)} \frac{d_{l,jl} \nabla\eta(w_i^n) \cdot w_i^n + d_{j,jl} \nabla\eta(w_i^n) \cdot w_i^n}{d_{j,l}} Q(B_{jl,jq}). \end{aligned}$$

Un calcul identique *mutatis mutandis* donne également

$$\sum_{(q,p) \in \mathcal{N}(j,q)} \frac{d_{p,qp} \nabla\eta(w_i^n) \cdot w_i^n + d_{q,qp} \nabla\eta(w_i^n) \cdot w_i^n}{d_{q,p}} Q(B_{qp,jq}) = 0.$$

Par conséquent, une reformulation équivalente de l'égalité (5.102) est exprimée par

$$\begin{aligned} \nabla\eta(w_i^n) \cdot Q(\overline{H_{x,w_j^n}}) &= \sum_{(q,l) \in \mathcal{N}(j,j)} \left( \frac{d_{l,jl}}{d_{j,l}} \nabla\eta(w_i^n) \cdot (w_j^n - w_i^n) + \frac{d_{j,jl}}{d_{j,l}} \nabla\eta(w_i^n) \cdot (w_l^n - w_i^n) \right) Q(B_{jl,jq}) \\ &+ \sum_{(q,p) \in \mathcal{N}(j,q)} \left( \frac{d_{p,qp}}{d_{q,p}} \nabla\eta(w_i^n) \cdot (w_q^n - w_i^n) + \frac{d_{q,qp}}{d_{q,p}} \nabla\eta(w_i^n) \cdot (w_p^n - w_i^n) \right) Q(B_{qp,jq}). \end{aligned}$$

Une réécriture des termes de la forme  $\nabla\eta(w_i^n) \cdot (w_l^n - w_i^n)$  à l'aide de l'expression de  $\mathbf{r}_{il}^n$  donnée par (5.91a), puis une nouvelle utilisation de la propriété  $\sum_{p \in \mathcal{N}(q)} B_{qp,jq} = 0$  permettent de déduire que l'égalité précédente est équivalente à

$$\begin{aligned} \nabla\eta(w_i^n) \cdot Q(\overline{H_{x,w_j^n}}) &= \sum_{(q,l) \in \mathcal{N}(j,j)} \frac{d_{l,jl} \eta(w_j^n) + d_{j,jl} \eta(w_l^n)}{d_{j,l}} Q(B_{jl,jq}) + \sum_{(q,p) \in \mathcal{N}(j,q)} \frac{d_{p,qp} \eta(w_q^n) + d_{q,qp} \eta(w_p^n)}{d_{q,p}} Q(B_{qp,jq}) \\ &- \sum_{(q,l) \in \mathcal{N}(j,j)} \frac{d_{l,jl} \mathbf{r}_{ij}^n + d_{j,jl} \mathbf{r}_{il}^n}{d_{j,l}} Q(B_{jl,jq}) - \sum_{(q,p) \in \mathcal{N}(j,q)} \frac{d_{p,qp} \mathbf{r}_{iq}^n + d_{q,qp} \mathbf{r}_{ip}^n}{d_{q,p}} Q(B_{qp,jq}). \end{aligned}$$

En utilisant la linéarité de  $Q$  dans l'expression ci-dessus ainsi que les définitions de  $\mathbf{H}$  et de  $\widetilde{\mathbf{H}}(\mathbf{r}_{i,j}^n)$  respectivement données par (5.98) et par (5.100), on en déduit le résultat attendu ce qui termine la preuve.  $\square$

À l'aide de ce premier résultat intermédiaire, le prochain lemme établit une reformulation de  $\eta(w_i^n) \cdot \overline{\partial_{x_\mu} w_j^n}^{O2}$ .

**Lemme 5.4.2.2** (Reformulation des quantités  $\nabla\eta(w_i^n) \cdot \overline{\partial_{x_\mu} w_j^n}^{O2}$ ). Soient  $\mathcal{T}$  un maillage admissible qui vérifie la Définition 5.2.1 et  $\eta$  une entropie strictement convexe de classe  $C^2(\Omega, \mathbb{R})$ . Considérons une séquence  $(w_j^n)_{j \in \mathcal{T}}$  incluse dans  $\Omega$ , les gradients discrets du second ordre  $(\overline{\nabla_x w_j^n}^{O2})_{j \in \mathcal{T}}$  donnés par (5.31b) ainsi que la fonction  $D^{O2}$  définie dans (5.99). Pour une cellule  $C_i$  fixée et en considérant la séquence de réels positifs  $(\mathbf{r}_{il}^n)_{l \in \mathcal{T}}$  définie par (5.91a), notons  $(\widetilde{D^{O2}(\mathbf{r}_{i,j}^n)})_{j \in \mathcal{T}}$  une séquence dans  $\mathbb{R}^2$  telle que

$$\widetilde{D^{O2}(\mathbf{r}_{i,j}^n)} = D^{O2}((\mathbf{r}_{il}^n)_{l \in \mathcal{T}, j}), \quad \forall j \in \mathcal{T} \quad (5.103)$$

L'égalité suivante est vérifiée :

$$\left( \nabla\eta(w_i^n) \cdot \overline{\partial_{x_1} w_j^n}^{O2} \quad \nabla\eta(w_i^n) \cdot \overline{\partial_{x_2} w_j^n}^{O2} \right)^\top = D^{O2}((\eta(w_l^n))_{l \in \mathcal{T}, j}) - \widetilde{D^{O2}(\mathbf{r}_{i,j}^n)}. \quad (5.104)$$

*Démonstration.* Pour  $\mu$  dans  $\{1, 2\}$ , et d'après la définition des gradients discrets du second ordre  $\overline{\partial_{x_\mu} w_j^n}^{O2}$  donnée par (5.31b), il suit

$$\nabla\eta(w_i^n) \cdot \overline{\partial_{x_\mu} w_j^n}^{O2} = \sum_{q \in \mathcal{N}(j)} \left( \frac{d_{q,jq}}{d_{j,q}} \nabla\eta(w_i^n) \cdot w_j^n + \frac{d_{j,jq}}{d_{j,q}} \nabla\eta(w_i^n) \cdot w_q^n \right) b_{jq\mu} \quad (5.105a)$$

$$+ \sum_{q \in \mathcal{N}(j)} \nabla\eta(w_i^n) \cdot \mathcal{L}_{jq}(\overline{H_{x,w_j^n}}) b_{jq\mu}. \quad (5.105b)$$

L'expression de droite dans (5.105a) s'identifie à la quantité  $\nabla\eta(w_i^n) \cdot \overline{\partial_{x_\mu} w_j^n}^{O1}$  dont une reformulation à été établie dans la preuve du Lemme 5.4.1.1 qui traite de  $\nabla\eta(w_i^n) \cdot A_{ij}^{O2}$ . Cette reformulation est donnée par (5.95) et elle s'écrit

$$\begin{aligned} \sum_{q \in \mathcal{N}(j)} \left( \frac{d_{q,jq}}{d_{j,q}} \nabla\eta(w_i^n) \cdot w_j^n + \frac{d_{j,jq}}{d_{j,q}} \nabla\eta(w_i^n) \cdot w_q^n \right) b_{jq\mu} = \\ \sum_{q \in \mathcal{N}(j)} \frac{d_{q,jq}\eta(w_j^n) + d_{j,jq}\eta(w_q^n)}{d_{j,q}} b_{jq\mu} - \sum_{q \in \mathcal{N}(j)} \frac{d_{q,jq}\mathbf{r}_{ij}^n + d_{j,jq}\mathbf{r}_{iq}^n}{d_{j,q}} b_{jq\mu}. \end{aligned} \quad (5.106)$$

De plus, et puisque que  $\mathcal{L}_{jq}$  est une forme linéaire, le Lemme 5.4.2.1 autorise une réécriture de la quantité  $\nabla\eta(w_i^n) \cdot \mathcal{L}_{jq}(\overline{H_{x,w_j^n}})$ . En considérant  $\widetilde{H(\mathbf{r}_{i,j}^n)}$  défini par (5.100), cette réécriture est donnée par

$$\nabla\eta(w_i^n) \cdot \mathcal{L}_{jq}(\overline{H_{x,w_j^n}}) = \mathcal{L}_{jq}(H((\eta(w_l^n))_{l \in \mathcal{T}, j})) - \mathcal{L}_{jq}(\widetilde{H(\mathbf{r}_{i,j}^n)}).$$

Ainsi, en introduisant l'égalité ci-dessus ainsi que le résultat intermédiaire (5.106) dans l'expression de  $\nabla\eta(w_i^n) \cdot \overline{\partial_{x_\mu} w_j^n}^{O2}$  donnée par (5.105), il vient

$$\begin{aligned} \nabla\eta(w_i^n) \cdot \overline{\partial_{x_\mu} w_j^n}^{O2} = \sum_{q \in \mathcal{N}(j)} \left( \frac{d_{q,jq}\eta(w_j^n) + d_{j,jq}\eta(w_q^n)}{d_{j,q}} + \mathcal{L}_{jq}(H((\eta(w_l^n))_{l \in \mathcal{T}, j})) \right) b_{jq\mu} \\ - \sum_{q \in \mathcal{N}(j)} \left( \frac{d_{q,jq}\mathbf{r}_{ij}^n + d_{j,jq}\mathbf{r}_{iq}^n}{d_{j,q}} + \mathcal{L}_{jq}(\widetilde{H(\mathbf{r}_{i,j}^n)}) \right) b_{jq\mu}. \end{aligned}$$

En considérant les définitions de  $D^{O2}$  et de  $\widetilde{D^{O2}(\mathbf{r}_{i,j}^n)}$  respectivement données par (5.99) et par (5.103) dans l'expression ci-dessus, on en déduit le résultat attendu ce qui conclut la preuve.  $\square$

Les deux lemmes précédents donnent des écritures de  $\nabla\eta(w_i^n) \cdot Q(\overline{H_{x,w_j^n}})$  et de  $\nabla\eta(w_i^n) \cdot \overline{\partial_{x_\mu} w_j^n}^{O2}$ . Puisque la correction  $A_{ij}^{O3}$  définie dans (5.84) dépend linéairement des quantités  $Q(\overline{H_{x,w_j^n}})$  et  $\overline{\partial_{x_\mu} w_j^n}^{O2}$ , une reformulation de  $\nabla\eta(w_i^n) \cdot A_{ij}^{O3}$  est maintenant possible. Pour établir cette reformulation, deux fonctions supplémentaires sont maintenant introduites, chacune relativement au couple cellule-interface



$(C_i, \Gamma_{ij})$  et  $(C_j, \Gamma_{ij})$ . En considérant un entier  $r$  dans  $\mathbb{N}^*$  et en utilisant la convention d'écriture de  $Q(D)$  définie dans (5.20a), ces fonctions sont indicées par  $\alpha$  dans  $\{i, j\}$  et elles sont données par

$$\mathcal{P}_{ij}^{C_\alpha} : \left( \begin{array}{c} \mathcal{M}_{2r,1}(\mathbb{R}) \times (\mathcal{M}_{2r,2}(\mathbb{R}))^3 \rightarrow \mathbb{R}^r \\ \left( \begin{array}{cc} a_1 \\ a_2 \end{array} \right), \left( \begin{array}{cc} b_{11} & b_{12} \\ b_{21} & b_{22} \end{array} \right), \left( \begin{array}{cc} c_{11} & c_{12} \\ c_{21} & c_{22} \end{array} \right), D \mapsto \\ \sum_{\mu=1}^2 (x_j - x_i)_\mu \left( a_\mu + \sum_{\nu=1}^2 \frac{d_{j,ij} b_{\mu\nu} + d_{i,ij} c_{\mu\nu}}{d_{i,j}} (x_{ij} - x_\alpha)_\nu \right) + \frac{1}{2} (Q_{ij}^{C_j} - Q_{ij}^{C_i})(D). \end{array} \right. \quad (5.107)$$

Les fonctions  $(\mathcal{P}_{ij}^{C_\alpha})_{\alpha \in \{i,j\}}$  sont définies relativement à un triplet  $(C_i, \Gamma_{ij}, C_j)$  et par conséquent, leurs expressions dépendent de la géométrie de l'interface orientée  $\Gamma_{ij}$ . Puisqu'il n'y a pas d'ambiguïté possible cette dépendance est omise dans les arguments des fonctions et elle figure seulement dans ses indices.

En s'appuyant sur les fonctions définies dans (5.107), le lemme ci-dessous introduit une reformulation de  $A_{ij}^{O3}$  puis propose une expression de  $\nabla \eta(w_i^n) \cdot A_{ij}^{O3}$ .

**Lemme 5.4.2.3** (Reformulation des quantités  $\eta(w_i^n) \cdot A_{ij}^{O3} d_{i,j}$ ). *Soient  $\mathcal{T}$  un maillage admissible qui vérifie la Définition 5.2.1 et  $\eta$  une entropie strictement convexe de classe  $C^2(\Omega, \mathbb{R})$ . Considérons une séquence  $(w_l^n)_{l \in \mathcal{T}}$  incluse dans  $\Omega$  ainsi que les fonctions  $H$  et  $D^{O2}$  respectivement définies dans (5.98), (5.99) et complétées par les notations*

$$H(\eta(w_i^n)) = H((\eta(w_l^n))_{l \in \mathcal{T}}, i) \in \mathcal{M}_{2,2}(\mathbb{R}), \quad (5.108a)$$

$$D^{O2}(\eta(w_i^n)) = D^{O2}((\eta(w_l^n))_{l \in \mathcal{T}}, i) \in \mathbb{R}^2, \quad \forall i \in \mathcal{T}. \quad (5.108b)$$

Pour un triplet cellule, interface orientée, cellule  $(C_i, \Gamma_{ij}, C_j)$  donné, considérons de plus les quantités discrètes  $(\overline{\nabla_x w_\alpha^n}, \overline{H_{x, w_\alpha^n}}, \widetilde{H(\mathbf{r}_{i,\alpha}^n)}, \widetilde{D^{O2}(\mathbf{r}_{i,\alpha}^n)})_{\alpha \in \{i,j\}}$  ainsi que les fonctions  $(\mathcal{P}_{ij}^{C_\alpha})_{\alpha \in \{i,j\}}$  respectivement définies dans (5.31b), (5.44), (5.100), (5.103), (5.107) et complétées par les notations

$$\mathcal{P}_{\alpha,ij} = \mathcal{P}_{ij}^{C_\alpha}(\overline{\nabla_x w_\alpha^n}, \overline{H_{x, w_\alpha^n}}, \overline{H_{x, w_\alpha^n}}, \overline{H_{x, w_\alpha^n}}) \in \mathbb{R}^m, \quad (5.109a)$$

$$\mathcal{P}_{\alpha,ij}^\eta = \mathcal{P}_{ij}^{C_\alpha}(D^{O2}(\eta(w_\alpha^n)), H(\eta(w_i^n)), H(\eta(w_j^n)), H(\eta(w_\alpha^n))) \in \mathbb{R}, \quad (5.109b)$$

$$\widetilde{\mathcal{P}_{i,\alpha,ij}^{\mathbf{r}}} = \mathcal{P}_{ij}^{C_\alpha}(\widetilde{D^{O2}(\mathbf{r}_{i,\alpha}^n)}, \widetilde{H(\mathbf{r}_{i,i}^n)}, \widetilde{H(\mathbf{r}_{i,j}^n)}, \widetilde{H(\mathbf{r}_{i,\alpha}^n)}) \in \mathbb{R}, \quad \forall \alpha \in \{i, j\}. \quad (5.109c)$$

Considérons enfin que la correction  $A_{ij}^{O3}$  définie dans (5.84) est associée à une matrice  $\Theta_{ij}^{O3}$  de la forme  $\Theta_{ij}^{O3} = \theta_{ij}^{O3} I_m$  avec  $\theta_{ij}^{O3}$  dans  $\mathbb{R}$ . Les assertions suivantes sont vérifiées.

i) Une reformulation équivalente de  $A_{ij}^{O3}$  est donnée par

$$\frac{2}{\lambda} A_{ij}^{O3} d_{i,j} = \frac{d_{j,ij}}{d_{i,j}} (\theta_{ij}^{O3} \mathcal{P}_{i,ij} + (1 - \theta_{ij}^{O3}) \mathcal{P}_{j,ij}) + \frac{d_{i,ij}}{d_{i,j}} (\theta_{ji}^{O3} \mathcal{P}_{j,ij} + (1 - \theta_{ji}^{O3}) \mathcal{P}_{i,ij}). \quad (5.110)$$

ii) L'égalité suivante est satisfaite

$$\begin{aligned} \frac{2}{\lambda} \eta(w_i^n) \cdot A_{ij}^{O3} d_{i,j} &= \frac{d_{j,ij}}{d_{i,j}} (\theta_{ij}^{O3} \mathcal{P}_{i,ij}^\eta + (1 - \theta_{ij}^{O3}) \mathcal{P}_{j,ij}^\eta) + \frac{d_{i,ij}}{d_{i,j}} (\theta_{ji}^{O3} \mathcal{P}_{j,ij}^\eta + (1 - \theta_{ji}^{O3}) \mathcal{P}_{i,ij}^\eta) \\ &\quad - \frac{d_{j,ij}}{d_{i,j}} (\theta_{ij}^{O3} \widetilde{\mathcal{P}_{i,i,ij}^{\mathbf{r}}} + (1 - \theta_{ij}^{O3}) \widetilde{\mathcal{P}_{i,j,ij}^{\mathbf{r}}}) - \frac{d_{i,ij}}{d_{i,j}} (\theta_{ji}^{O3} \widetilde{\mathcal{P}_{j,i,ij}^{\mathbf{r}}} + (1 - \theta_{ji}^{O3}) \widetilde{\mathcal{P}_{j,i,ij}^{\mathbf{r}}}). \end{aligned}$$

*Démonstration.* En associant les notations  $\mathcal{P}_{\alpha,ij}$  données dans (5.109a) à l'expression de  $A_{ij}^{O3}$  telle que définie dans (5.84), l'assertion i) est immédiate.

Concernant l'assertion *ii*), et puisque  $\mathcal{P}_{ij}^{C\alpha}$  définie par (5.107) est linéaire selon toutes ses variables, il existe des réels  $\gamma_{\alpha,ij,\mu}$ ,  $\sigma_{\alpha,ij,\mu\nu}$  et  $\tau_{\alpha,ij,\mu\nu}$  tels que

$$\mathcal{P}_{\alpha,ij} = \sum_{\mu=1}^2 \gamma_{\alpha,ij,\mu} \overline{\partial_{x_\mu} w_\alpha^n}^{O2} + \sum_{(\mu,\nu) \in \{1,2\}^2} \sigma_{\alpha,ij,\mu\nu} \overline{\partial_{x_\mu x_\nu} w_i^n} + \sum_{(\mu,\nu) \in \{1,2\}^2} \tau_{\alpha,ij,\mu\nu} \overline{\partial_{x_\mu x_\nu} w_j^n}. \quad (5.111)$$

Par conséquent, en distribuant le produit scalaire de  $\mathbb{R}^m$  puis en utilisant successivement les résultats des Lemmes 5.4.2.1 et 5.4.2.3 qui donnent respectivement des réécritures de  $\nabla\eta(w_i^n) \cdot Q(\overline{H_{x,w_\alpha^n}})$  et de  $\nabla\eta(w_i^n) \cdot \overline{\partial_{x_\mu} w_\alpha^n}^{O2}$ , il suit

$$\begin{aligned} \nabla\eta(w_i^n) \cdot \mathcal{P}_{\alpha,ij} &= \sum_{\mu=1}^2 \gamma_{\alpha,ij,\mu} \nabla\eta(w_i^n) \cdot \overline{\partial_{x_\mu} w_\alpha^n}^{O2} + \sum_{(\mu,\nu) \in \{1,2\}^2} \sigma_{\alpha,ij,\mu} \nabla\eta(w_i^n) \cdot \overline{\partial_{x_\mu x_\nu} w_i^n} \\ &\quad + \sum_{(\mu,\nu) \in \{1,2\}^2} \tau_{\alpha,ij,\mu} \nabla\eta(w_i^n) \cdot \overline{\partial_{x_\mu x_\nu} w_j^n}, \\ &= \sum_{\mu=1}^2 \gamma_{\alpha,ij,\mu} \left( (D^{O2}((\eta(w_l^n))_{l \in \mathcal{T}}, \alpha))_\mu - (D^{O2}(\widetilde{\mathbf{r}}_{i,\alpha}^n))_\mu \right) \\ &\quad + \sum_{(\mu,\nu) \in \{1,2\}^2} \sigma_{\alpha,ij,\mu} \left( (H((\eta(w_l^n))_{l \in \mathcal{T}}, i))_{\mu\nu} - (H(\widetilde{\mathbf{r}}_{i,i}^n))_{\mu\nu} \right) \\ &\quad + \sum_{(\mu,\nu) \in \{1,2\}^2} \tau_{\alpha,ij,\mu} \left( (H((\eta(w_l^n))_{l \in \mathcal{T}}, j))_{\mu\nu} - (H(\widetilde{\mathbf{r}}_{i,j}^n))_{\mu\nu} \right). \end{aligned}$$

En réorganisant les termes de l'équation ci-dessus et en utilisant les notations  $\mathcal{P}_{\alpha,ij}^\eta$  et  $\widetilde{\mathcal{P}}_{i,\alpha,ij}^\mathbf{r}$  définies dans (5.109b)-(5.109c), il vient

$$\nabla\eta(w_i^n) \cdot \mathcal{P}_{\alpha,ij} = \mathcal{P}_{\alpha,ij}^\eta - \widetilde{\mathcal{P}}_{i,\alpha,ij}^\mathbf{r}, \quad \forall \alpha \in \{i, j\}.$$

L'association de l'égalité ci-dessus à l'expression équivalente de  $A_{ij}^{O3}$  donnée par (5.110) permet d'écrire

$$\begin{aligned} \frac{2}{\lambda} \nabla\eta(w_i^n) \cdot A_{ij}^{O3} d_{i,j} &= \frac{d_{j,ij}}{d_{i,j}} (\theta_{ij}^{O3} \nabla\eta(w_i^n) \cdot \mathcal{P}_{i,ij} + (1 - \theta_{ij}^{O3}) \nabla\eta(w_i^n) \cdot \mathcal{P}_{j,ij}) \\ &\quad + \frac{d_{i,ij}}{d_{i,j}} (\theta_{ji}^{O3} \nabla\eta(w_i^n) \cdot \mathcal{P}_{j,ij} + (1 - \theta_{ji}^{O3}) \nabla\eta(w_i^n) \cdot \mathcal{P}_{i,ij}), \\ &= \frac{d_{j,ij}}{d_{i,j}} (\theta_{ij}^{O3} (\mathcal{P}_{i,ij}^\eta - \widetilde{\mathcal{P}}_{i,i,ij}^\mathbf{r}) + (1 - \theta_{ij}^{O3}) (\mathcal{P}_{j,ij}^\eta - \widetilde{\mathcal{P}}_{i,j,ij}^\mathbf{r})) \\ &\quad + \frac{d_{i,ij}}{d_{i,j}} (\theta_{ji}^{O3} (\mathcal{P}_{j,ij}^\eta - \widetilde{\mathcal{P}}_{i,j,ij}^\mathbf{r}) + (1 - \theta_{ji}^{O3}) (\mathcal{P}_{i,ij}^\eta - \widetilde{\mathcal{P}}_{i,i,ij}^\mathbf{r})). \end{aligned}$$

En réorganisant cette expression, on en déduit le résultat attendu ce qui conclut la démonstration.  $\square$

Puisque le lemme ci dessus propose une reformulation de  $\nabla\eta(w_i^n) \cdot A_{ij}^{O3} d_{i,j}$ , l'inégalité de dissipation associée à  $(A_{ij}^{O3})_{ij \in \mathcal{T}}$  définie par (5.97) peut maintenant être établie.

**Proposition 5.4.2.1** (Existence de l'inégalité (5.97)). *Soient  $\mathcal{T}$  un maillage admissible qui vérifie la Définition 5.2.1,  $\eta$  une entropie strictement convexe de classe  $C^2(\Omega, \mathbb{R})$  et  $K^n$  un compact de  $\Omega$ . Considérons une séquence  $(w_i^n)_{i \in \mathcal{T}}$  non constante, à valeurs et à support compact étendu dans  $K^n$ . Considérons ensuite les quantités réelles  $\widetilde{\mathcal{P}}_{i,i,ij}^\mathbf{r}$  et  $\widetilde{\mathcal{P}}_{i,j,ij}^\mathbf{r}$  définies par (5.109c) et notons*

$$\zeta_{i,j}^{O3} = \widetilde{\mathcal{P}}_{i,i,ij}^\mathbf{r} - \widetilde{\mathcal{P}}_{i,j,ij}^\mathbf{r}. \quad (5.112)$$

Supposons que  $(w_i^n)_{i \in \mathcal{T}}$  est telle que  $\sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| |\zeta_{i,j}^{O3} - \zeta_{j,i}^{O3}| \neq 0$ . Pour chaque interface orientée  $\Gamma_{ij}$ , considérons enfin que les matrices  $\Theta_{ij}^{O3}$  qui définissent  $A_{ij}^{O3}$  dans (5.84) s'écrivent

$$\Theta_{ij}^{O3} = \theta_{ij}^{O3} I_m, \quad \text{avec } \theta_{ij}^{O3} \in \mathbb{R}.$$

Si, les réels  $\theta_{ij}^{O3}$  sont tels que

$$\theta_{ij}^{O3} = -\theta_{ji}^{O3} = -\kappa^{O3} \operatorname{sign}(\zeta_{i,j}^{O3} - \zeta_{j,i}^{O3}), \quad (5.113a)$$

$$\kappa^{O3} > \frac{2 \max\left(0, \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{d_{i,j}} (d_{i,ij} \widetilde{\mathcal{P}}_{i,i,ij}^{\mathbf{r}} + d_{j,ij} \widetilde{\mathcal{P}}_{i,j,ij}^{\mathbf{r}})\right)}{\sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| |\zeta_{i,j}^{O3} - \zeta_{j,i}^{O3}|}, \quad (5.113b)$$

alors l'inégalité de dissipation (5.97) associée aux corrections  $(A_{ij}^{O3})_{ij \in \mathcal{E}}$  est satisfaite. Pour rappel, cette inégalité se lit

$$\int_0^1 \sum_{ij \in \mathcal{E}} |\Gamma_{ij}| \nabla^2 \eta(w_i^n + s(w_j^n - w_i^n))(w_j^n - w_i^n) \cdot A_{ij}^{O3} d_{i,j} ds < 0. \quad (5.114)$$

*Démonstration.* La preuve de l'implication annoncée repose sur une réécriture du membre de gauche de l'expression (5.114). En utilisant successivement une intégration par parties discrète ainsi que le résultat du Lemme 5.4.2.3 qui reformule la quantité  $\nabla \eta(w_i^n) \cdot A_{ij}^{O3} d_{i,j}$ , il suit

$$\begin{aligned} & \frac{2}{\lambda} \int_0^1 \sum_{ij \in \mathcal{E}} \nabla^2 \eta(w_i^n + s(w_j^n - w_i^n))(w_j^n - w_i^n) \cdot A_{ij}^{O3} d_{i,j} ds \\ &= -\frac{2}{\lambda} \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \nabla \eta(w_i^n) \cdot A_{ij}^{O3} d_{i,j}, \\ &= -\sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \left( \frac{d_{j,ij}}{d_{i,j}} (\theta_{ij}^{O3} \mathcal{P}_{i,ij}^\eta + (1 - \theta_{ij}^{O3}) \mathcal{P}_{j,ij}^\eta) + \frac{d_{i,ij}}{d_{i,j}} (\theta_{ji}^{O3} \mathcal{P}_{j,ij}^\eta + (1 - \theta_{ji}^{O3}) \mathcal{P}_{i,ij}^\eta) \right) \\ & \quad + \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \left( \frac{d_{j,ij}}{d_{i,j}} (\theta_{ij}^{O3} \widetilde{\mathcal{P}}_{i,i,ij}^{\mathbf{r}} + (1 - \theta_{ij}^{O3}) \widetilde{\mathcal{P}}_{i,j,ij}^{\mathbf{r}}) + \frac{d_{i,ij}}{d_{i,j}} (\theta_{ji}^{O3} \widetilde{\mathcal{P}}_{i,j,ij}^{\mathbf{r}} + (1 - \theta_{ji}^{O3}) \widetilde{\mathcal{P}}_{i,i,ij}^{\mathbf{r}}) \right). \end{aligned} \quad (5.115)$$

La contribution (5.115) est constituée de termes conservatifs. Par conséquent, cette somme est nulle et l'expression ci-dessus devient

$$\begin{aligned} & \frac{2}{\lambda} \int_0^1 \sum_{ij \in \mathcal{E}} \nabla^2 \eta(w_i^n + s(w_j^n - w_i^n))(w_j^n - w_i^n) \cdot A_{ij}^{O3} d_{i,j} ds = \\ & \quad \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \left( \frac{d_{j,ij}}{d_{i,j}} (\theta_{ij}^{O3} \widetilde{\mathcal{P}}_{i,i,ij}^{\mathbf{r}} + (1 - \theta_{ij}^{O3}) \widetilde{\mathcal{P}}_{i,j,ij}^{\mathbf{r}}) + \frac{d_{i,ij}}{d_{i,j}} (\theta_{ji}^{O3} \widetilde{\mathcal{P}}_{i,j,ij}^{\mathbf{r}} + (1 - \theta_{ji}^{O3}) \widetilde{\mathcal{P}}_{i,i,ij}^{\mathbf{r}}) \right). \end{aligned}$$

En récrivant cette dernière égalité avec la définition de  $\zeta_{i,j}^{O3}$  donnée par (5.112) puis en réorganisant les termes, il vient

$$\begin{aligned} & \frac{2}{\lambda} \int_0^1 \sum_{ij \in \mathcal{E}} \nabla^2 \eta(w_i^n + s(w_j^n - w_i^n))(w_j^n - w_i^n) \cdot A_{ij}^{O3} d_{i,j} ds = \\ & \quad \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{d_{i,j}} (d_{i,ij} \theta_{ij}^{O3} - d_{j,ij} \theta_{ji}^{O3}) \zeta_{i,j}^{O3} + \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{d_{i,j}} (d_{i,ij} \widetilde{\mathcal{P}}_{i,i,ij}^{\mathbf{r}} + d_{j,ij} \widetilde{\mathcal{P}}_{i,j,ij}^{\mathbf{r}}). \end{aligned} \quad (5.116)$$

En introduisant la quantité  $\zeta_{j,i}^{O3}$  dans la première somme du membre de droite de l'égalité ci-dessus, on trouve

$$\frac{1}{2} \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{d_{i,j}} (d_{j,ij} \theta_{ij}^{O3} - d_{i,ij} \theta_{ji}^{O3}) \zeta_{i,j}^{O3} =$$

$$\underbrace{\sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{d_{i,j}} (d_{j,ij} \theta_{ij}^{O3} - d_{i,ij} \theta_{ji}^{O3}) \frac{\zeta_{i,j}^{O3} + \zeta_{j,i}^{O3}}{2} - \frac{1}{2} \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{d_{i,j}} (d_{j,ij} \theta_{ij}^{O3} - d_{i,ij} \theta_{ji}^{O3}) \zeta_{j,i}^{O3}}_{=0 \text{ (somme de termes conservatifs)}}$$

Par conséquent, l'égalité (5.116) se reformule également de la manière suivante :

$$\begin{aligned} \frac{2}{\lambda} \int_0^1 \sum_{ij \in \mathcal{E}} \nabla^2 \eta(w_i^n + s(w_j^n - w_i^n))(w_j^n - w_i^n) \cdot A_{ij}^{O3} d_{i,j} ds = \\ \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{2d_{i,j}} (d_{j,ij} \theta_{ij}^{O3} - d_{i,ij} \theta_{ji}^{O3}) (\zeta_{i,j}^{O3} - \zeta_{j,i}^{O3}) + \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{d_{i,j}} (d_{i,ij} \widetilde{\mathcal{P}}_{i,i,ij}^r + d_{j,ij} \widetilde{\mathcal{P}}_{i,j,ij}^r). \end{aligned}$$

En s'appuyant sur la définition de  $\theta_{ij}^{O3}$  donnée par (5.113a), cette dernière expression permet de déduire l'inégalité (5.97) ce qui conclut la preuve.  $\square$

Le lemme précédent montre l'existence de corrections  $(A_{ij}^{O3})_{ij \in \mathcal{E}}$  telles que l'inégalité de dissipation (5.114) est satisfaite or le Lemme 5.3.3.2 a établi en Section 5.3.3 que ces corrections vérifient également l'estimation

$$A_{ij}^{O3} d_{i,j} = \frac{\lambda}{2} (w_j^n - w_i^n) + \mathcal{O}(h^3).$$

Par conséquent, ces deux résultats montrent l'existence de corrections dissipatives du troisième ordre sur maillages non structurés. En associant ces résultats aux analogues obtenus pour le second ordre en Section 5.4.1, il est possible de proposer un résultat de stabilité entropique discrète globale sur maillage non structurés.

### 5.4.3 Stabilité entropique discrète globale

En notant  $|P_i| > 0$ , le périmètre de la cellule  $C_i$ , les schémas conservatifs du second et du troisième définis dans la section 5.3 s'écrivent

$$w_i^{n+1} = \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{|P_i|} \left( w_i^n - \frac{\Delta t |P_i|}{|C_i|} \mathcal{F}_{ij}^{Ok} \right), \quad (5.117a)$$

$$\mathcal{F}_{ij}^{Ok} = \mathfrak{F}_{ij}^{Ok} - \frac{\lambda}{2} (w_j^n - w_i^n) + A_{ij}^{Ok} d_{i,j}, \quad (5.117b)$$

où les quantités  $(\mathcal{F}_{ij}^{Ok}, A_{ij}^{Ok})$  sont respectivement données par (5.70), (5.79) lorsque  $k = 2$  et par (5.71), (5.84) lorsque  $k = 3$ . Le résultat de stabilité entropique de ces schémas est le suivant.

**Théorème 5.4.1** (Stabilité entropique globale discrète sur maillages non structurés). *Soient  $\mathcal{T}$  un maillage admissible qui vérifie la Définition 5.2.1 et  $\eta$  une entropie strictement convexe de classe  $C^2(\Omega, \mathbb{R})$ . Considérons une séquence  $(w_i^n)_{i \in \mathcal{T}}$  non constante, incluse dans un compact de  $\Omega$  et telle que*

$$\left| \sum_{i \in \mathcal{T}} \eta(w_i^n) |C_i| \right| < +\infty, \quad (5.118a)$$

$$\left| \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| |\nabla \eta(w_i^n) \cdot \mathcal{F}_{ij}^{Ok}| \right| < +\infty, \quad (5.118b)$$

$$\left| \sum_{i \in \mathcal{T}} \frac{|P_i|}{|C_i|} \int_0^1 (1-s) \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| |\nabla^2 \eta(w_i^n - s \frac{\Delta t |P_i|}{|C_i|} \mathcal{F}_{ij}^{Ok}) \mathcal{F}_{ij}^{Ok} \cdot \mathcal{F}_{ij}^{Ok} ds \right| < +\infty, \quad (5.118c)$$

Si les quantités conservatives  $A_{ij}^{Ok}$  données par (5.79), (5.84), vérifient l'inégalité stricte globale

$$\int_0^1 \sum_{ij \in \mathcal{E}} |\Gamma_{ij}| \nabla^2 \eta(w_i^n + s(w_j^n - w_i^n))(w_j^n - w_i^n) \cdot A_{ij}^{Ok} d_{i,j} ds < 0, \quad \forall k \in \{2, 3\}, \quad (5.119)$$

alors, il existe  $\lambda > 0$  suffisamment grand et  $\Delta t > 0$  suffisamment petit tels que la séquence mise à jour  $(w_i^{n+1})_{i \in \mathcal{T}}$  donnée par les schémas d'ordre élevé (5.117) est incluse dans  $\Omega$  et tels que cette séquence vérifie

$$\sum_{i \in \mathcal{T}} \frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} |C_i| \leq 0. \quad (5.120)$$

*Démonstration.* Avec l'hypothèse de compacité énoncée sur la séquence  $(w_i^n)_{i \in \mathcal{T}}$ , des arguments topologiques tels que ceux déployés dans le Lemme 4.1.1 montreraient l'existence de  $\Delta t > 0$  tel que  $(w_i^{n+1})_{i \in \mathcal{T}} \subset \Omega$ . En supposant que cette condition est vérifiée, la convexité de l'entropie  $\eta$  donne

$$\eta(w_i^{n+1}) \leq \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{|P_i|} \eta\left(w_i^n - \frac{\Delta t |P_i|}{|C_i|} \mathcal{F}_{ij}^{Ok}\right).$$

En notant  $\cdot$  le produit scalaire usuel de  $\mathbb{R}^m$ , et en associant la régularité  $C^2$  de l'entropie  $\eta$  à un développement de Taylor, il vient

$$\begin{aligned} \eta(w_i^{n+1}) &\leq \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{|P_i|} \eta\left(w_i^n - \frac{\Delta t |P_i|}{|C_i|} \mathcal{F}_{ij}^{Ok}\right), \\ &\leq \eta(w_i^n) - \frac{\Delta t}{|C_i|} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \nabla \eta(w_i^n) \cdot \mathcal{F}_{ij}^{Ok} \\ &\quad + \left(\frac{\Delta t |P_i|}{|C_i|}\right)^2 \int_0^1 (1-s) \sum_{j \in \mathcal{N}(i)} \frac{|\Gamma_{ij}|}{|P_i|} \nabla^2 \eta\left(w_i^n - s \frac{\Delta t |P_i|}{|C_i|} \mathcal{F}_{ij}^{Ok}\right) \mathcal{F}_{ij}^{Ok} \cdot \mathcal{F}_{ij}^{Ok} ds. \end{aligned}$$

En réorganisant les termes de l'inégalité ci-dessus puis en sommant sur toutes les cellules  $C_i$ , il suit

$$\begin{aligned} \sum_{i \in \mathcal{T}} \frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} |C_i| &\leq - \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \nabla \eta(w_i^n) \cdot \mathcal{F}_{ij}^{Ok} \\ &\quad + \Delta t \sum_{i \in \mathcal{T}} \frac{|P_i|}{|C_i|} \int_0^1 (1-s) \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \nabla^2 \eta\left(w_i^n - s \frac{\Delta t |P_i|}{|C_i|} \mathcal{F}_{ij}^{Ok}\right) \mathcal{F}_{ij}^{Ok} \cdot \mathcal{F}_{ij}^{Ok} ds. \end{aligned} \quad (5.121)$$

En vertu des hypothèses de sommabilité énoncées dans (5.118), toutes les quantités ci-dessus sont finies. Par conséquent, l'objectif à ce stade est de récrire la première somme du membre de droite de l'inégalité (5.121) afin d'en proposer une majoration. Puisque le flux numérique est conservatif, une intégration par parties discrète telle que décrite dans le Lemme 5.1.5.1 permet d'écrire

$$\begin{aligned} & - \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \mathcal{F}_{ij}^{Ok} \cdot \nabla \eta(w_i^n) \\ &= \sum_{ij \in \mathcal{E}} |\Gamma_{ij}| \mathcal{F}_{ij}^{Ok} \cdot \left( \nabla \eta(w_j^n) - \nabla \eta(w_i^n) \right), \\ &= \int_0^1 \sum_{ij \in \mathcal{E}} |\Gamma_{ij}| \nabla^2 \eta(w_i^n + s(w_j^n - w_i^n))(w_j^n - w_i^n) \cdot \mathfrak{F}_{ij}^{Ok} ds \\ &\quad + \int_0^1 \sum_{ij \in \mathcal{E}} |\Gamma_{ij}| \nabla^2 \eta(w_i^n + s(w_j^n - w_i^n))(w_j^n - w_i^n) \cdot \left( -\frac{\lambda}{2} (w_j^n - w_i^n) + A_{ij}^{Ok} d_{i,j} \right) ds. \end{aligned} \quad (5.122)$$

En utilisant l'inégalité (5.119) dans l'expression ci dessus, et en notant  $\delta_{ij}^n = w_j^n - w_i^n$  dans  $\mathbb{R}^m$ , on en déduit

$$\begin{aligned} & - \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \mathcal{F}_{ij}^{Ok} \cdot \nabla \eta(w_i^n) \\ & \leq \int_0^1 \sum_{ij \in \mathcal{E}} |\Gamma_{ij}| \nabla^2 \eta(w_i^n + s \delta_{ij}^n) \delta_{ij}^n \cdot \mathfrak{F}_{ij}^{Ok} ds - \frac{\lambda}{2} \int_0^1 \sum_{ij \in \mathcal{E}} |\Gamma_{ij}| \nabla^2 \eta(w_i^n + s \delta_{ij}^n) \delta_{ij}^n \cdot \delta_{ij}^n ds. \end{aligned}$$

Cette dernière inégalité donne le majorant recherché et son introduction dans le résultat intermédiaire (5.121) permet d'écrire

$$\begin{aligned} & \sum_{i \in \mathcal{T}} \frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} |C_i| \\ & \leq \int_0^1 \sum_{ij \in \mathcal{E}} |\Gamma_{ij}| \nabla^2 \eta(w_i^n + s \delta_{ij}^n) \delta_{ij}^n \cdot \mathfrak{F}_{ij}^{Ok} ds - \frac{\lambda}{2} \int_0^1 \sum_{ij \in \mathcal{E}} |\Gamma_{ij}| \nabla^2 \eta(w_i^n + s \delta_{ij}^n) \delta_{ij}^n \cdot \delta_{ij}^n ds \\ & \quad + \Delta t \sum_{i \in \mathcal{T}} \frac{|P_i|}{|C_i|} \int_0^1 (1-s) \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \nabla^2 \eta \left( w_i^n - s \frac{\Delta t |P_i|}{|C_i|} \mathcal{F}_{ij}^{Ok} \right) \mathcal{F}_{ij}^{Ok} \cdot \mathcal{F}_{ij}^{Ok} ds. \quad (5.123) \end{aligned}$$

Il est maintenant nécessaire de montrer que si  $\lambda > 0$  et  $\Delta t > 0$  sont convenablement choisis, alors le membre de droite de l'inégalité ci-dessus peut être négatif.

Tout d'abord, un raisonnement par l'absurde immédiat associant l'égalité (5.122) à la sommabilité (5.118b) donne

$$\left| \int_0^1 \sum_{ij \in \mathcal{E}} |\Gamma_{ij}| \nabla^2 \eta(w_i^n + s \delta_{ij}^n) \delta_{ij}^n \cdot \mathfrak{F}_{ij}^{Ok} ds \right| < +\infty.$$

Ensuite, comme la séquence  $(w_i^n)_{i \in \mathcal{T}}$  n'est pas constante et puisque que l'entropie  $\eta$  est strictement convexe, il suit

$$0 < \int_0^1 \sum_{ij \in \mathcal{E}} |\Gamma_{ij}| \nabla^2 \eta(w_i^n + s \delta_{ij}^n) \delta_{ij}^n \cdot \delta_{ij}^n ds.$$

Enfin, en arguant une nouvelle fois que la séquence  $(w_i^n)_{i \in \mathcal{T}}$  n'est pas constante, la convexité stricte de  $\eta$  et la sommabilité (5.118c) permettent d'écrire

$$0 < \sum_{i \in \mathcal{T}} \frac{|P_i|}{|C_i|} \int_0^1 (1-s) \sum_{j \in \mathcal{N}(i)} |\Gamma_{ij}| \nabla^2 \eta \left( w_i^n - s \frac{\Delta t |P_i|}{|C_i|} \mathcal{F}_{ij}^{Ok} \right) \mathcal{F}_{ij}^{Ok} \cdot \mathcal{F}_{ij}^{Ok} ds < +\infty.$$

Ainsi, si  $\lambda > 0$  et  $\Delta t > 0$  sont tels que

$$\lambda > \frac{2 \max \left( 0, \int_0^1 \sum_{ij \in \mathcal{E}} |\Gamma_{ij}| \nabla^2 \eta(w_i^n + s \delta_{ij}^n) \delta_{ij}^n \cdot \mathfrak{F}_{ij}^{Ok} ds \right)}{\int_0^1 \sum_{ij \in \mathcal{E}} |\Gamma_{ij}| \nabla^2 \eta(w_i^n + s \delta_{ij}^n) \delta_{ij}^n \cdot \delta_{ij}^n ds}, \quad (5.124a)$$

$$0 < \Delta t \leq \frac{\int_0^1 \sum_{ij \in \mathcal{E}} |\Gamma_{ij}| \nabla^2 \eta(w_i^n + s \delta_{ij}^n) \delta_{ij}^n \cdot \delta_{ij}^n ds}{\int_0^1 (1-s) \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{N}(i)} \frac{|P_i| |\Gamma_{ij}|}{|C_i|} \nabla^2 \eta \left( w_i^n - s \frac{\Delta t |P_i|}{|C_i|} \mathcal{F}_{ij}^{Ok} \right) \mathcal{F}_{ij}^{Ok} \cdot \mathcal{F}_{ij}^{Ok} ds}, \quad (5.124b)$$

alors, d'après l'inégalité (5.123), on en déduit le résultat de stabilité annoncé dans (5.120) ce qui termine la preuve.  $\square$

L'originalité du Théorème (5.4.1) est donnée par l'inégalité de dissipation (5.119). Cette inégalité peut s'interpréter comme une contrainte sur les corrections  $(A_{ij}^{O_k})_{ij \in \mathcal{E}}$  et elle s'obtient en sélectionnant les séquences de matrices  $(\Theta_{ij}^{O_k})_{ij \in \mathcal{E}}$  qui définissent  $(A_{ij}^{O_k})_{ij \in \mathcal{E}}$  dans (5.79), (5.84). Avec la donnée de telles séquences, le résultat de stabilité entropique (5.120) est ensuite établi sous la condition CFL (5.124).

Ainsi, le Théorème 5.4.1 montre l'existence de schémas numériques explicites en temps, du second et du troisième ordre, sans limiteur de pente, définis sur des maillages non structurés et qui vérifient une inégalité d'entropie discrète globale donnée par (5.120).

Toutefois, si l'existence théorique des schémas est démontrée, il est clair qu'une mise en œuvre pratique avec les éléments des démonstrations demeure extrêmement difficile. En effet, les familles de séquence  $(\Theta_{ij}^{O_k})_{ij \in \mathcal{E}}$  proposées dans (5.92) pour le second ordre et dans (5.113) pour le troisième ordre sont toutes globales et leurs évaluations à chaque itération temporelle conduiraient nécessairement à des temps de calcul rédhibitoires. Cette remarque est également valable pour la condition CFL (5.124) qui n'est, de plus, très certainement pas optimale.

Par conséquent, la programmation et la mise en œuvre des schémas énoncés dans le Théorème 5.4.1 demandent des précisions complémentaires qui sont maintenant détaillées puis accompagnées de résultats numériques.

## 5.5 Résultats de simulations numériques

### 5.5.1 Méthodes et paramétrages

Les schémas développés dans ce chapitre ont été programmés en langage C sans librairie particulière puis parallélisés avec quelques instructions `OpenMP`. Les éléments de géométrie du maillage tels que les distances  $d$ , les vecteurs  $b$ , les matrices  $B$ , etc, sont tous calculés puis stockés avant la boucle temporelle. Ainsi, les objets définis à l'aide de ces éléments comme les gradients discrets (5.31) ou les matrices hessiennes discrètes (5.44) s'obtiennent simplement avec des opérations de types produits scalaires. La séquence mise à jour  $(w_i^{n+1})_{i \in \mathcal{T}}$  est calculée à l'aide d'une boucle itérative sur les interfaces non orientées.

Les schémas numériques ainsi programmés sont testés sur trois types de maillages : les grilles cartésiennes uniformes, les maillages définis par des triangles structurés et les maillages non structurés générés par le logiciel `FreeFem++` [120]. Quelques cellules de ces maillages sont représentées par la Figure 5.3.

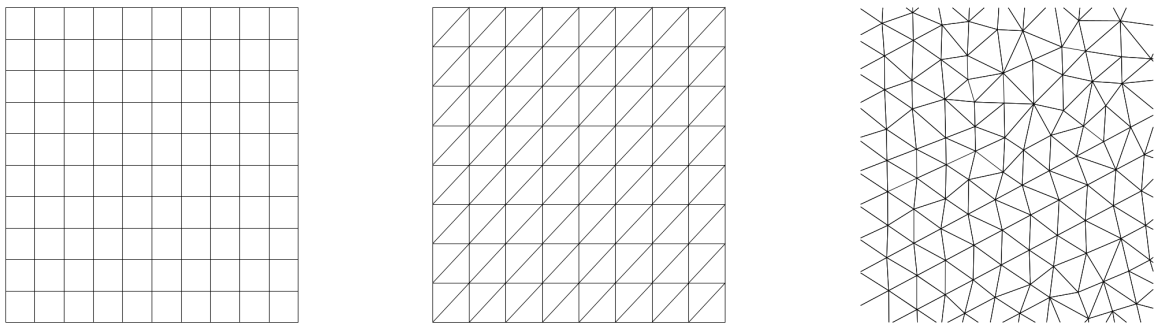


FIGURE 5.3 – De gauche à droite, illustration de cellules uniformes d'une grille cartésienne, de cellules triangulaires structurées et de cellules triangulaires non structurées générées par le logiciel `FreeFem++` [120].

Les grilles cartésiennes satisfont la définition de maillage admissible donnée en Section 5.2.1, Définition 5.2.1. Il est admis que les autres maillage employés satisfont cette définition s'il est possible de calculer numériquement l'inverse des matrices  $M_i$  et  $I_4 - A_i$  respectivement définies par (5.21) et par (5.51).

Ces maillages et les schémas numériques d'ordre élevé (5.117) sont utilisés pour simuler les solutions des équations d'Euler pour un gaz parfait diatomique. L'inconnue  $w$  de ces équations est  $(\rho, \rho u, \rho v, \rho E)^T$  et elle prend ses valeurs dans l'espace

$$\Omega = \left\{ (\rho, \rho u, \rho v, \rho E)^T \in \mathbb{R}^4 \mid \rho > 0, E - (u^2 + v^2)/2 > 0 \right\}. \quad (5.125)$$

Pour un réel  $\gamma = 1,4$  et pour une fonction  $p : \Omega \rightarrow \mathbb{R}_*^+$  telle que

$$p(w) = (\gamma - 1) \left( \rho E - \frac{\rho u^2}{2} - \frac{\rho v^2}{2} \right),$$

les deux fonctions de flux  $f_1, f_2 : \Omega \rightarrow \mathbb{R}^4$  des équations d'Euler pour un gaz parfait diatomique sont données par

$$f_1(w) = (\rho u, \rho u^2 + p, \rho uv, (\rho E + p)u)^T, \quad f_2(w) = (\rho v, \rho uv, \rho v^2 + p, (\rho E + p)v)^T.$$

Parmi les entropies possibles [22, 88, 138], le couple entropie convexe, flux d'entropie  $(\eta, G)$  retenu pour les équations d'Euler est le suivant :

$$\eta(w) = -\rho \ln\left(\frac{p}{\rho^\gamma}\right), \quad G(w) = -\rho \ln\left(\frac{p}{\rho^\gamma}\right)(u, v)^T.$$

À partir de ces conditions, les schémas numériques énoncés dans le Théorème 5.4.1 nécessitent la donnée d'un triplet  $(\Theta_{ij}^{Ok}, \lambda, \Delta t)$ . Bien que des formules explicites soient établies pour ces trois paramètres dans (5.92), (5.113), (5.124), il est clair que les expressions proposées sont pratiquement rédhibitoires. Par conséquent, une approche heuristique est utilisée pour sélectionner le triplet  $(\Theta_{ij}^{Ok}, \lambda, \Delta t)$ . Cette approche heuristique est fondée sur l'analyse menée en Section 4.3 pour le cas d'une dimension d'espace et elle est définie par les éléments suivants.

Tout d'abord, deux configurations sont proposées pour les matrices  $\Theta_{ij}^{Ok}$  dans  $\mathcal{M}_{4,4}(\mathbb{R})$ . La première ne dépend que de la géométrie de l'interface  $\Gamma_{ij}$  et elle est donnée par

$$\Theta_{a,ij}^{Ok} = \frac{d_{j,ij}}{d_{i,j}} I_4. \quad (5.126)$$

La seconde configuration est notée  $\Theta_{b,ij}^{Ok}$  et elle dépend non linéairement des gradients discrets (5.31) par l'intermédiaire des réels  $\zeta_{ij,l}^{O(k-1)}$  tels que

$$\zeta_{ij,l}^{O(k-1)} = \left( \sum_{\mu=1}^2 \frac{(x_j - x_i)_\mu}{d_{i,j}} (\partial_\mu w_i^n)^{O(k-1)(l)} \right)^2, \quad \forall l \in \{1, \dots, 4\}.$$

À partir de ces coefficients,  $\Theta_{b,ij}^{Ok}$  est donné par

$$\Theta_{b,ij}^{Ok} = \text{diag}_{1 \leq l \leq 4} \tanh(\zeta_{ij,l}^{O(k-1)} - \zeta_{ji,l}^{O(k-1)}). \quad (5.127)$$

Ensuite, pour sélectionner les paramètres  $\lambda > 0$  et  $\Delta t > 0$ , une montée en ordre temporel avec les méthodes Runge Kutta *Strong Stability Preserving* [107, 109, 110] est effectuée. Grâce à leurs combinaisons convexes et d'après [76, 230], ces méthodes préservent le résultat du Théorème 5.4.1. En notant  $|\mu_{n_{ij}}(w)|$  le rayon spectral de la matrice  $\nabla f_1(w)n_{ij1} + \nabla f_2(w)n_{ij2}$ , le couple  $(\lambda, \Delta t)$  est choisi selon les restrictions

$$\lambda = \max_{ij \in \mathcal{E}} (|\mu_{n_{ij}}(w_j^n)|, |\mu_{n_{ij}}(w_i^n)|), \quad \frac{\lambda \Delta t |P_i|}{|C_i|} \leq \frac{1}{2}, \quad \forall i \in \mathcal{T}. \quad (5.128)$$

Enfin, en notant  $\partial \mathcal{E}$  les interfaces du bord de  $\mathcal{T}$  et  $j^*$  les indices des cellules fantômes, les choix heuristiques exposés ci-dessus sont vérifiés numériquement à l'aide de la quantité

$$\frac{\Delta \mathcal{S}}{\Delta t} = \sum_{i \in \mathcal{T}} \frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} |C_i| + \sum_{\Gamma_{ij^*} \in \partial \mathcal{E}} |\Gamma_{ij^*}| n_{ij^*}^T G(w_{j^*}^n). \quad (5.129)$$



La quantité  $\Delta\mathcal{S}/\Delta t$  est une version modifiée du membre de gauche de (5.120) qui comprend les conditions de bord. En effet, et si par simplicité le Théorème 5.4.1 est établi sur un domaine sans bord, la quantité  $\Delta\mathcal{S}/\Delta t$  est nécessaire en pratique car les conditions de bords d'un domaine fermé influencent le bilan d'entropie globale [200]. Ainsi, le triplet  $(\Theta_{ij}^{Ok}, \lambda, \Delta t)$  est convenablement choisi lorsque la quantité  $\Delta\mathcal{S}/\Delta t$  définie par (5.129) est négative : dans ce cas, la stabilité entropique discrète globale des schémas numériques (5.117) est satisfaite.

Pour préserver l'espace  $\Omega$  défini par (5.125) et accroître qualitativement la stabilité générale, une correction *a posteriori* de type MOOD [60, 79] est de plus adoptée. Par conséquent, si des quantités calculées par les schémas d'ordre élevé (5.117) ne sont pas dans  $\Omega$  alors cette procédure de limitation substitue localement ces quantités par des valeurs issues d'un schéma du premier ordre.

Avec ces paramétrages, les cas tests considérés sont le problème du vortex isentropique [55, 199], des problèmes de Riemann bi-dimensionnels [136], l'écoulement d'un fluide dans un tunnel possédant une marche [224] et l'évolution d'une onde de choc le long d'une rampe inclinée [224]. Pour chacun de ces problèmes, les résultats sont donnés en trois temps.

Tout d'abord, le cas test est défini puis une allure de la solution est donnée par le résultat d'un schéma usuel du premier ordre.

Ensuite, les résultats du schéma du second ordre sont présentés par un affichage de la variable  $\rho$ . Dans le cas de solutions régulières, ce résultat graphique est complété par une mesure de l'erreur commise avec les normes  $L^1$ ,  $L^2$  et  $L^\infty$ . Dans le cas de solutions discontinues, l'affichage de la variable  $\rho$  est complété par l'évolution temporelle de  $\Delta\mathcal{S}/\Delta t$  défini par (5.129).

Enfin, les résultats du schéma du troisième ordre sont donnés avec une présentation analogue à celle utilisée pour l'ordre deux.

La prochaine section utilise ces conventions pour présenter le cas test du vortex isentropique qui permet une mesure de l'ordre sur des solutions régulières.

## 5.5.2 Vortex isentropique

### 5.5.2.1 Définition du cas test

Le premier cas test est celui du vortex isentropique tel que décrit dans [55, 199]. Le domaine spatial est le carré  $[0, 10]^2$  et pour  $\phi(x) = (5/2\pi) \exp((1 - \|x - x_0\|_2^2)/2)$  avec  $x_0 = (5; 5)^\top$  dans  $\mathbb{R}^2$ , les conditions initiales sont

$$\rho = \left(1 - \frac{\gamma - 1}{2\gamma} \phi^2(x)\right)^{\frac{1}{\gamma-1}}, \quad (1 - u; 1 - v)^\top = \phi(x)(x - x_0)^\top, \quad p = \left(1 - \frac{\gamma - 1}{2\gamma} \phi^2(x)\right)^{\frac{\gamma}{\gamma-1}}. \quad (5.130)$$

Les conditions de bords sont périodiques. La solution exacte est un vortex régulier transporté le long d'une diagonale du domaine. La Figure 5.4 illustre ce problème avec le résultat d'un schéma usuel du premier ordre sur une grille cartésienne de 6400 cellules et au temps final de simulation  $t = 0,1$ .

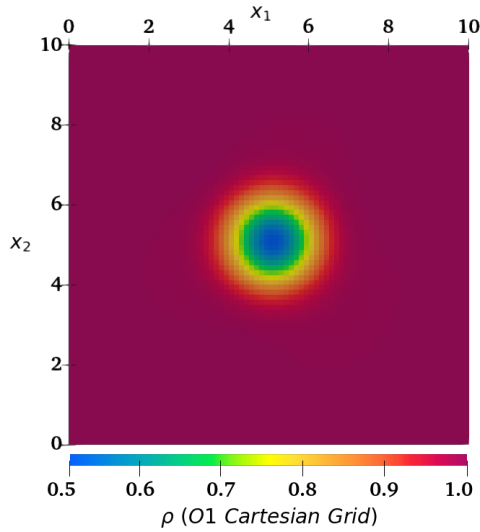
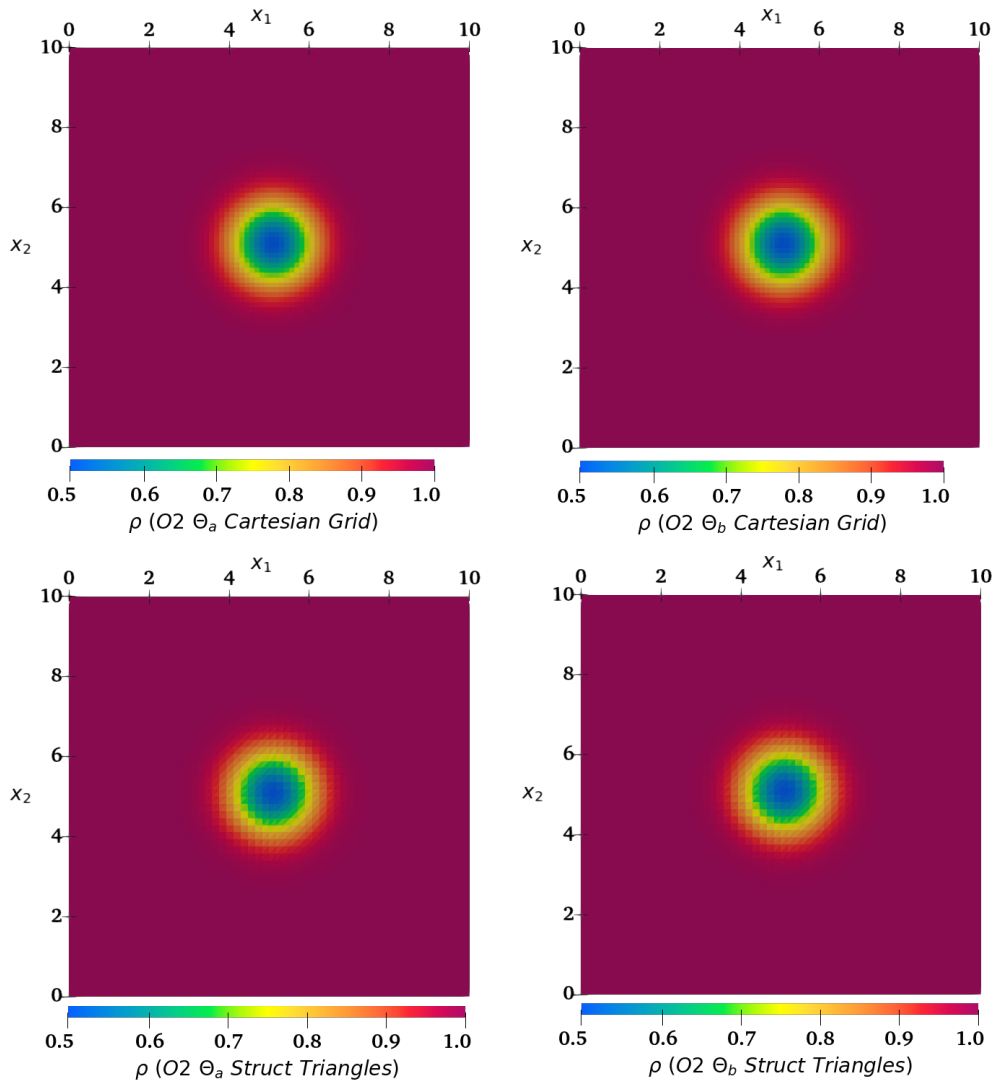


FIGURE 5.4 – Résultat numérique d'un schéma usuel du premier ordre sur une grille cartésienne de 6400 cellules et pour le problème du vortex isentropique (5.130) au temps final  $t = 0,1$ .

### 5.5.2.2 Résultats du schéma du second ordre

Les résultats graphiques du schéma du second ordre (5.117) pour le problème du vortex isentropique (5.130) sont tout d'abord donnés par la Figure 5.5. Ensuite, les tableaux 5.1-5.2 rapportent ses erreurs numériques ainsi que ses ordres pour différents maillages.



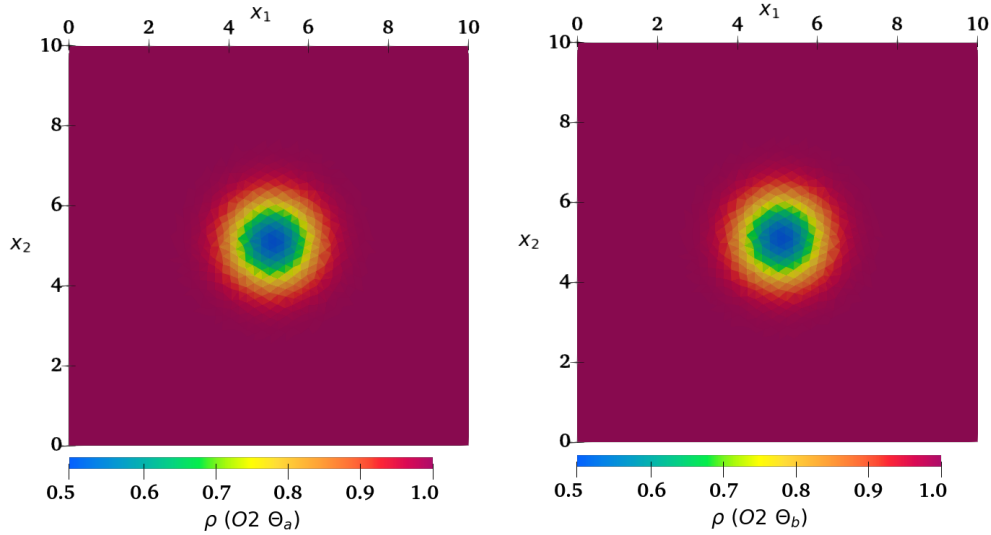


FIGURE 5.5 – Résultats numériques du schéma du second ordre (5.117) pour le problème du vortex isentropique (5.130) au temps final  $t = 0,1$  avec la légende : de gauche à droite résultats pour  $(\Theta_{a,ij}^{O2})_{ij \in \mathcal{E}}$  définie par (5.126) puis pour  $(\Theta_{b,ij}^{O2})_{ij \in \mathcal{E}}$  définie par (5.127). De haut en bas, résultats sur grille cartésienne uniforme de 6400 cellules puis sur 6400 triangles structurés et sur 5262 triangles non structurés générés par le logiciel FreeFem++ [120].

Erreurs du schéma du second ordre avec $(\Theta_{a,ij}^{O2})_{ij \in \mathcal{E}}$ sur grilles cartésiennes							
cellules	$h$	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	1,0E+00	3,0E-01	-	7,3E-02	-	5,1E-02	-
400	5,0E-01	9,7E-02	1,6	2,6E-02	1,5	2,3E-02	1,1
1600	2,5E-01	2,1E-02	2,2	5,3E-03	2,3	4,4E-03	2,4
6400	1,2E-01	4,4E-03	2,3	1,1E-03	2,3	8,6E-04	2,3
10000	1,0E-01	2,7E-03	2,2	6,6E-04	2,2	5,3E-04	2,2

Erreurs du schéma du second ordre avec $(\Theta_{a,ij}^{O2})_{ij \in \mathcal{E}}$ sur triangles structurés							
cellules	$h$	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
128	9,3E-01	3,4E-01	-	7,6E-02	-	3,4E-02	-
512	4,7E-01	9,9E-02	1,8	2,4E-02	1,7	1,5E-02	1,1
1682	2,6E-01	2,9E-02	2,1	6,8E-03	2,1	5,6E-03	1,7
6272	1,3E-01	7,3E-03	2,1	1,7E-03	2,1	1,4E-03	2,1
10658	1,0E-01	4,3E-03	2,0	9,8E-04	2,1	7,9E-04	2,1

Erreurs du schéma du second ordre avec $(\Theta_{a,ij}^{O2})_{ij \in \mathcal{E}}$ sur triangles non structurés							
cellules	$h$	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
116	1,1E+00	3,4E-01	-	7,8E-02	-	4,0E-02	-
524	5,7E-01	9,1E-02	2,0	2,0E-02	2,1	1,3E-02	1,7
1148	4,2E-01	4,7E-02	2,1	1,1E-02	1,9	7,3E-03	1,8
5262	1,9E-01	1,1E-02	1,8	2,7E-03	1,8	3,3E-03	1,0
10394	1,4E-01	5,9E-03	2,0	1,5E-03	1,9	1,7E-03	2,1

TABLE 5.1 – Évaluation de l'erreur et des ordres pour le schéma du second ordre (5.117) paramétré par  $(\Theta_{a,ij}^{O2})_{ij \in \mathcal{E}}$  définie dans (5.126), avec  $h = \max_{ij \in \mathcal{E}} d_{ij}$  et pour le problème du vortex isentropique (5.130) au temps final  $t = 0,1$ .

Erreurs du schéma du second ordre avec  $(\Theta_{b,ij}^{O2})_{ij \in \mathcal{E}}$  sur grilles cartésiennes

cellules	$h$	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	1,0E+00	3,0E-01	-	7,3E-02	-	5,2E-02	-
400	5,0E-01	9,9E-02	1,6	2,8E-02	1,4	2,7E-02	0,9
1600	2,5E-01	2,2E-02	2,2	5,8E-03	2,3	5,3E-03	2,4
6400	1,2E-01	4,6E-03	2,2	1,2E-03	2,2	1,2E-03	2,1
10000	1,0E-01	2,9E-03	2,0	7,9E-04	2,0	8,5E-04	1,7

Erreurs du schéma du second ordre avec  $(\Theta_{b,ij}^{O2})_{ij \in \mathcal{E}}$  sur triangles structurés

cellules	$h$	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
128	9,3E-01	3,5E-01	-	7,9E-02	-	3,5E-02	-
512	4,7E-01	1,0E-01	1,8	2,6E-02	1,6	1,7E-02	1,0
1682	2,6E-01	3,0E-02	2,0	7,2E-03	2,1	6,7E-03	1,6
6272	1,3E-01	7,5E-03	2,1	1,8E-03	2,1	1,9E-03	1,9
10658	1,0E-01	4,4E-03	2,1	1,0E-03	2,1	1,1E-03	2,1

Erreurs du schéma du second ordre avec  $(\Theta_{b,ij}^{O2})_{ij \in \mathcal{E}}$  sur triangles non structurés

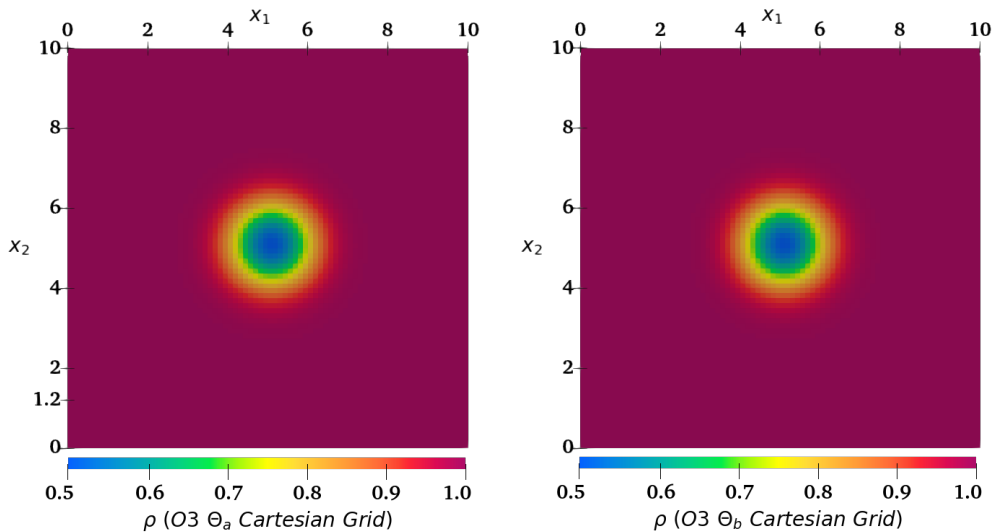
cellules	$h$	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
116	1,1E+00	3,4E-01	-	8,2E-02	-	4,9E-02	-
524	5,7E-01	9,3E-02	2,0	2,1E-02	2,1	1,3E-02	2,0
1148	4,2E-01	4,4E-02	2,4	1,0E-02	2,3	6,8E-03	2,1
5262	1,9E-01	1,0E-02	1,8	2,4E-03	1,8	2,2E-03	1,5
10394	1,4E-01	5,0E-03	2,3	1,2E-03	2,3	1,1E-03	2,1

TABLE 5.2 – Évaluation de l’erreur et des ordres pour le schéma du second ordre (5.117) paramétré par  $(\Theta_{b,ij}^{O2})_{ij \in \mathcal{E}}$  définie dans (5.127), avec  $h = \max_{ij \in \mathcal{E}} d_{ij}$  et pour le problème du vortex isentropique (5.130) au temps final  $t = 0,1$ .

Puisque le schéma du second ordre (5.117) n’utilise pas de limitation de pente, les ordres théoriques sont exactement observés dès que le maillage est suffisamment fin.

### 5.5.2.3 Résultats du schéma du troisième ordre

Les résultats du schéma du troisième ordre simulant le problème du vortex isentropique (5.130) sont donnés dans cette section. Les résultats graphiques sont illustrés par la Figure 5.6 et les tableaux 5.3-5.4 rapportent les erreurs numériques ainsi que les ordres pour différents maillages.



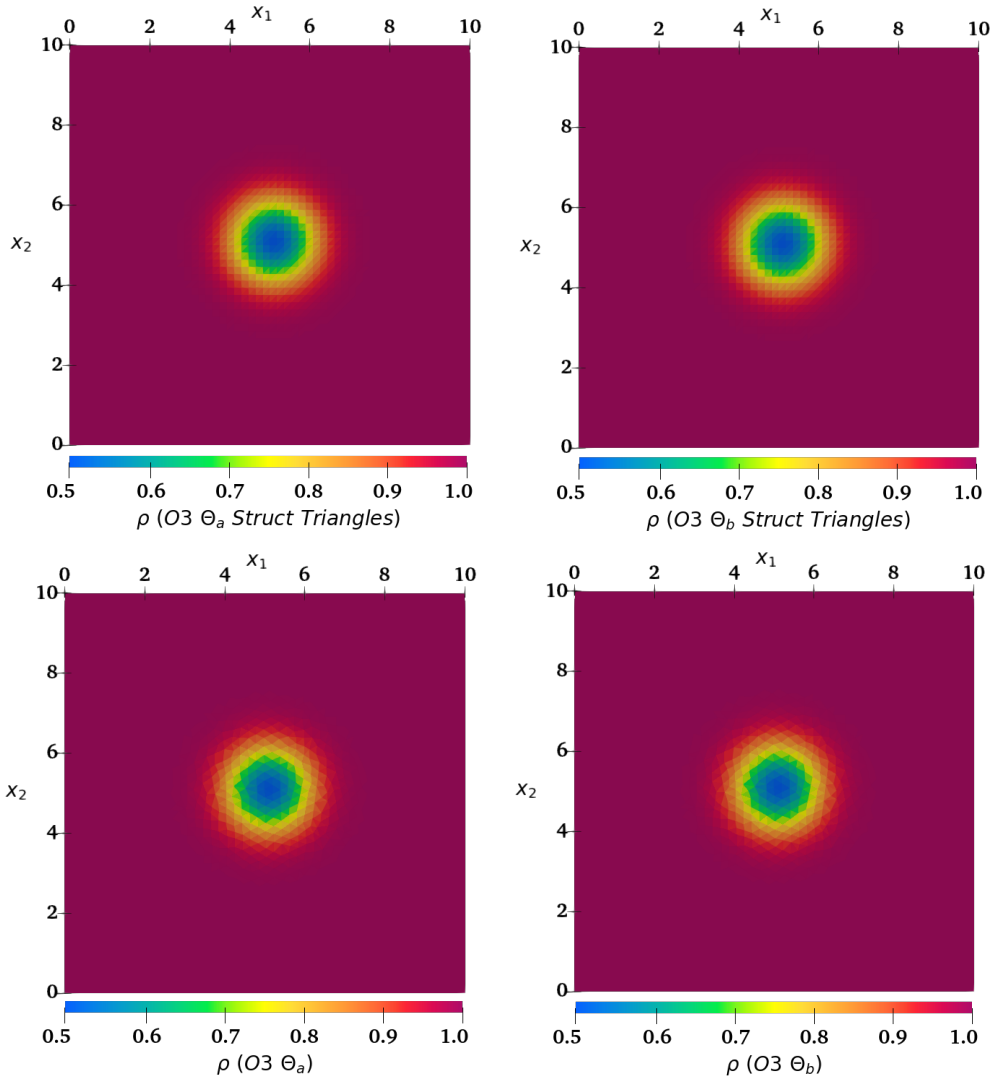


FIGURE 5.6 – Résultats numériques du schéma du troisième ordre (5.117) pour le problème du vortex isentropique (5.130) au temps final  $t = 0,1$  avec la légende : de gauche à droite résultats pour  $(\Theta_{a,ij}^{O3})_{ij \in \mathcal{E}}$  définie par (5.126) puis pour  $(\Theta_{b,ij}^{O3})_{ij \in \mathcal{E}}$  définie par (5.127). De haut en bas, résultats sur grille cartésienne uniforme de 6400 cellules puis sur 6400 triangles structurés et sur 5262 triangles non structurés générés par le logiciel FreeFem<sub>++</sub> [120].

Erreurs du schéma du troisième ordre avec $(\Theta_{a,ij}^{O3})_{ij \in \mathcal{E}}$ sur grilles cartésiennes							
cellules	$h$	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	1,0E+00	2,4E-01	-	5,3E-02	-	3,7E-02	-
400	5,0E-01	3,8E-02	2,7	9,8E-03	2,5	7,0E-03	2,4
1600	2,5E-01	3,0E-03	3,7	8,1E-04	3,6	6,4E-04	3,4
6400	1,2E-01	2,2E-04	3,8	6,1E-05	3,7	6,0E-05	3,4
10000	1,0E-01	1,0E-04	3,5	3,0E-05	3,2	2,9E-05	3,3

Erreurs du schéma du troisième ordre avec $(\Theta_{a,ij}^{O3})_{ij \in \mathcal{E}}$ sur triangles structurés							
cellules	$h$	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
128	9,3E-01	2,7E-01	-	5,8E-02	-	2,9E-02	-
512	4,7E-01	4,1E-02	2,7	1,0E-02	2,5	6,6E-03	2,2
1682	2,6E-01	5,5E-03	3,4	1,5E-03	3,3	1,1E-03	3,0
6272	1,3E-01	5,7E-04	3,4	1,6E-04	3,4	1,1E-04	3,6
10658	1,0E-01	2,5E-04	3,1	6,9E-05	3,1	4,8E-05	3,0

Erreurs du schéma du troisième ordre avec $(\Theta_{a,ij}^{O3})_{ij \in \mathcal{E}}$ sur triangles non structurés							
cellules	$h$	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
116	1,1E+00	2,4E-01	-	5,7E-02	-	2,9E-02	-
524	5,7E-01	2,9E-02	3,3	7,3E-03	3,2	4,9E-03	2,7
1148	4,2E-01	1,1E-02	2,9	2,8E-03	3,1	1,9E-03	2,9
5262	1,9E-01	1,3E-03	2,8	3,2E-04	2,8	3,3E-04	2,3
10394	1,4E-01	5,4E-04	2,7	1,5E-04	2,4	2,0E-04	1,5

TABLE 5.3 – Évaluation de l’erreur et des ordres pour le schéma du troisième ordre (5.117) paramétré par  $(\Theta_{a,ij}^{O3})_{ij \in \mathcal{E}}$  définie dans (5.126), avec  $h = \max_{ij \in \mathcal{E}} d_{ij}$  et pour le problème du vortex isentropique (5.130) au temps final  $t = 0,1$ .

Erreurs du schéma du troisième ordre avec $(\Theta_{b,ij}^{O3})_{ij \in \mathcal{E}}$ sur grilles cartésiennes							
cellules	$h$	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
100	1,0E+00	2,4E-01	-	5,3E-02	-	3,7E-02	-
400	5,0E-01	3,8E-02	2,6	1,0E-02	2,4	7,8E-03	2,2
1600	2,5E-01	3,0E-03	3,7	8,2E-04	3,6	7,1E-04	3,5
6400	1,2E-01	2,3E-04	3,7	7,0E-05	3,6	6,3E-05	3,5
10000	1,0E-01	1,1E-04	3,4	3,4E-05	3,2	2,9E-05	3,4

Erreurs du schéma du troisième ordre avec $(\Theta_{b,ij}^{O3})_{ij \in \mathcal{E}}$ sur triangles structurés							
cellules	$h$	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
128	9,3E-01	2,7E-01	-	5,8E-02	-	3,0E-02	-
512	4,7E-01	4,1E-02	2,7	1,1E-02	2,5	6,9E-03	2,1
1682	2,6E-01	5,5E-03	3,4	1,5E-03	3,3	1,2E-03	3,0
6272	1,3E-01	5,7E-04	3,4	1,6E-04	3,4	1,1E-04	3,6
10658	1,0E-01	2,5E-04	3,1	6,9E-05	3,1	4,9E-05	3,1

Erreurs du schéma du troisième ordre avec $(\Theta_{b,ij}^{O3})_{ij \in \mathcal{E}}$ sur triangles non structurés							
cellules	$h$	$L^1$	ordre	$L^2$	ordre	$L^\infty$	ordre
116	1,1E+00	2,5E-01	-	5,9E-02	-	3,2E-02	-
524	5,7E-01	2,9E-02	3,3	7,3E-03	3,2	4,9E-03	2,9
1148	4,2E-01	1,1E-02	2,9	2,8E-03	3,1	1,7E-03	3,4
5262	1,9E-01	1,3E-03	2,8	3,2E-04	2,8	3,3E-04	2,1
10394	1,4E-01	5,5E-04	2,7	1,5E-04	2,4	2,0E-04	1,6

TABLE 5.4 – Évaluation de l’erreur et des ordres pour le schéma du troisième ordre (5.117) paramétré par  $(\Theta_{b,ij}^{O3})_{ij \in \mathcal{E}}$  définie dans (5.127), avec  $h = \max_{ij \in \mathcal{E}} d_{ij}$  et pour le problème du vortex isentropique (5.130) au temps final  $t = 0,1$ .

Les ordres théoriques sont une nouvelle fois exactement restitués sur grilles cartésiennes et sur maillages triangulaires structurés. En revanche, une légère dégradation de la précision est observée sur maillages non structurés. L’origine exacte de cette dégradation est inconnue mais les ordres observés demeurent satisfaisants.

### 5.5.3 Cas des problèmes de Riemann en dimension deux

#### 5.5.3.1 Définitions des cas tests

Cette section concerne trois problèmes de Riemann en dimension deux tels que décrits par A. Kurganov et E. Tadmor dans [136]. Pour ces problèmes, le carré unité est d’abord divisé en quatre cadrants délimités par les droites  $x_1 = 1/2$  et  $x_2 = 1/2$ . Ensuite, dans chaque cadrant, la quantité  $(\rho, u, v, p)^T$  est initialisée par une constante de  $\Omega$ . Les trois problèmes considérés sont notés KT3, KT5 et KT6 et le tableau 5.5 donne les valeurs des constantes d’initialisation.

Quadrant	problème KT3	problème KT5	problème KT6
$[0, 1/2] \times [0, 1/2]$	$(0,138; 1,206; 1,206; 0,029)^T$	$(1; 0,75; 0,5; 1)^T$	$(1; -0,75; 0,5; 1)^T$
$[0, 1/2] \times [1/2, 1]$	$(0,5323; 1,206; 0; 0,3)^T$	$(2; -0,75; 0,5; 1)^T$	$(2; 0,75; 0,5; 1)^T$
$[1/2, 1] \times [0, 1/2]$	$(0,5323; 0; 1,206; 0,3)^T$	$(3; 0,75; -0,5; 1)^T$	$(3; -0,75; -0,5; 1)^T$
$[1/2, 1] \times [1/2, 1]$	$(1,5; 0; 0; 1,5)^T$	$(1; -0,75; -0,5; 1)^T$	$(1; 0,75; -0,5; 1)^T$

TABLE 5.5 – Données initiales des problèmes de Riemann bidimensionnels.

Des conditions de bords de types Neumann homogènes sont imposées sur toutes les frontières du domaine et pour chaque problème, le temps final de simulation est  $t = 0,3$ . Pour illustrer ces définitions, une solution obtenue avec un schéma usuel du premier ordre est donnée par la Figure 5.7.

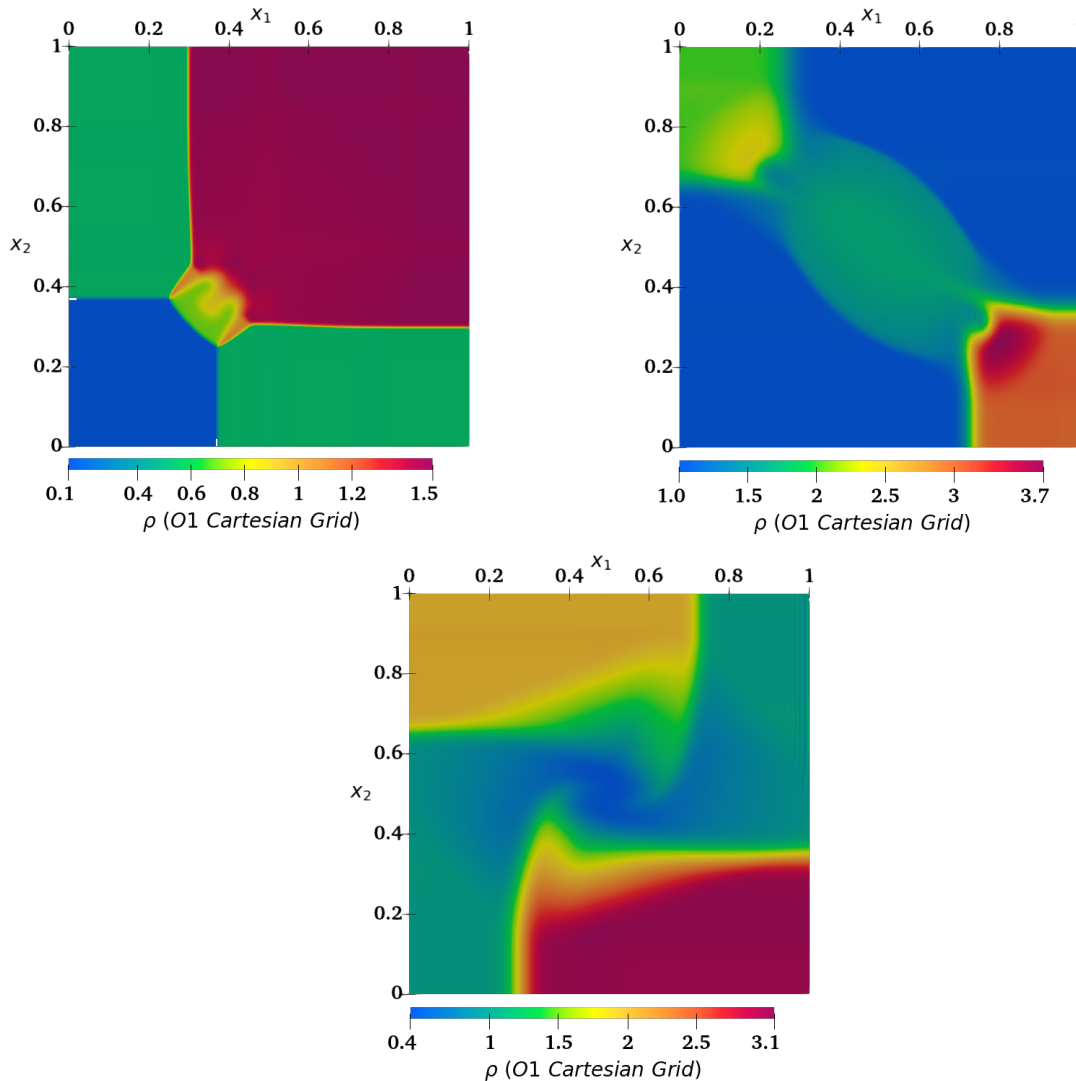


FIGURE 5.7 – En haut à gauche (*resp.* en haut à droite, en bas) résultats numériques d'un schéma usuel du premier ordre sur une grille cartésienne ayant 1000 cellules par direction, au temps final  $t = 0,3$  et pour le problème de Riemann KT3 (*resp.* KT5, KT6) donné par le tableau 5.5.

### 5.5.3.2 Résultats du schéma du second ordre

Cette section concerne les résultats du schéma du second ordre (5.117) pour les problèmes de Riemann définis ci-avant. La Figure 5.8 donne les résultats pour le problème KT3 et la Figure 5.9 montre ceux obtenus pour le problème KT5.

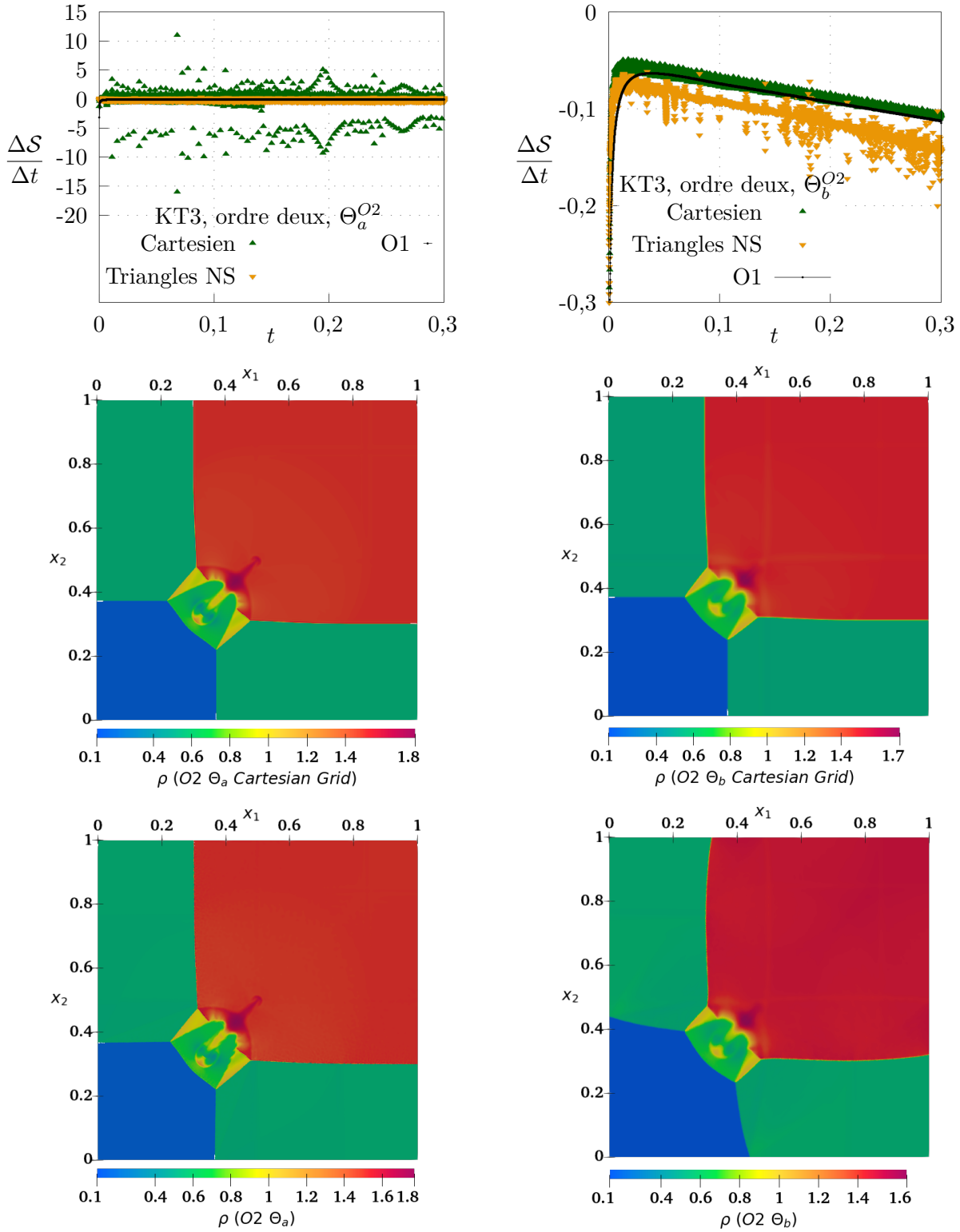


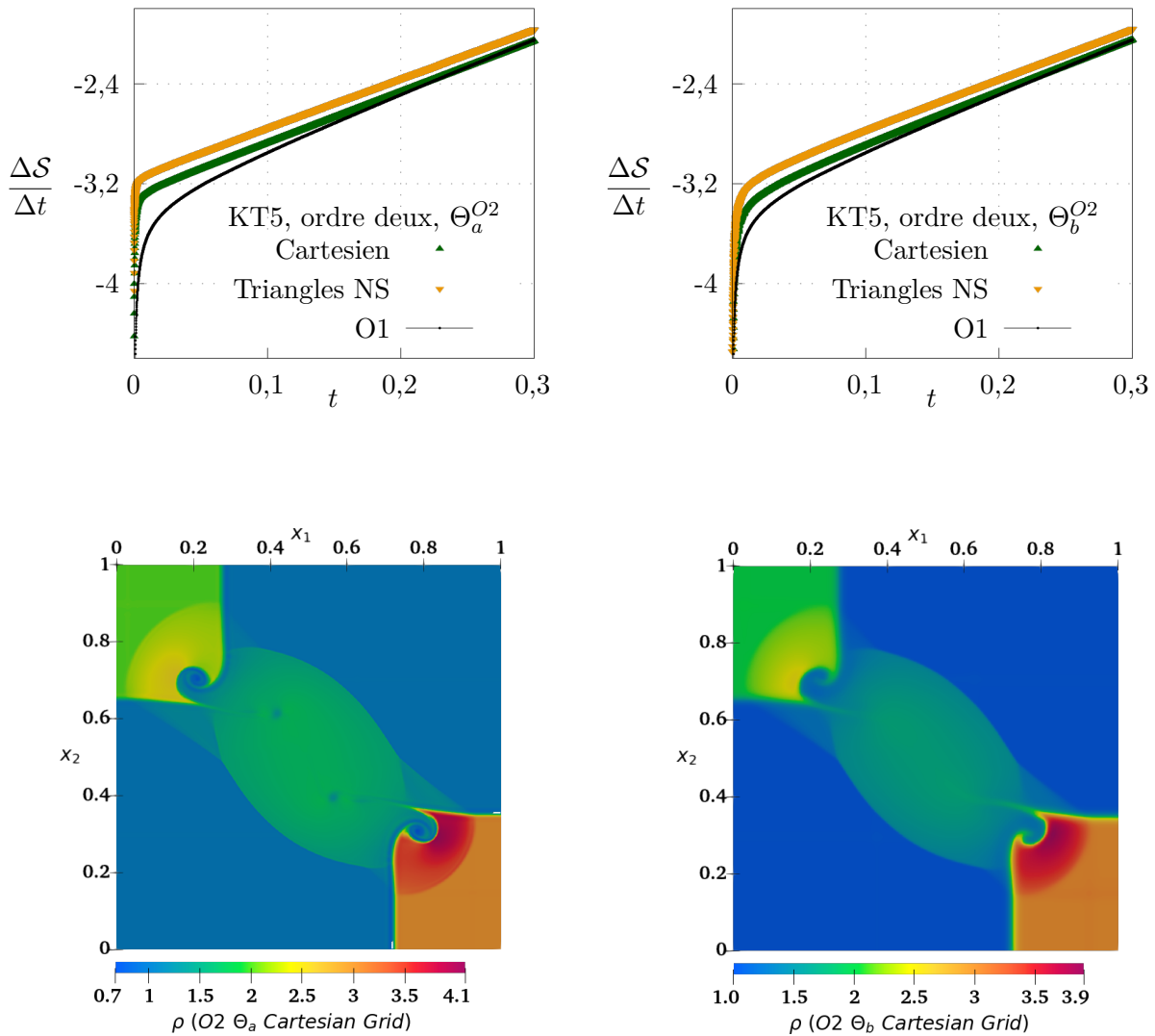
FIGURE 5.8 – Résultats numériques du schéma du second ordre (5.117) au temps final  $t = 0,3$  pour le problème de Riemann KT3 donné par le tableaux 5.5. De gauche à droite, résultats pour  $(\Theta_{a,ij}^{O2})_{ij \in \mathcal{E}}$  définie par (5.126) puis pour  $(\Theta_{b,ij}^{O2})_{ij \in \mathcal{E}}$  définie par (5.127). De haut en bas, évolution temporelle de  $\Delta\mathcal{S}/\Delta t$  défini dans (5.129), résultats sur grille cartésienne uniforme de 1 000 000 cellules puis sur 1 002 422 triangles non structurés générés par le logiciel FreeFem<sub>++</sub> [120].

Les solutions numériques obtenues pour le problème KT3 avec le triplet  $(\Theta_{a,ij}^{O2}, \lambda, \Delta t)$  défini par (5.126), (5.128) sont globalement moins diffuses que les solutions standard de la Figure 5.7. Pour autant, et d'après l'évolution temporelle de  $\Delta\mathcal{S}/\Delta t$  tracée en haut à gauche de la Figure 5.8, les solutions calculées



avec ce triplet ne vérifient pas l'inégalité d'entropie discrète globale.

Ce défaut de stabilité est corrigé avec les matrices  $\Theta_{b,ij}^{O2}$  données par (5.127). Néanmoins, avec ces matrices et sur les deux types de maillages, des oscillations de  $\Delta S/\Delta t$  sont observés en haut à droite de la Figure 5.8. L'origine de ces oscillations est inconnue mais leur aspect erratique sur maillage non structuré pourrait être attribué à l'utilisation de la procédure MOOD. En effet, une mesure directe montre que cette procédure est seulement activée sur ces maillages et elle concerne au plus  $10^{-4}\%$  du total des cellules au cours de la simulation. Ce léger défaut de robustesse sur maillages non structurés pourrait également expliquer les déformations des ondes observées dans les régions proches de  $(x_1; x_2)^\top = (0,4; 0)^\top$  et de  $(x_1; x_2)^\top = (0; 0,4)^\top$ .



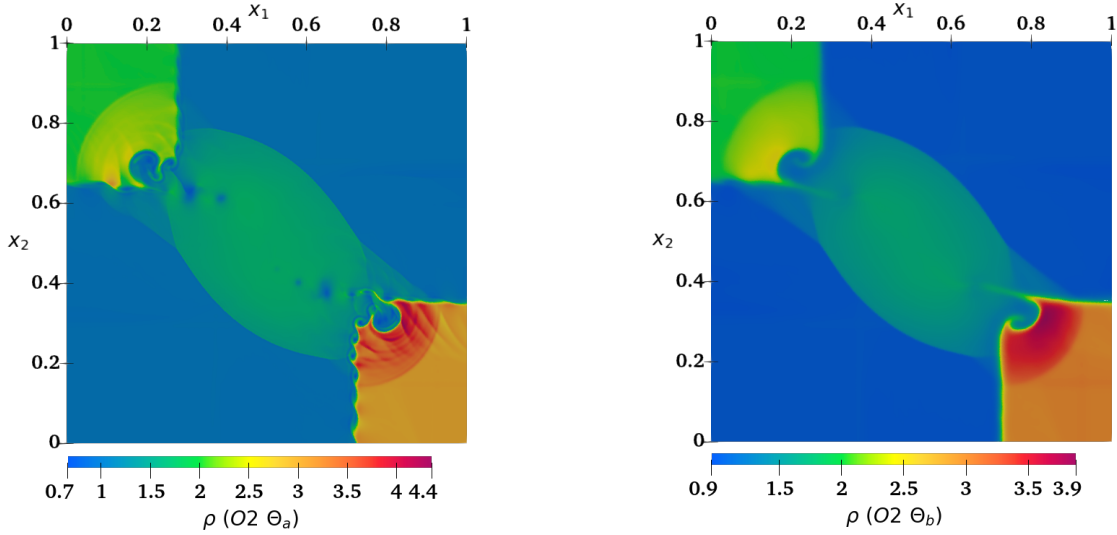
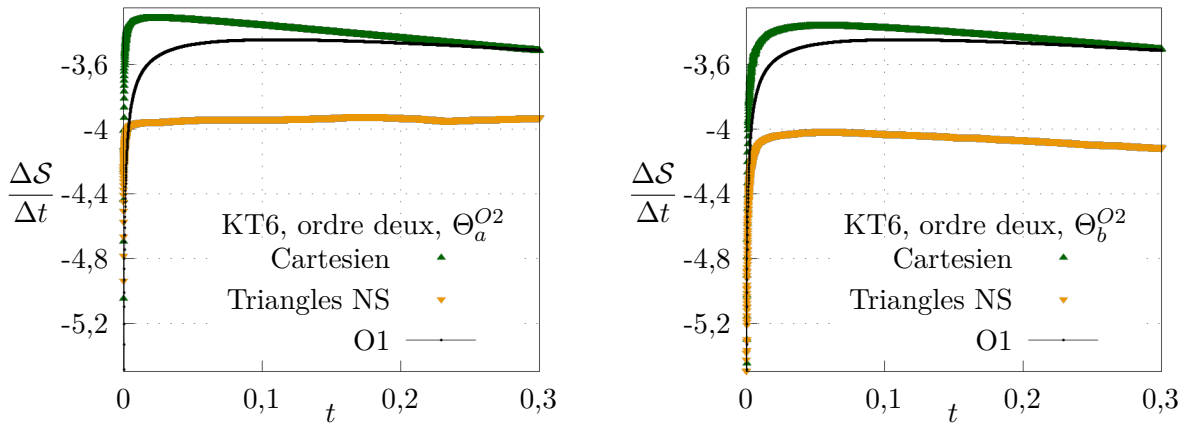


FIGURE 5.9 – Résultats numériques du schéma du second ordre (5.117) au temps final  $t = 0,3$  pour le problème de Riemann KT5 défini dans le tableaux 5.5. De gauche à droite, résultats pour  $(\Theta_{a,ij}^{O2})_{ij \in \mathcal{E}}$  définie par (5.126) puis pour  $(\Theta_{b,ij}^{O2})_{ij \in \mathcal{E}}$  définie par (5.127). De haut en bas, évolution temporelle de  $\Delta S / \Delta t$  défini par (5.129), résultats sur grille cartésienne uniforme de 1 000 000 cellules puis sur 1 002 422 triangles non structurés générés par le logiciel FreeFem<sub>++</sub> [120].

Pour le problème KT5, les paramétrages  $\Theta_{a,ij}^{O2}$  et  $\Theta_{b,ij}^{O2}$  respectivement donnés par (5.126) et par (5.127) assurent tous deux la stabilité entropique discrète sur grilles cartésiennes et sur maillages non structurés. De plus, la procédure de limitation MOOD n’est pas utilisée.

Par ailleurs, des irrégularités sont observées sur les fronts des ondes calculés par le paramétrage  $\Theta_{a,ij}^{O2}$  sur maillages non structurés (voir Fig 5.9 en bas à gauche). Ces irrégularités sont des instabilités de Kelvin-Helmholtz et d’après [60], leurs apparitions est une mesure qualitative de la diffusion numérique. D’après la Figure 5.9, ces instabilités n’apparaissent pas sur grilles cartésiennes et par conséquent, les résultats les plus précis sont probablement ceux donnés par les matrices  $\Theta_{a,ij}^{O2}$  définies par (5.126) et sur maillages non structuré.

Les mêmes remarques peuvent être formulées pour le problème KT6 dont les résultats sont présentés par la Figure 5.10.



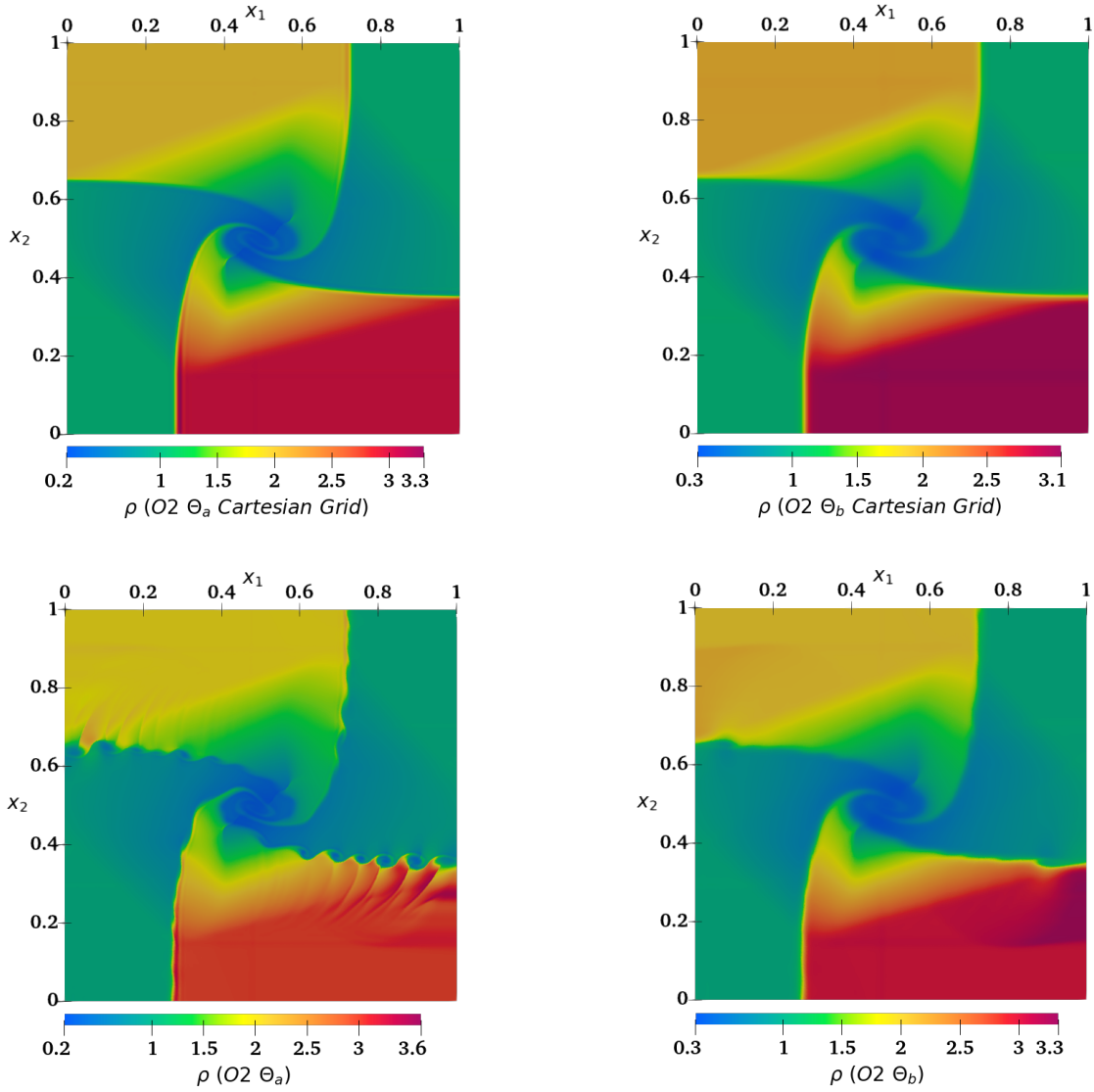


FIGURE 5.10 – Résultats numériques du schéma du second ordre (5.117) au temps final  $t = 0,3$  pour le problème de Riemann KT6 donné par le tableau 5.5. De gauche à droite, résultats pour  $(\Theta_{a,ij}^{O2})_{ij \in \mathcal{E}}$  définie par (5.126) puis pour  $(\Theta_{b,ij}^{O2})_{ij \in \mathcal{E}}$  définie par (5.127). De haut en bas, évolution temporelle de  $\Delta S / \Delta t$  défini par (5.129), résultats sur grille cartésienne uniforme de 1 000 000 cellules puis sur 1 002 422 triangles non structurés générés par le logiciel FreeFem<sub>++</sub> [120].

### 5.5.3.3 Résultats du schéma du troisième ordre

Cette section présente les résultats du schéma du troisième ordre (5.117) pour les problèmes de Riemann bi-dimensionnels définis dans le tableau 5.5.

Malgré les efforts déployés, les triplets  $((\Theta_{\alpha,ij}^{O3})_{\alpha \in \{a,b\}}, \lambda, \Delta t)$  proposés dans (5.126), (5.127), (5.128) demeurent parfois instables pour le schéma du troisième ordre et certaines simulations n'ont pas pu être réalisées sur maillages non structurés. Dans ces cas, les résultats sont présentés sur des maillages triangulaires structurés tel que décrit en Section 5.5.1, Figure 5.3.

La Figure 5.11 montre les résultats du schéma du troisième ordre pour le problème KT3 et les résultats obtenus pour le problème KT5 sont donnés en Figure 5.12.

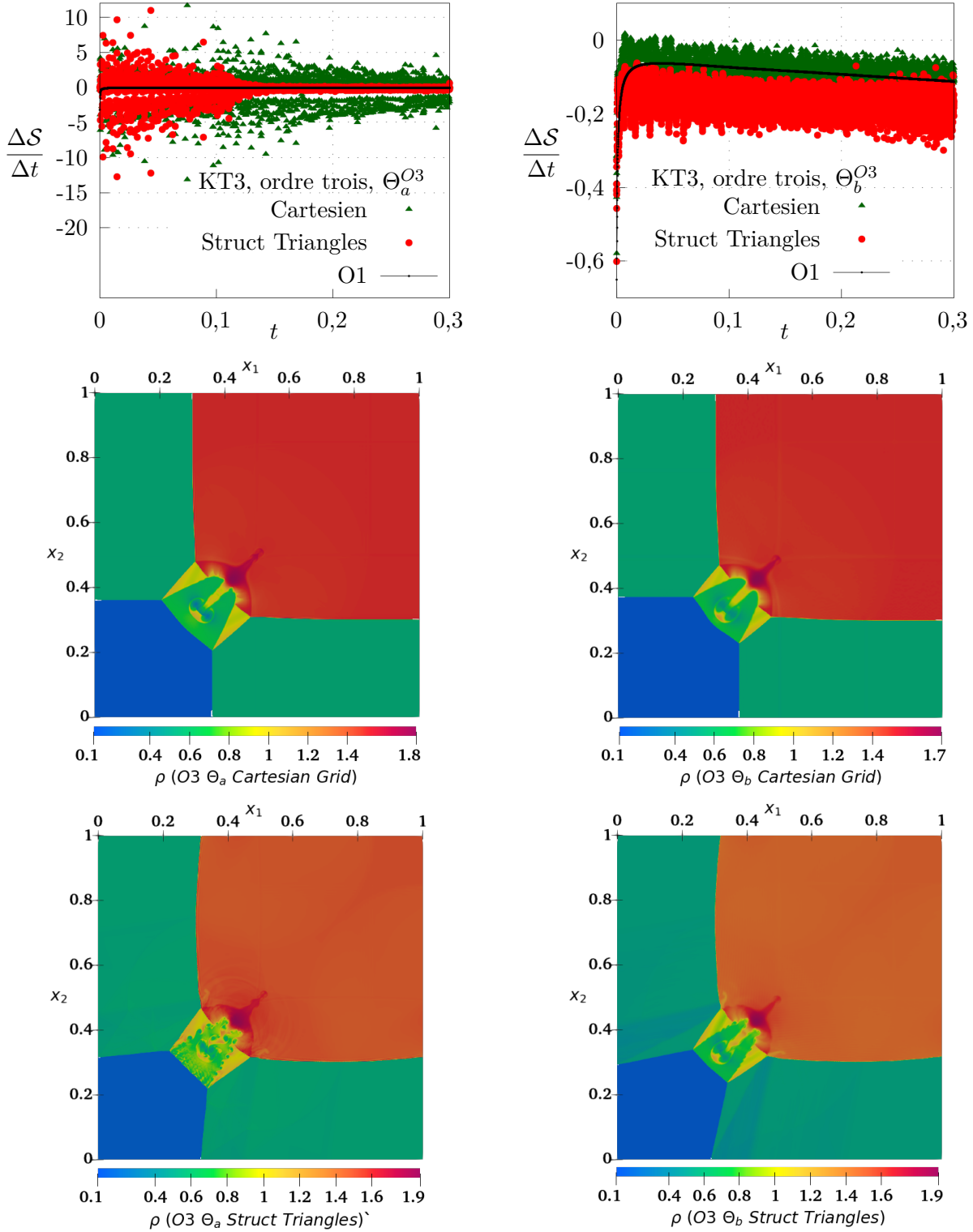
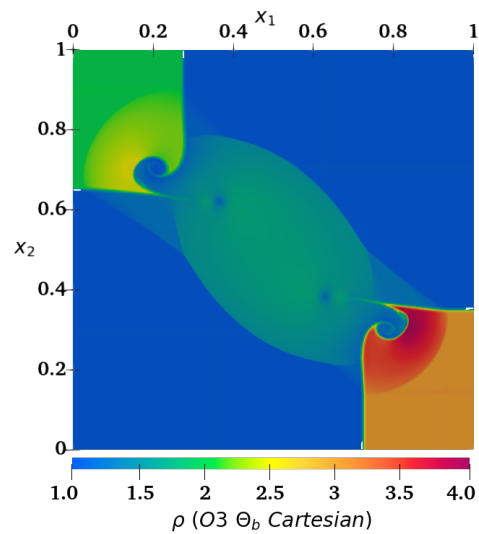
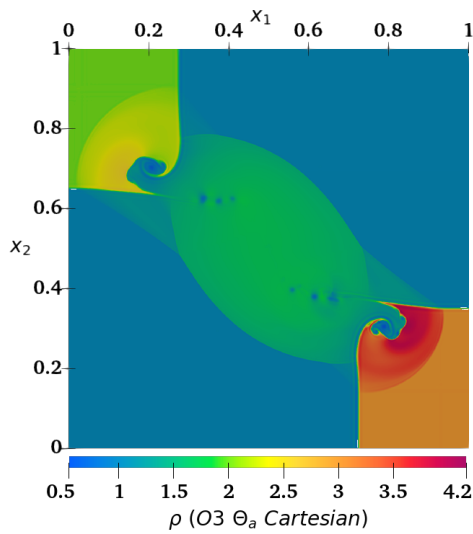
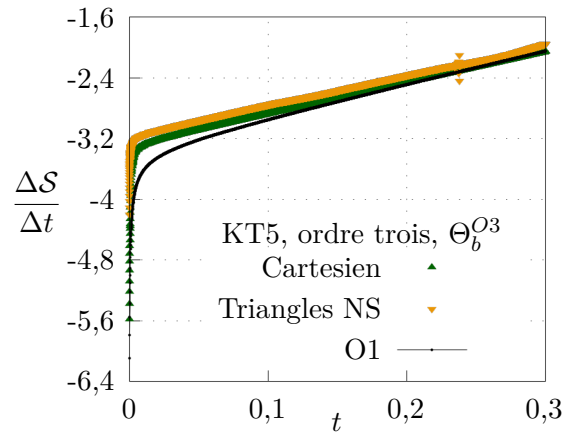
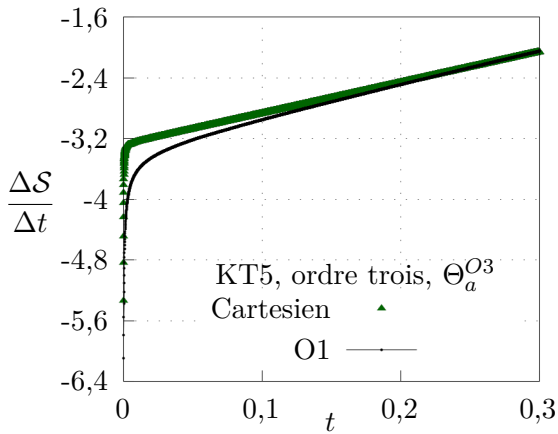


FIGURE 5.11 – Résultats numériques du schéma du troisième ordre (5.117) au temps final  $t = 0,3$  pour le problème de Riemann KT3 donné dans le tableau 5.5. De gauche à droite, résultats pour  $(\Theta_{a,ij}^{O3})_{ij \in \mathcal{E}}$  définie par (5.126) puis pour  $(\Theta_{b,ij}^{O3})_{ij \in \mathcal{E}}$  définie par (5.127). De haut en bas, évolution temporelle de  $\Delta S/\Delta t$  défini par (5.129), résultats sur grille cartésienne uniforme de 1 000 000 cellules puis sur 1 008 200 triangles structurés.

Concernant le problème KT3, les quatre simulations donnent des résultats différents et pour toutes la procédure de limitation MOOD est nécessaire. Avec les matrices  $\Theta_{a,ij}^{O3}$  définies par (5.126), cette procédure s'active pour au plus 0,1% du total des cellules et avec les matrices  $\Theta_{b,ij}^{O3}$  exprimées dans (5.126), elle concerne au plus 0,0025% du total des cellules. Ces proportions suggèrent que les simula-

tions numériques sont plus stables avec le choix de  $\Theta_{b,ij}^{O3}$ . Cette tendance est clairement confirmée par le tracé de la quantité  $\Delta S/\Delta t$  qui est erratique pour  $\Theta_{a,ij}^{O3}$  mais demeure négative pour  $\Theta_{b,ij}^{O3}$  sur maillage triangulaire structuré.

L'association de ces maillages au schéma d'ordre trois fait de plus apparaître des instabilités de Kelvin-Helmholtz dans la zone d'interaction centrale approximativement située dans le rectangle  $[0,4; 1] \times [0,2; 0,5]$ . Ces instabilités n'étaient pas présentes dans la Figure 5.8 qui illustre le schéma du second ordre. Par conséquent, le gain de précision obtenue par une montée à l'ordre trois semble réel mais des déformations des ondes sont une nouvelle fois observées dans les régions proches de  $(x_1; x_2)^T = (0,4; 0)^T$  et de  $(x_1; x_2)^T = (0; 0,4)^T$ .



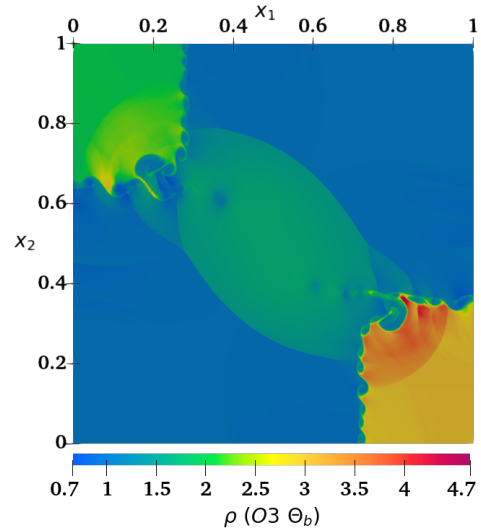


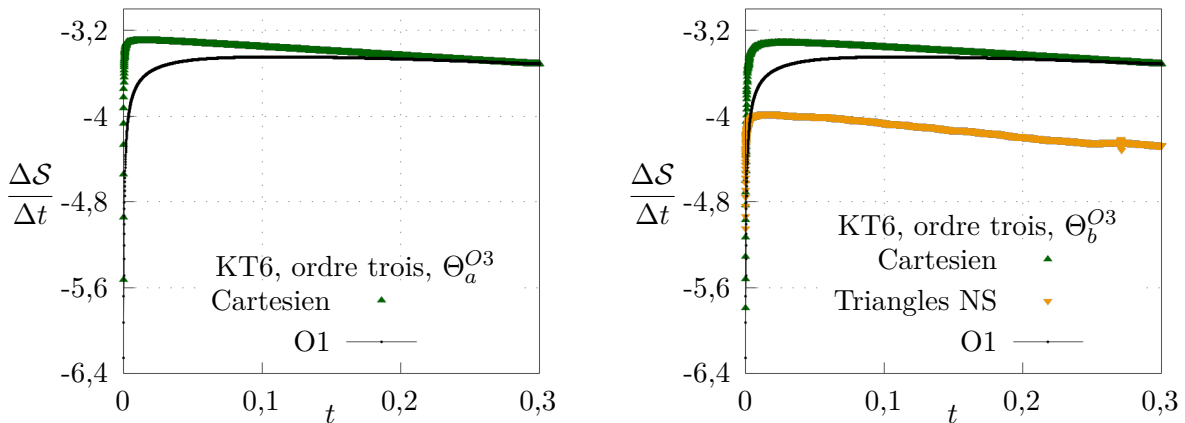
FIGURE 5.12 – Résultats numériques du schéma du troisième ordre (5.117) au temps final  $t = 0,3$  pour le problème de Riemann KT5 donné par le tableau 5.5. De gauche à droite, résultats pour  $(\Theta_{a,ij}^{O3})_{ij \in \mathcal{E}}$  définie par (5.126) puis pour  $(\Theta_{b,ij}^{O3})_{ij \in \mathcal{E}}$  définie par (5.127). De haut en bas, évolution temporelle de  $\Delta S/\Delta t$  défini par (5.129), résultats sur grille cartésienne uniforme de 1 000 000 cellules puis sur 1 002 422 triangles non structurés générés par le logiciel FreeFem<sub>++</sub> [120].

Pour le problème KT5, le choix de  $\Theta_{a,ij}^{O3}$  défini par (5.126) n'est ni fonctionnel sur maillage triangulaire structuré ni sur maillage non structuré. En revanche, ce choix donne des résultats satisfaisants sur grille cartésienne.

L'ensemble des résultats présentés en Figure 5.12 pour le cas test KT5 vérifie la stabilité entropique discrète globale. La procédure MOOD n'est nécessaire que sur les maillages triangulaires non structurés. Dans ce cas, son activation concerne au plus  $10^{-4}\%$  du total des cellules et elle se manifeste par les oscillations de la quantité  $\Delta S/\Delta t$  visibles à la date  $t \approx 0,25$ .

La montée à l'ordre trois pour le cas test KT5 est bénéfique sur les maillages non structurés puisque le choix de  $\Theta_{b,ij}^{O3}$  assure la stabilité entropique et il permet d'observer les instabilités de Kelvin-Helmholtz.

Les mêmes remarques peuvent être formulées pour le problème KT6 dont les résultats sont présentés par la Figure 5.13.



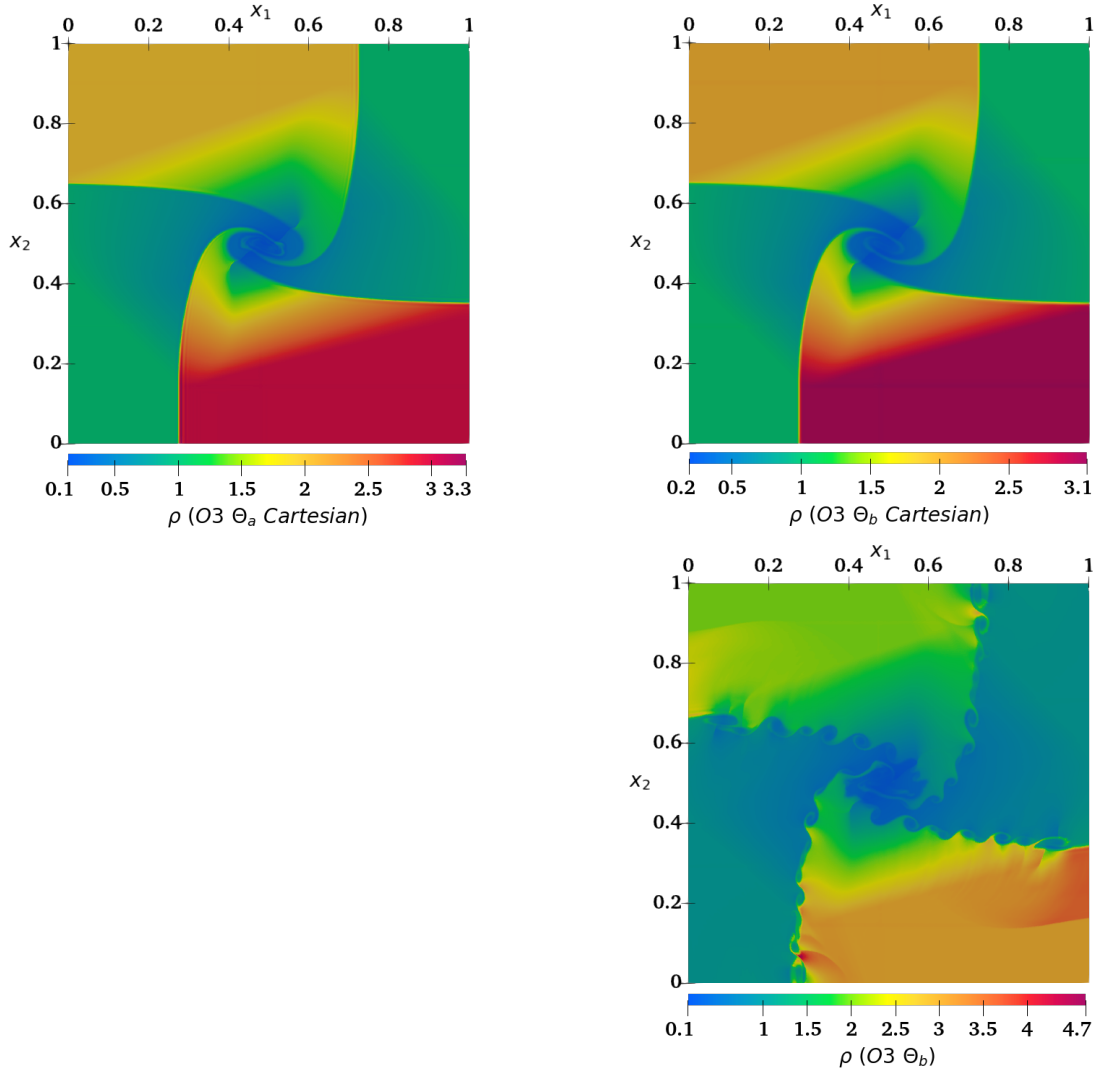


FIGURE 5.13 – Résultats numériques du schéma du troisième ordre (5.117) au temps final  $t = 0,3$  pour le problème de Riemann KT6 donné par le tableau 5.5. De gauche à droite, résultats pour  $(\Theta_{a,ij}^{O3})_{ij \in \mathcal{E}}$  définie par (5.126) puis pour  $(\Theta_{b,ij}^{O3})_{ij \in \mathcal{E}}$  définie par (5.127). De haut en bas, évolution temporelle de  $\Delta S / \Delta t$  défini par (5.129), résultats sur grille cartésienne uniforme de 1 000 000 cellules puis sur 1 002 422 triangles non structurés générés par le logiciel FreeFem++ [120].

## 5.5.4 Écoulement dans un tunnel avec marche

### 5.5.4.1 Définition du cas test

Cette section concerne l'écoulement d'un gaz évoluant à mach 3 dans un tunnel avec marche [224]. La géométrie du tunnel est incluse dans le rectangle  $[0, 3] \times [0, 1]$  mais à l'abscisse  $x_1 = 0,6$ , une marche de hauteur 0,2 réduit le profil longitudinal. Les conditions initiales sont  $\rho = 1,4$ ,  $(u; v)^T = (3; 0)^T$ ,  $p = 1$ . Une condition de Dirichlet avec les valeurs initiales est imposée sur le bord gauche tandis qu'une condition de Neumann homogène est imposée à droite. Sur toutes les autres frontières, le fluide vérifie une condition de réflexion. Le temps final de simulation est  $t = 4$ . Ce cas test est illustré par la Figure 5.14 qui présente les résultats d'un schéma usuel du premier ordre sur maillage triangulaire non structuré possédant 101 182 cellules.

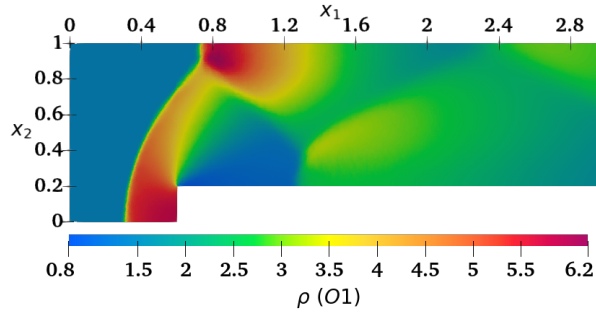


FIGURE 5.14 – Résultat numérique d'un schéma usuel du premier ordre pour le problème de la marche [224] au temps final  $t = 4$  sur 101 182 triangles non structurés générés par le logiciel FreeFem++ [120].

### 5.5.4.2 Résultats du schéma du second ordre

Cette section présente les résultats du schéma du second ordre pour le problème de la marche décrit ci-dessus. Les matrices  $(\Theta_{a,ij}^{O2})_{ij \in \mathcal{E}}$  données par (5.126) ne permettent pas de simuler ce problème et les résultats pour les matrices  $(\Theta_{b,ij}^{O2})_{ij \in \mathcal{E}}$  définies par (5.127) sont donnés par la Figure 5.15.

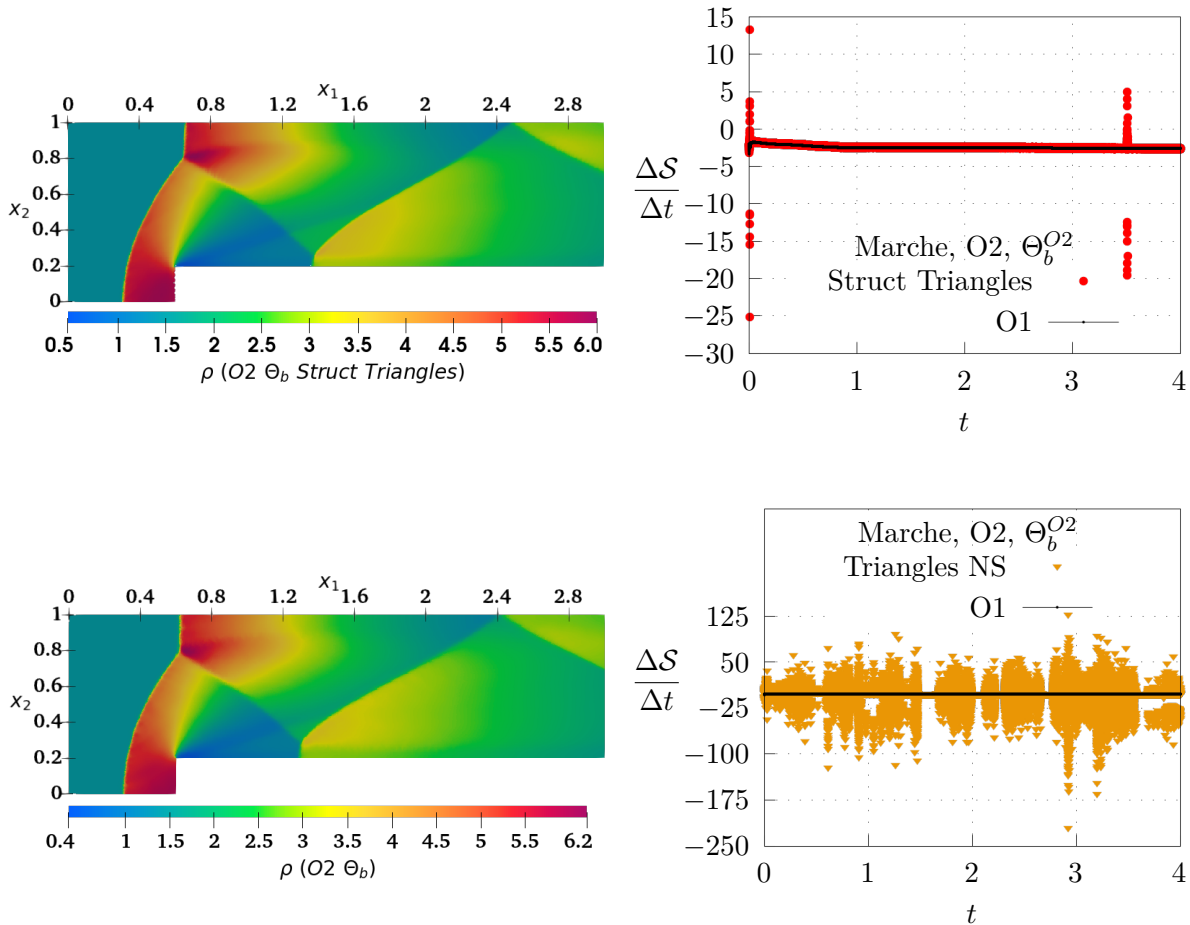


FIGURE 5.15 – À gauche, résultats numériques du schéma du second ordre (5.117) avec les matrices  $(\Theta_{b,ij}^{O2})_{ij \in \mathcal{E}}$  définies par (5.127) pour le problème de la marche [224] au temps final  $t = 4$  puis à droite, évolution temporelle de  $\Delta S / \Delta t$  défini par (5.129). De haut en bas, résultats numériques sur maillage constitué 100 330 triangles structurés puis sur 101 182 triangles non structurés générés par le logiciel FreeFem++ [120].

Graphiquement, les résultats sont semblables sur les deux types de maillages. En revanche, et bien qu'imparfaite, la stabilité entropique discrète globale est nettement mieux satisfaite sur le maillage



triangulaire structuré. La procédure MOOD concerne au plus 0.025% du total des cellules d'un tel maillage et elle est nécessaire pour au plus  $2 \cdot 10^{-3}\%$  du total des cellules triangulaires non structurés.

### 5.5.4.3 Résultats du schéma du troisième ordre

Cette section donne les résultats du schéma du troisième ordre (5.117) pour le problème de la marche introduit dans [224] et décrit ci-avant. Les matrices  $(\Theta_{a,ij}^{O3})_{ij \in \mathcal{E}}$  données par (5.126) ne permettent pas de simuler ce problème et les matrices  $(\Theta_{b,ij}^{O3})_{ij \in \mathcal{E}}$  définies par (5.127) sont fonctionnelles seulement sur les maillages triangulaires structurés. La Figure 5.16 montre les résultats obtenus.

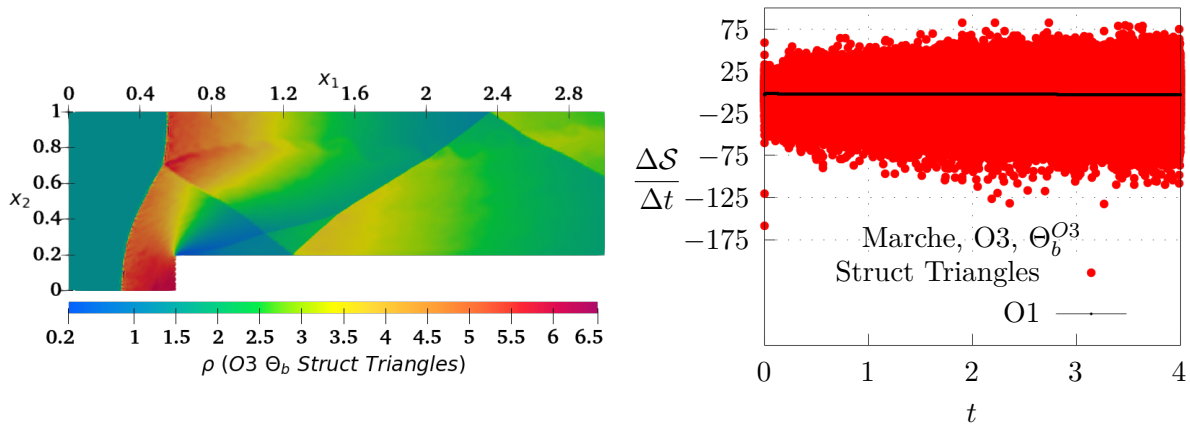


FIGURE 5.16 – À gauche, résultat numérique du schéma du troisième ordre (5.117) avec les matrices  $(\Theta_{b,ij}^{O3})_{ij \in \mathcal{E}}$  définies par (5.127), sur maillage constitué 100 330 triangles structurés et pour le problème de la marche [224] au temps final  $t = 4$ . À droite, évolution temporelle de  $\Delta S/\Delta t$  défini par (5.129).

Les résultats graphiques semblent précis avec une montée à l'ordre trois puisque des instabilités de Kelvin-Helmholtz sont légèrement visibles près de la droite  $x_2 = 0,8$ . L'onde verticale près de la droite  $x_1 = 0,5$  est néanmoins restituée à une position différente de celle observée jusqu'alors dans les Figures 5.14 et 5.15. Sa taille verticale est également plus grande. Malgré ces observations, la stabilité entropique discrète globale du schéma n'est clairement pas satisfaite pour ce cas test. La procédure MOOD s'active sur au plus 0.015% du total des cellules du maillage.

## 5.5.5 Double réflexion sur une rampe

### 5.5.5.1 Définition du cas test

Cette section est consacrée à la simulation d'une onde de choc évoluant le long d'une rampe inclinée telle que proposée dans [224]. La géométrie du domaine est incluse dans le rectangle  $[0; 1,9] \times [0; 1,3]$  mais à l'abscisse  $x_1 = 1/6$ , une pente formant un angle  $30^\circ$  avec l'horizontale connecte les bords du bas et de droite. Les conditions initiales sont  $\rho = 1,4$ ,  $(u; v)^T = (0; 0)^T$ ,  $p = 1$ . Sur le bord gauche, une condition de type Dirichlet est imposée avec les valeurs  $\rho = 8$ ,  $(u; v)^T = (8,25; 0)^T$  et  $p = 116,5$ , tandis que des conditions de Neumann homogènes sont imposées sur le bord droit du domaine. Sur toutes les autres frontières, des conditions de réflexions sont prescrites. Le temps final de simulation est  $t = 0,15$ . La Figure 5.17 illustre ce problème avec le résultat d'un schéma usuel du premier ordre sur maillage triangulaire non structuré possédant 1 003 193 cellules.

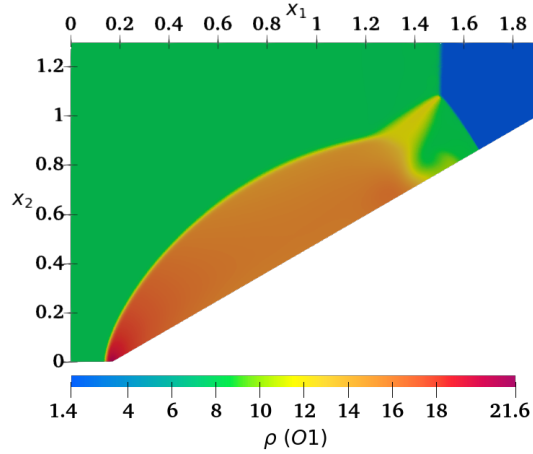


FIGURE 5.17 – Résultat numérique d’un schéma usuel du premier ordre pour le problème de la rampe [224] au temps final  $t = 0,15$  et sur 1 003 193 triangles non structurés générés par le logiciel FreeFem++ [120].

### 5.5.5.2 Résultats du schéma du second ordre

Cette section donne les résultats du schéma du second ordre (5.117) pour le problème de la rampe défini ci-dessus. Les matrices  $(\Theta_{a,ij}^{O2})_{ij \in \mathcal{E}}$  données par (5.126) ne permettent pas de simuler ce problème et les résultats pour les matrices  $(\Theta_{b,ij}^{O2})_{ij \in \mathcal{E}}$  définies par (5.127) sont donnés par la Figure 5.18.

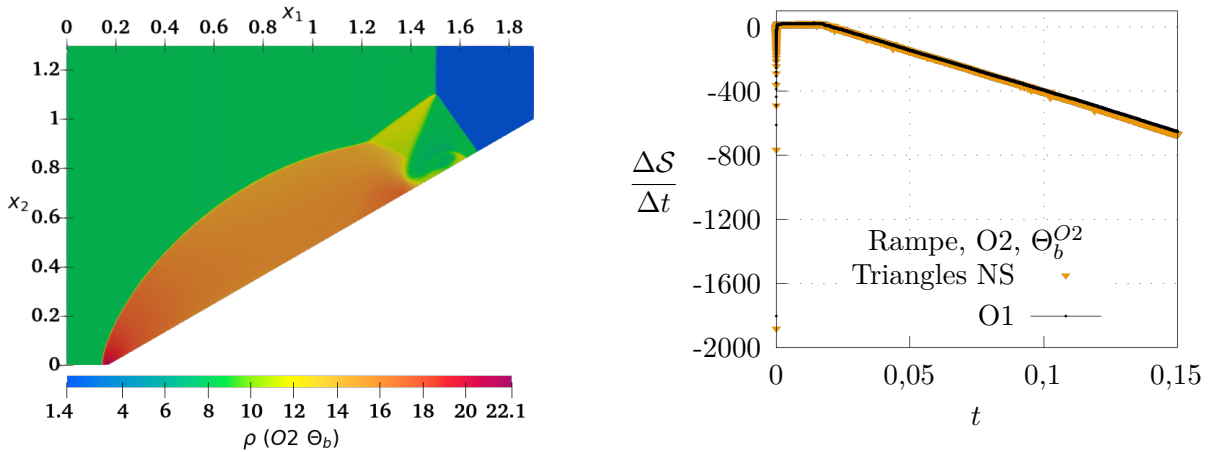


FIGURE 5.18 – À gauche, résultat numérique du schéma du second ordre (5.117) avec les matrices  $(\Theta_{b,ij}^{O2})_{ij \in \mathcal{E}}$  définies par (5.127), sur 1 003 193 triangles non structurés générés par le logiciel FreeFem++ [120] et pour le problème de la rampe [224] au temps final  $t = 0,15$ . À droite, évolution temporelle de  $\Delta S / \Delta t$  défini par (5.129).

La montée au second est ordre est bénéfique puisque les résultats sont moins diffusifs que ceux d’ordre un donnés par la Figure 5.17. Cependant, les instabilités de Kelvin-Helmholtz qui se développent le long de rampe après le passage de l’onde de choc (voir par exemple ([76, Figure 2.19], [60, Fig 14])) ne sont pas visibles sur la Figure 5.18. Au début de la simulation et approximativement jusqu’à la date  $t \approx 0,02$ , la quantité  $\Delta S / \Delta t$  défini par (5.129) est légèrement positive mais elle reste négative après cette date. Pour ce cas test, la procédure de limitation MOOD est appliquée au plus à  $10^{-4}\%$  du total des cellules.

### 5.5.5.3 Résultats du schéma du troisième ordre

Pour conclure, cette section s’intéresse aux résultats du schéma du troisième ordre (5.117) pour le problème de la rampe donné ci-avant. Les matrices  $(\Theta_{a,ij}^{O2})_{ij \in \mathcal{E}}$  définies par (5.126) ne sont pas

fonctionnelles pour ce problème et les résultats pour les matrices  $(\Theta_{b,ij}^{O2})_{ij \in \mathcal{E}}$  définies par (5.127) sont présentés par la Figure 5.19.

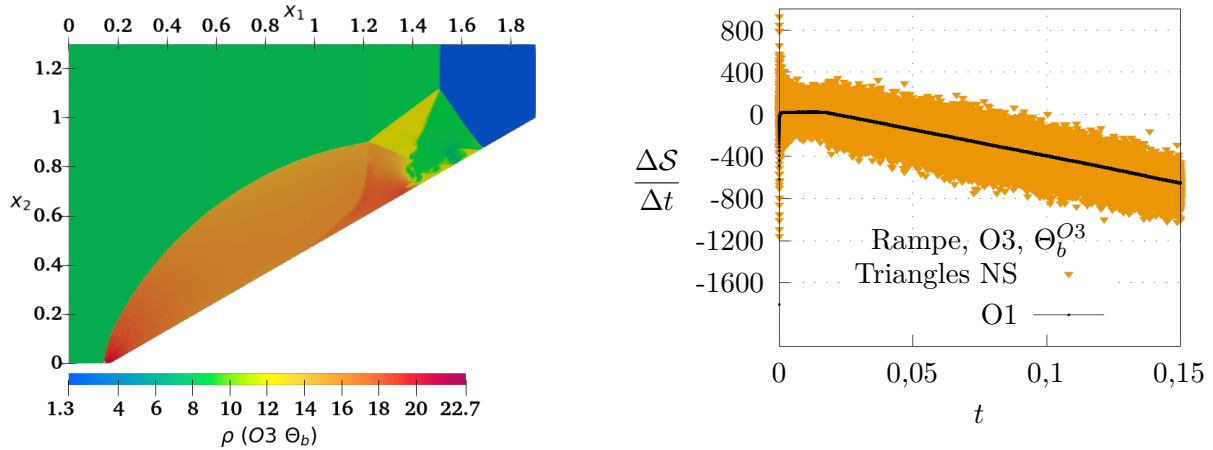


FIGURE 5.19 – À gauche, résultat numérique du schéma du troisième ordre (5.117) avec les matrices  $(\Theta_{b,ij}^{O3})_{ij \in \mathcal{E}}$  définies par (5.127), sur 1 003 193 triangles non structurés générés par le logiciel FreeFem++ [120] et pour le problème de la rampe [224] au temps final  $t = 0,15$ . À droite, évolution temporelle de  $\Delta S / \Delta t$  défini par (5.129).

La montée à l'ordre trois permet de restituer des instabilités de Kelvin-Helmholtz. La stabilité entropique du schéma est tendanciellement correcte mais elle demeure nettement moins satisfaisante que celle du second ordre. Cette difficulté se traduit également par un recours plus fréquent à la procédure MOOD puisque pour ce cas test, elle est appliquée au plus à 0,08% du total des cellules.

## 5.6 Synthèse et conclusion

Ce chapitre présente les extensions bi-dimensionnelles sur maillages non structurés des schémas numériques proposés dans le chapitre 4. Cette extension est réalisée en trois temps.

Tout d'abord, et en restreignant l'anisotropie locale des cellules du maillage, des approximations des matrices hessiennes et des gradients de  $w$  sont proposées. Les approximations des matrices hessiennes sont connues après résolution d'un système linéaire de taille quatre défini à chaque cellule du maillage. Les approximations obtenues sont consistantes au premier ordre, linéaires, locales et elles ne nécessitent pas de technique de limitation additionnelle. Ces matrices hessiennes discrètes permettent également de déduire des approximations d'ordre deux du gradient de  $w$ . Ces gradients discrets possèdent les mêmes propriétés que les hessiennes discrètes mais ils sont de plus totalement explicites et ils donnent, par dégradation immédiate, une approximation au premier ordre.

Ensuite, ces dérivées discrètes sont utilisées pour définir des schémas numériques explicites en temps, d'ordre deux et trois sur maillages non structurés. Les flux numériques associés aux schémas reposent sur une discrétisation d'ordre élevé des intégrales des fonctions de flux puis sur des corrections dissipatives du même ordre paramétrées par des matrices  $(\Theta_{ij}^{Ok})_{ij \in \mathcal{E}}$  dans  $\mathcal{M}_{m,m}(\mathbb{R})$ . Par conséquent, les schémas proposés n'utilisent pas de limiteur de pente et ils sont définis par la donnée d'un triplet  $(\Theta_{ij}^{Ok}, \lambda, \Delta t)$  dans  $\mathcal{M}_{m,m}(\mathbb{R}) \times (\mathbb{R}_*^+)^2$ .

Enfin, l'étude de stabilité montre l'existence d'un triplet  $(\Theta_{ij}^{Ok}, \lambda, \Delta t)$  tel que la séquence mise à jour  $(w_i^{n+1})_{i \in \mathcal{T}}$  vérifie une inégalité d'entropie discrète globale pour toute entropie strictement convexe de classe  $C^2(\Omega, \mathbb{R})$  donnée. Ce résultat d'existence est publié pour le second ordre dans [160].

Des expressions explicites des triplets  $(\Theta_{ij}^{Ok}, \lambda, \Delta t)$  sont respectivement données dans (5.92), (5.113), (5.124) mais il est clair que ces expressions sont malheureusement inutilisables en pratique. D'une manière générale, la mise en œuvre des schémas proposés n'est pas simple et par conséquent les paramètres  $(\Theta_{ij}^{Ok}, \lambda, \Delta t)$  sont choisis par une approche heuristique décrite en Section 5.5.1.

Pour les schémas du second et du troisième ordre, les résultats obtenus pour la simulation de solutions régulières sont satisfaisants. L'absence de limiteur de pente permet de retrouver exactement les ordres théoriques. Une très légère dégradation de l'ordre est toutefois observée pour le schéma du troisième ordre sur les maillages non structurés. L'origine de ce défaut est inconnue.

La simulation de solutions discontinues donne des résultats graphiques satisfaisants. La stabilité entropique du schéma du second ordre est globalement correcte avec l'une des deux heuristiques proposées. Néanmoins, des instabilités sont parfois observées sur maillages non structurés. La montée à l'ordre trois permet un gain de précision souligné par l'apparition d'instabilités de Kelvin-Helmholtz, mais les performances de stabilité du schéma sont décevantes avec les heuristiques proposées.

# Conclusion

L'idée directrice de ce mémoire de thèse est le développement de schémas numériques aux volumes finis à partir d'éléments qui permettent de retrouver une stabilité entropique discrète. Dans les chapitres 2 et 3, ces éléments s'expriment sous la forme de conditions suffisantes directement incluses dans les définitions des solveurs de Riemann. Dans les chapitres 4 et 5, ces éléments sont des paramètres additionnels laissés volontairement libres lors de la définition du schéma.

La démarche des chapitres 2 et 3, appliquée aux équations de Saint Venant, de Ripa et d'Euler, conduit à des équations non linéaires qui définissent des solveurs de Riemann à deux états intermédiaires constants. Sauf pour le cas des équations de Ripa, ces équations non linéaires sont quadratiques et leur discriminant est toujours positif. Dans le chapitre 2, ce signe est une conséquence directe du schéma HLL standard [118] et dans le chapitre 3, il est assuré par une condition implicite de type CFL et par une discrétisation spécifique du terme source. Par conséquent, les équations quadratiques proposées dans ces chapitres peuvent toujours être résolues explicitement dans  $\mathbb{R}$ . Une résolution explicite analogue n'a malheureusement pas pu être établie pour le cas des équations de Ripa dans lequel l'existence des solutions au système proposé est seulement donnée sous la forme d'une conjecture étayée par les résultats numériques.

Les schémas de type Godunov ainsi obtenus sont consistants, par construction, ils vérifient une inégalité d'entropie discrète locale et ils sont également *well-balanced* lorsque le système d'équations est équipé d'un terme source. Ces schémas sont de plus explicites sauf pour le cas des équations de Ripa qui nécessite la résolution d'un système d'équations algébriques à l'aide d'une méthode de Newton ou de point fixe par exemple.

Les résultats numériques sont globalement convenables et très satisfaisant pour les simulations de solutions discontinues. Toutefois, dans le cas des systèmes avec terme source, les schémas proposés peuvent converger vers des solutions qui ne sont pas celles de référence. Ces défauts de convergence ont déjà été observés avec d'autres schémas dans [34, 42, 59] et dans ce mémoire de thèse, ils se manifestent dans les simulations des équilibres mouvants atteints après régimes transitoires. Dans ces simulations, la propriété *well-balanced* n'est donc pas restituée à la précision machine attendue.

Par ailleurs, et en considérant un système hyperbolique muni d'une entropie suffisamment régulière, les chapitres 4 et 5 proposent d'obtenir une stabilité entropique discrète globale à partir d'une séquence de matrice formellement notée  $\Theta$ . Ces matrices sont libres lors de la définition des schémas puis contraintes en vertu de la stabilité souhaitée. Les contraintes sont explicitement données et il est établi que plusieurs matrices  $\Theta$  sont admissibles.

Cette stratégie permet de proposer des schémas d'ordre élevé, sans limiteurs pente et qui vérifient une inégalité d'entropie discrète globale d'abord en une dimension d'espace dans le chapitre 4 puis sur des maillages non structurés bi-dimensionnels dans le chapitre 5. Ces résultats de stabilité sont obtenus sous une condition CFL restrictive mais l'analyse du cas linéaire et les simulations numériques montrent que cette CFL peut être relaxée en utilisant une discrétisation temporelle d'ordre élevé de type *Runge-Kutta Strong Stability Preserving* [107, 109, 110]. D'après [76, 230], ces discrétisations temporelles préservent les résultats de stabilité entropique.

Puisque les schémas d'ordre élevé proposés n'utilisent pas de limiteur de pente, les ordres théoriques sont exactement retrouvés dans les simulations de solutions régulières. Une très légère dégradation de l'ordre est toutefois observée pour le schéma du troisième ordre sur les maillages non structurés. L'origine de ce défaut est inconnue.

Dans le cas de solutions discontinues, les simulations numériques réalisées montrent clairement que les matrices  $\Theta$  ne sont pas toutes équivalentes. En une dimension d'espace, certaines d'entre elles génèrent peu d'oscillations parasites et semblent assurer le principe du maximum dans le cas de l'équation scalaire de Burgers. Dans le cas bi-dimensionnel, le choix des matrices  $\Theta$  est plus délicat et puisque les expressions théoriques établies sont pratiquement réductrices, des choix heuristiques sont proposés. Si les résultats du schéma du second ordre sont corrects avec l'une des deux heuristiques, les performances du schéma du troisième ordre demeurent fragiles avec les choix de matrices  $\Theta$  proposés.

Ainsi, il est possible de définir des schémas numériques à partir des conditions usuelles de consistance complétées par des éléments liées à la stabilité entropique discrète. Les deux démarches présentées en ce sens donnent des résultats numériques variables mais elles permettent néanmoins de proposer plusieurs schémas qui vérifient une stabilité entropique discrète d'abord locale puis globale.

En particulier, l'ajout de conditions suffisantes de stabilité permet de proposer un schéma numérique pour les équations de Saint Venant, à stabilité entropique locale et *well-balanced* pour toutes les solutions stationnaires régulières. L'utilisation de paramètres libres permet par ailleurs d'établir des schémas numériques explicites en temps, du second et du troisième ordre, sans limiteur de pente, définis sur des maillages non structurés et qui vérifient une inégalité d'entropie discrète globale.

Les perspectives et les poursuites de ce mémoire de thèse sont diverses et quelques pistes sont énumérées ci-dessous.

La suite immédiate du chapitre 3 consisterait à démontrer la conjecture (3.117) qui admet l'existence des solutions d'un système d'équations algébriques non linéaires. La démonstration de cette conjecture permettrait de conclure quant à l'existence d'un schéma numérique entropique et *well-balanced* pour les équations de Ripa.

Ensuite, un prolongement possible des chapitres 4 et 5 serait d'identifier l'influence exacte des matrices  $\Theta$  sur les oscillations parasites observées. Cette identification pourrait être conduite en étudiant les éventuelles propriétés de stabilité entropique locale des schémas proposés. Ces recherches seraient motivées car la stabilité entropique discrète globale n'exclue pas la stabilité entropique locale et certaines matrices  $\Theta$  donnent des résultats numériques qui pourraient corroborer une telle stabilité dans le cas de l'équation scalaire de Burgers. Ces investigations pourraient être débutées en s'appuyant sur les récents résultats de N. Aguillon *et al* qui établissent dans [10] des tests numériques qui permettent de détecter les schémas non entropiques.

Enfin, il conviendrait de confronter les gradients et les matrices hessiennes discrets proposés en Section 5.2 aux méthodes présentées dans la bibliographie dédiée en Section 5.1.3. D'après [66, 201], plusieurs critères généraux de comparaisons pourraient être retenus et il pourrait également être intéressant de coupler ces dérivées discrètes avec d'autres schémas numériques existants.

## Annexe A

# Solutions stationnaires régulières des équations de Saint Venant

Cette annexe concerne les solutions stationnaires régulières des équations de Saint Venant. Les équations algébriques et une propriété vérifiées par ces solutions stationnaires sont établies ci-dessous.

**Lemme A.1** (Solutions stationnaires régulières des équations de Saint Venant). *Soit  $I$  un intervalle fermé et borné dans  $\mathbb{R}$ . Pour une fonction régulière indépendante du temps  $z : \mathbb{R} \rightarrow \mathbb{R}$  donnée, considérons le système d'équations de Saint Venant en une dimension d'espace donné par*

$$\partial_t \begin{pmatrix} h \\ q \end{pmatrix} + \partial_x \begin{pmatrix} q \\ \frac{q^2}{h} + \frac{gh^2}{2} \end{pmatrix} = \begin{pmatrix} 0 \\ -gh \end{pmatrix} \partial_x z, \quad \forall (x, t) \in I \times ]0, +\infty[. \quad (\text{A.1})$$

La fonction  $(x, t) \mapsto (h(x, t), q(x, t))^T$  prend ses valeurs dans l'espace convexe  $\Omega$  défini par

$$\Omega = \left\{ (h, q) \in \mathbb{R}^2 \mid h > 0, q \in \mathbb{R} \right\}.$$

Soit  $x \mapsto (h(x), q(x))$  une solution stationnaire régulière du système de Saint Venant (A.1). Les assertions suivantes sont vérifiées.

i) Il existe  $(q_0, B_0)$  dans  $\mathbb{R} \times \mathbb{R}_+^*$  tel que pour tout  $x$  dans  $I$

$$q(x) = q_0, \quad (\text{A.2a})$$

$$\frac{q_0}{2h^2(x)} + g(h(x) + z(x)) = B_0. \quad (\text{A.2b})$$

ii) En considérant  $q_0$  dans  $\mathbb{R}$  défini ci-dessus ainsi que l'ensemble  $J_{q_0}$  inclus dans  $I$  et tel que

$$J_{q_0} = \left\{ x \in I \mid 1 - \frac{q_0^2}{gh^3(x)} \neq 0 \right\}, \quad (\text{A.3})$$

la dérivée de  $h$  vérifie

$$\partial_x h = -\frac{\partial_x z}{1 - \frac{q_0^2}{gh^3(x)}}, \quad \forall x \in J_{q_0}. \quad (\text{A.4})$$

*Démonstration.* Par définition,  $x \mapsto h(x)$  et  $x \mapsto q(x)$  satisfont les deux équations suivantes pour tout  $x$  dans  $I$  :

$$\partial_x q = 0, \quad (\text{A.5a})$$

$$\partial_x \left( \frac{q^2}{h} + \frac{gh^2}{2} \right) = -gh \partial_x z. \quad (\text{A.5b})$$

Pour l'assertion *i*), et d'après l'équation (A.5a), le résultat (A.2a) est immédiat. En considérant ce résultat dans l'équation (A.5b) puis en appliquant les règles usuelles de dérivation, il suit

$$\begin{aligned}
0 &= \partial_x \left( \frac{q_0^2}{h} + \frac{gh^2}{2} \right) + gh \partial_x z, \\
&= -(\partial_x h) \frac{q_0^2}{h^2} + gh \partial_x (h + z), \\
&= h \left( -(\partial_x h) \frac{q_0^2}{h^3} \right) + gh \partial_x (h + z), \\
&= h \partial_x \left( \frac{q_0^2}{2h^2} \right) + gh \partial_x (h + z).
\end{aligned}$$

Puisque  $h > 0$ , il est possible de multiplier l'équation ci-dessus par  $1/h$  pour obtenir

$$\partial_x \left( \frac{q_0^2}{2h^2} + g(h + z) \right) = 0.$$

Cette dernière égalité permet de déduire le résultat (A.2b) et elle est conclut donc la preuve de l'assertion *i*).

En ce qui concerne l'assertion *ii*), et en utilisant les règles usuelles de dérivation, l'équation (A.5b) s'écrit également pour tout  $x$  dans  $I$  de la manière suivante :

$$\begin{aligned}
-gh \partial_x z &= \partial_x \left( \frac{q_0^2}{h} + \frac{gh^2}{2} \right), \\
&= -\partial_x h \frac{q_0^2}{h^2} + gh \partial_x h, \\
&= gh \partial_x h \left( 1 - \frac{q_0^2}{gh^3} \right).
\end{aligned}$$

En divisant l'équation ci-dessus par  $gh > 0$ , il suit

$$-\partial_x z = \partial_x h \left( 1 - \frac{q_0^2}{gh^3} \right). \tag{A.6}$$

En restreignant  $x$  à l'ensemble  $J_{q_0}$  défini par (A.3), il suit le résultat (A.4) ce qui conclut la preuve de l'assertion *ii*) ainsi que la preuve du lemme.  $\square$

D'après l'assertion *ii*) du lemme ci-dessus, en régime stationnaire et tant que  $x$  est dans l'ensemble  $J_{q_0}$ , les dérivées de  $h$  et de  $z$  sont proportionnelles. Ainsi, si  $z$  est localement constant alors  $h$  est aussi constant. De plus, le résultat (A.2a) montre que  $q$  est également constant en régimes stationnaires. Par conséquent, cette assertion permet de déduire que pour  $x$  dans  $J_{q_0}$ , si  $z$  est localement constante alors seules les constantes de  $\hat{\Omega}$  sont des solutions stationnaires régulières du système (A.1).

Lorsque  $x$  n'est pas dans  $J_{q_0}$ , l'équation (A.6) devient singulière. Dans ce cas, et comme  $z$  est une donnée, les équations stationnaires écrites sous la forme de (A.5) pourraient être mal posées.



# Bibliographie

- [1] H. Abdelwahed, M. Abdelrahman, A. Alsarhana, and K. K. Mohamed. Investigation of the Ripa model via NHRS scheme with its wide-ranging applications. *Fractal and Fractional*, 6 :745, 12 2022.
- [2] R. Abgrall. Residual distribution schemes : Current status and future trends. *Computers and Fluids*, 35(7) :641–669, 2006. Special Issue Dedicated to Professor Stanley G. Rubin on the Occasion of his 65th Birthday.
- [3] R. Abgrall. A general framework to construct schemes satisfying additional conservation relations. Application to entropy conservative and entropy dissipative schemes. *Journal of Computational Physics*, 372 :640–666, 2018.
- [4] R. Abgrall, P. Bacigaluppi, and S. Tokareva. High-order residual distribution scheme for the time-dependent Euler equations of fluid dynamics. *Computers and Mathematics with Applications*, 78(2) :274–297, 2019. Proceedings of the Eight International Conference on Numerical Methods for Multi-Material Fluid Flows (MULTIMAT 2017).
- [5] R. Abgrall, M. Lukacova-Medvidova, and P. Öffner. On the convergence of residual distribution schemes for the compressible Euler equations via dissipative weak solutions. *Mathematical Models and Methods in Applied Sciences*, 33, 12 2022.
- [6] R. Abgrall and M. Mezzine. Construction of second order accurate monotone and stable residual distribution schemes for unsteady flow problems. *Journal of Computational Physics*, 188(1) :16–55, 2003.
- [7] R. Abgrall, J. Nordström, P. Öffner, and S. Tokareva. Analysis of the SBP-SAT stabilization for finite element methods part I : Linear problems. *Journal of Scientific Computing*, 85 :43, 11 2020.
- [8] R. Abgrall, J. Nordström, P. Öffner, and S. Tokareva. Analysis of the SBP-SAT stabilization for finite element methods part II : Entropy stability. *Communications on Applied Mathematics and Computation*, 5, 06 2020.
- [9] R. Abgrall, P. Öffner, and H. Ranocha. Reinterpretation and extension of entropy correction terms for residual distribution and discontinuous Galerkin schemes : Application to structure preserving discretization. *Journal of Computational Physics*, 453 :110955, 2022.
- [10] N. Aguillon, E. Audusse, V. Desveaux, and J. Salomon. How to find a discrete entropy inequality when you don't know if it exists. working paper or preprint, Dec. 2022.
- [11] F. Alcrudo and F. Benkhaldoun. Exact solutions to the Riemann problem of the shallow water equations with a bottom step. *Computers and Fluids*, 30(6) :643–671, 2001.
- [12] A. Aleksyuk, M. Malakhov, and V. Belikov. The exact Riemann solver for the shallow water equations with a discontinuous bottom. *Journal of Computational Physics*, 450 :110801, 2022.

- [13] E. Audusse, F. Bouchut, M. Bristeau, R. Klein, and B. Perthame. A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows. *Siam Journal on Scientific Computing*, 25, 01 2004.
- [14] E. Audusse, F. Bouchut, M. Bristeau, and J. Sainte-Marie. Kinetic entropy inequality and hydrostatic reconstruction scheme for the Saint-Venant system. *Mathematics of Computation*, 85 :2815–2837, Nov. 2016.
- [15] E. Audusse, C. Chalons, and P. Ung. A simple well-balanced and positive numerical scheme for the shallow-water system. *Communications in Mathematical Sciences*, 13 :1317–1332, 2015.
- [16] M. Badsì, C. Berthon, and L. Martaud. A family of second-order dissipative finite volume schemes for hyperbolic systems of conservation laws. *The SMAI Journal of computational mathematics*, 9 :31–60, 01 2023.
- [17] D. S. Balsara, S. Garain, V. Florinski, and W. Boscheri. An efficient class of WENO schemes with adaptive order for unstructured meshes. *Journal of Computational Physics*, 404 :109062, 2020.
- [18] M. Ben-Artzi and J. Li. Consistency and convergence of finite volume approximations to nonlinear hyperbolic balance laws. *arXiv : Numerical Analysis*, 2019.
- [19] A. Bermudez and E. Vazquez. Upwind methods for hyperbolic conservation laws with source terms. *Computers and Fluids*, 23(8) :1049–1071, 1994.
- [20] R. Berneti, V. Titarev, and E. Toro. Exact solution of the Riemann problem for the shallow water equations with discontinuous bottom geometry. *Journal of Computational Physics*, 227(6) :3212–3243, 2008.
- [21] C. Berthon. Numerical approximations of the 10-moment gaussian closure. *Mathematics of Computation*, 75(256) :1809–1832, 2006.
- [22] C. Berthon. Robustness of MUSCL schemes for 2D unstructured meshes. *Journal of Computational Physics*, 218(2) :495–509, 2006.
- [23] C. Berthon and C. Chalons. A fully well-balanced, positive and entropy-satisfying Godunov-type method for the shallow-water equations. *Mathematics of Computation*, 85, 02 2015.
- [24] C. Berthon, B. Dubroca, and A. Sangam. An entropy preserving relaxation scheme for ten-moments equations with source terms. *Communications in Mathematical Sciences*, 13(8) :2119–2154, 2015.
- [25] C. Berthon, M. M’Baye, M. Le, and D. Seck. A well-defined moving steady states capturing Godunov-type scheme for shallow-water model. *Int. J. Finite Volumes*, 01 2021.
- [26] C. Berthon and V. Michel-Dansac. A simple fully well-balanced and entropy preserving scheme for the shallow-water equations. *Applied Mathematics Letters*, 86 :284–290, 2018.
- [27] C. Berthon, Y. Y. Coudière, and V. Desveaux. Second-order MUSCL schemes based on dual mesh gradient reconstruction (DMGR). *European Series in Applied and Industrial Mathematics (ESAIM) : Mathematical Modelling and Numerical Analysis*, 48, 03 2014.
- [28] L. Betchen and A. Straatman. An accurate gradient and hessian reconstruction method for cell centered finite volume discretizations on general unstructured grids. *International Journal for Numerical Methods in Fluids*, 62 :945 – 962, 01 2009.
- [29] F. Bianco, G. Puppo, and G. Russo. High-order central schemes for hyperbolic systems of conservation laws. In M. Fey and R. Jeltsch, editors, *Hyperbolic Problems : Theory, Numerics, Applications*, pages 55–64, Basel, 1999. Birkhäuser Basel.

- [30] A. Bihlo and S. MacLachlan. Well-balanced mesh-based and meshless schemes for the shallow-water equations, 2017.
- [31] F. Bouchut. Construction of BGK models with a family of kinetic entropies for a given system of conservation laws. *Journal of statistical physics*, 95(1-2) :113–170, 1999.
- [32] F. Bouchut. *Nonlinear stability of finite volume methods for hyperbolic conservation laws and well-balanced schemes for sources*. Frontiers in Mathematics. Birkhäuser Verlag, Basel, 2004.
- [33] F. Bouchut, C. Bourdarias, and B. Perthame. A MUSCL method satisfying all the numerical entropy inequalities. *Math. Comput.*, 65 :1439–1461, 10 1996.
- [34] F. Bouchut and T. M. de Luna. An entropy satisfying scheme for two-layer shallow water equations with uncoupled treatment. *ESAIM : Mathematical Modelling and Numerical Analysis*, 42 :683 – 698, 07 2008.
- [35] F. Bouchut and T. M. de Luna. A subsonic-well-balanced reconstruction scheme for shallow water flows. *SIAM J. Numerical Analysis*, 48 :1733–1758, 01 2010.
- [36] Y. Brenier and S. Osher. Approximate Riemann solvers and numerical flux functions. *SIAM Journal on Numerical Analysis*, 23, 01 1985.
- [37] J. Brezina, E. Chiodaroli, and O. Kreml. On contact discontinuities in multi-dimensional isentropic Euler equations, 2017.
- [38] J. Britton and Y. Xing. High order still-water and moving-water equilibria preserving discontinuous Galerkin methods for the Ripa model. *Journal of Scientific Computing*, 82 :1–37, 2020.
- [39] S. E. Buckley and M. C. Leverett. Mechanism of fluid displacement in sands. *Transactions of the AIIME*, 146 :107–116, 1942.
- [40] S. Busto, M. Dumbser, I. Peshkov, and E. Romenski. On thermodynamically compatible finite volume schemes for continuum mechanics. *SIAM Journal on Scientific Computing*, 44(3) :A1723–A1751, 2022.
- [41] J. Butcher. The numerical analysis of ordinary differential equations : Runge-Kutta and general linear methods. *Mathematics of Computation*, 51 :377, 1987.
- [42] C. C. Madroñal and E. Pimentel-García. The Riemann problem for the shallow water equations with discontinuous topography : The wet-dry case. *Journal of Computational Physics*, 378, 11 2018.
- [43] D. Caraeni and D. Hill. Unstructured-grid third order finite volume discretization using a multistep quadratic data reconstruction method. *AIAA Journal*, 48, 04 2010.
- [44] M. Carpenter and T. Fisher. High-order entropy stable formulations for computational fluid dynamics. 06 2013.
- [45] M. Carpenter, T. Fisher, and S. Frankel. Entropy stable spectral collocation schemes for the Navier-Stokes equations : discontinuous interfaces. *SIAM Journal on Scientific Computing*, 36 :B835–B867, 01 2014.
- [46] C. Castro and E. Toro. Solvers for the high-order Riemann problem for hyperbolic balance laws. *Journal of Computational Physics*, 227 :2482–2513, 2008.
- [47] M. J. Castro, U. S. Fjordholm, S. Mishra, and C. Parés. Entropy conservative and entropy stable schemes for nonconservative hyperbolic systems. *SIAM Journal on Numerical Analysis*, 51(3) :1371–1391, 2013.

- [48] M. Castro Díaz, J. López-García, and C. Parés. High-order exactly well-balanced numerical methods for shallow water systems. *Journal of Computational Physics*, 246 :242–264, 2013.
- [49] C. Chalons and P. G. LeFloch. A fully discrete scheme for diffusive-dispersive conservation laws. *Numerische Mathematik*, 89(3) :493–509, 2001.
- [50] J. Chan. On discretely entropy conservative and entropy stable discontinuous Galerkin methods. *Journal of Computational Physics*, 362, 08 2017.
- [51] P. Chandrashekar. Kinetic energy preserving and entropy stable finite volume schemes for compressible Euler and Navier-Stokes equations. *Communications in Computational Physics*, 14, 09 2012.
- [52] P. Chandrashekar and C. Klingenberg. Entropy stable finite volume scheme for ideal compressible MHD on 2-D cartesian meshes. *SIAM Journal on Numerical Analysis*, 54 :1313–1340, 01 2016.
- [53] J. Charney, R. Fjortoft, and J. Neumann. Numerical integration of the barotropic vorticity equation. *Tellus A*, 2, 11 1950.
- [54] G. Chen and S. Noelle. A new hydrostatic reconstruction scheme based on subcell reconstructions. *SIAM Journal on Numerical Analysis*, 55(2) :758–784, 2017.
- [55] T. Chen and C. Shu. Entropy stable high-order discontinuous Galerkin methods with suitable quadrature rules for hyperbolic conservation laws. *Journal of Computational Physics*, 345 :427–461, 2017.
- [56] T. Chen and C. Shu. Review of entropy stable discontinuous Galerkin methods for systems of conservation laws on unstructured simplex meshes. *CSIAM Transactions on Applied Mathematics*, 1(1) :1–52, 2020.
- [57] X. Cheng. A fourth order entropy stable scheme for hyperbolic conservation laws. *Entropy (Basel, Switzerland)*, 21(5), May 2019.
- [58] A. Chertock, A. Kurganov, and Y. Liu. Central-upwind schemes for the system of shallow water equations with horizontal temperature gradients. *Numerische Mathematik*, 127, 08 2014.
- [59] A. Chinnayya, A. Leroux, and N. Seguin. A well-balanced numerical scheme for the approximation of the shallow-water equations with topography : The resonance phenomenon. *International Journal on Finite Volumes*, 1, 01 2004.
- [60] S. Clain, S. Diot, and R. Loubère. A high-order finite volume method for hyperbolic systems : Multi-dimensional optimal order detection (MOOD). *Journal of Computational Physics, Elsevier*, 2011.
- [61] B. Cockburn, F. Coquel, and P. G. LeFloch. Convergence of the finite volume method for multidimensional conservation laws. *SIAM Journal on Numerical Analysis*, 32(3) :687–705, 1995.
- [62] C. Contarino, Q. Zhang, E. Toro, and G. Montecinos. Towards high-order methods for blood flow. 06 2015.
- [63] F. Coquel, P. Helluy, and J. Schneider. Second-order entropy diminishing scheme for the Euler equations. *International Journal for Numerical Methods in Fluids*, 50 :1029–1061, 2006.
- [64] F. Coquel, J.-M. Hérard, and K. Saleh. A splitting method for the isentropic baer-nunziato two-phase flow model. In *ESAIM : Proceedings*, volume 38, pages 241–256. EDP Sciences, 2012.
- [65] F. Coquel and P. LeFloch. An entropy satisfying MUSCL scheme for systems of conservation laws. *Numerische Mathematik*, 74 :1–33, 07 1996.

- [66] C. Correa, R. Hero, and K. Ma. A comparison of gradient estimation methods for volume rendering on unstructured meshes. *IEEE transactions on visualization and computer graphics*, 17 :305–19, 03 2011.
- [67] Y. Coudière and G. Manzini. The discrete duality finite volume method for convection diffusion problems. *SIAM Journal on Numerical Analysis*, 47(6) :4163–4192, 2010.
- [68] L. Cueto-Felgueroso, I. Colominas, J. Fe, F. Navarrina, and M. Casteleiro. High-order finite volume schemes on unstructured grids using moving least-squares reconstruction. application to shallow water dynamics. *International Journal for Numerical Methods in Engineering*, 65(3) :295–331, 2006.
- [69] H. Deconinck and M. Ricchiuto. *Residual Distribution Schemes : Foundations and Analysis*. 10 2007.
- [70] M. Deka, S. Brahmachary, R. Thirumalaisamy, A. Dalal, and G. Natarajan. A new green-gauss reconstruction on unstructured meshes. part i : gradient reconstruction. *Journal of Computational Physics*, 422, 10 2018.
- [71] O. Delestre, C. Lucas, P. Ksinant, F. Darboux, C. Laguerre, T. Vo, F. James, and S. Cordier. SWASHES : a compilation of Shallow Water Analytic Solutions for Hydraulic and Environmental Studies. *International Journal for Numerical Methods in Fluids*, 72(3) :269–300, May 2013.
- [72] S. Dellacherie. Analysis of Godunov type schemes applied to the compressible euler system at low mach number. *Journal of Computational Physics*, 229(4) :978–1016, 2010.
- [73] E. Deriaz. Stability conditions for the numerical solution of convection-dominated problems with skew-symmetric discretizations. *SIAM Journal on Numerical Analysis*, 50(3) :1058–1085, 2012.
- [74] E. Deriaz and P. Haldenwang. Non-linear CFL conditions issued from the von Neumann stability analysis for the transport equation. *Journal of Scientific Computing*, 85(1) :5, 2020.
- [75] B. Després. Quadratic stability of flux limiters. working paper or preprint, July 2021.
- [76] V. Desveaux. *Contribution à l’approximation numérique des systèmes hyperboliques*. Theses, Université de Nantes, Nov. 2013.
- [77] V. Desveaux and C. Berthon. An entropy preserving MOOD scheme for the Euler equations. *International Journal On Finite Volumes*, 11 :1–39, 2014.
- [78] V. Desveaux, M. Zenk, C. Berthon, and C. Klingenberg. Well-balanced schemes to capture non-explicit steady states. Ripa model. *Mathematics of Computation*, 85 :1, 10 2015.
- [79] S. Diot, S. Clain, and R. Loubère. Improved detection criteria for the multi-dimensional optimal order detection (MOOD) on unstructured meshes with very high-order polynomials. *Computers and Fluids*, 64 :43–63, 2012.
- [80] B. Diskin and J. Thomas. Accuracy of gradient reconstruction on grids with high aspect ratio. *National Institute of Aerospace, Report No. 2008-12*, 01 2009.
- [81] K. Domelevo and P. Omnes. A finite volume method for the laplace equation on almost arbitrary two-dimensional grids. *ESAIM : M2AN*, 39 :1203–1249, 11 2005.
- [82] J. Dong and X. Qian. Well-balanced and positivity-preserving surface reconstruction schemes solving Ripa systems with nonflat bottom topography. *SIAM Journal on Scientific Computing*, 44 :A3098–A3129, 09 2022.
- [83] M. Dumbser, D. S. Balsara, E. Toro, and C. D. Munz. A unified framework for the construction of one-step finite volume and discontinuous Galerkin schemes on unstructured meshes. *Journal of Computational Physics*, 227(18) :8209 – 8253, 2008.

- [84] M. Dumbser, C. Enaux, and E. Toro. Finite volume schemes of very high-order of accuracy for stiff hyperbolic balance laws. *Journal of Computational Physics*, 227 :3971–4001, 2008.
- [85] A. Duran and F. Marche. Recent advances on the discontinuous Galerkin method for shallow water equations with topography source terms. *Comput. & Fluids*, 101 :88–104, 2014.
- [86] A. Duran, J. Vila, and R. Baraille. Semi-implicit staggered mesh scheme for the multi-layer shallow water system. *C. R. Acad. Sci. Paris.*, 2017.
- [87] B. Engquist and S. Osher. One-sided difference approximations for nonlinear conservation laws. *Journal of Computational Physics*, 36 :321–351, 1981. NR 20140805.
- [88] E. Tadmor. A minimum entropy principle in the gas dynamics equations. *Appl. Numer. Math*, pages 211–219, 1986.
- [89] R. Eymard, T. Gallouët, and R. Herbin. A cell-centred finite volume approximation for second order partial derivative operators with full matrix on unstructured meshes in any space dimension. *IMA Journal of Numerical Analysis*, 26 :326–353, 2006.
- [90] P. S. Farmakis, P. Tsoutsanis, and X. Nogueira. WENO schemes on unstructured meshes using a relaxed a posteriori MOOD limiting approach. *Computer Methods in Applied Mechanics and Engineering*, 363 :112921, 2020.
- [91] D. C. D. R. Fernández, J. Crean, M. Carpenter, and J. Hicken. Staggered-grid entropy-stable multidimensional summation-by-parts discretizations on curvilinear coordinates. *Journal of Computational Physics*, 392 :161–186, 2019.
- [92] T. Fisher and M. Carpenter. High-order entropy stable finite difference schemes for nonlinear conservation laws : Finite domains. *Journal of Computational Physics*, 252 :518–557, 2013.
- [93] U. Fjordholm, S. Mishra, and E. Tadmor. Well-balanced and energy stable schemes for the shallow water equations with discontinuous topography. *Journal of Computational Physics*, 230(14) :5587–5609, 2011.
- [94] U. Fjordholm, S. Mishra, and E. Tadmor. Arbitrarily high-order accurate entropy stable essentially nonoscillatory schemes for systems of conservation laws. *SIAM Journal on Numerical Analysis*, 50 :544–573, 03 2012.
- [95] E. Franck and L. Navoret. Semi-implicit two-speed well-balanced relaxation scheme for Ripa model. In *FVCA 2020, Finite Volumes for Complex Applications IX - Methods, Theoretical Aspects, Examples*, pages pp 735–743. Springer, Jun 2020.
- [96] L. Friedrich, G. Schnucke, A. Winters, D. D. R. Fernandez, G. Gassner, and M. Carpenter. Entropy stable space-time discontinuous Galerkin schemes with summation-by-parts property for hyperbolic conservation laws. *arXiv e-prints*, page arXiv :1808.08218, aug 2018.
- [97] J. Fürst and K. Kozel. *Second and Third Order Weighted ENO Scheme on Unstructured Meshes*, pages 737–744. 01 2003.
- [98] T. Gallouët, J. Hérard, and N. Seguin. Some approximate Godunov schemes to compute shallow-water equations with topography. *Computers and Fluids*, 32(4) :479–513, 2003.
- [99] T. Gallouët, R. Herbin, J. C. Latché, and N. Therme. Consistent internal energy based schemes for the compressible Euler equations. *Numerical Simulation in Physics and Engineering : Trends and Applications : Lecture Notes of the XVIII ‘Jacques-Louis Lions’ Spanish-French School*, pages 119–154, 2021.

- [100] G. Gassner, A. Winters, and D. Kopriva. A well balanced and entropy conservative discontinuous Galerkin spectral element method for the shallow water equations. *Applied Mathematics and Computation*, 272 :291–308, 2016. Recent Advances in Numerical Methods for Hyperbolic Partial Differential Equations.
- [101] L. Gastaldo, R. Herbin, J.-C. Latché, and N. Therme. A MUSCL-type segregated–explicit staggered scheme for the Euler equations. *Computers & Fluids*, 175 :91–110, 2018.
- [102] W. Gautschi. *Numerical Analysis*. SpringerLink : Bücher. Birkhäuser Boston, 2011.
- [103] D. George. Augmented Riemann solvers for the shallow water equations over variable topography with steady states and inundation. *Journal of Computational Physics*, 227(6) :3089–3113, 2008.
- [104] E. Godlewski and P.-A. Raviart. *Hyperbolic systems of conservation laws*, volume 3/4 of *Mathématiques & Applications (Paris) [Mathematics and Applications]*. Ellipses, Paris, 1991.
- [105] E. Godlewski and P.-A. Raviart. *Numerical approximation of hyperbolic systems of conservation laws*, volume 118 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1996.
- [106] S. K. Godunov. A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Matematičeskij sbornik*, 47(89) :271–306, 1959.
- [107] S. Gottlieb. On high order strong stability preserving Runge-Kutta and multi step time discretizations. *Journal of scientific computing*, 25(1) :105–128, 2005.
- [108] S. Gottlieb and D. Ketcheson. Chapter 21 - time discretization techniques. In R. Abgrall and C. Shu, editors, *Handbook of Numerical Methods for Hyperbolic Problems*, volume 17 of *Handbook of Numerical Analysis*, pages 549–583. Elsevier, 2016.
- [109] S. Gottlieb and C.-W. Shu. Total variation diminishing Runge-Kutta schemes. *Mathematics of Computation*, 67(221) :73–85, 1998.
- [110] S. Gottlieb, C.-W. Shu, and E. Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM review*, 43(1) :89–112, 2001.
- [111] N. Goutal and F. Maurel. *Proceedings of the 2nd workshop on Dam-Break Wave Simulation*. Department Laboratoire National d’Hydraulique, Groupe Hydraulique Fluviale, 1997.
- [112] J. Greenberg and A. Leroux. A well-balanced scheme for the numerical processing of source terms in hyperbolic equations. *SIAM Journal on Numerical Analysis*, 33(1) :1–16, 1996.
- [113] D. Hally. Calculation of the moments of polygons. Technical Report ADA183444, Defence Research Establishment Suffield Ralston (Alberta), Canadian National Defense, 06 1987. Technical memo.
- [114] E. Han and G. Warnecke. Exact Riemann solutions to shallow water equations. *Quarterly of Applied Mathematics*, 72, 07 2014.
- [115] X. Han and G. Li. Well-balanced finite difference WENO schemes for the Ripa model. *Computers and Fluids*, 134-135 :1–10, 2016.
- [116] A. Harten. High resolution schemes for hyperbolic conservation laws. *Journal of Computational Physics*, 135(2) :260 – 278, 1982.
- [117] A. Harten, B. Engquist, S. Osher, and S. R. Chakravarthy. Uniformly high-order accurate essentially non-oscillatory schemes, III. *Journal of Computational Physics*, 71(2) :231–303, 1987.
- [118] A. Harten, P. Lax, and B. Van Leer. On upstream differencing and Godunov-type schemes for hyperbolic conservation laws. *SIAM review*, 25 :35–61, 1983.

- [119] A. Harten and P. D. Lax. A random choice finite difference scheme for hyperbolic conservation laws. *SIAM Journal on Numerical Analysis*, 18(2) :289–315, 1981.
- [120] F. Hecht. New development in FreeFem++. *J. Numer. Math.*, 20(3-4) :251–265, 2012.
- [121] R. Herbin and F. Hubert. Benchmark on discretization schemes for anisotropic diffusion problems on general grids. In ISTE, editor, *Finite volumes for complex applications V*, pages 659–692, France, June 2008. Wiley.
- [122] R. Herbin, W. Kheriji, and J. Latché. On some implicit and semi-implicit staggered schemes for the shallow water and Euler equations. *Mathematical Modelling and Numerical Analysis*, 48 :1807–1857, 2014.
- [123] R. Herbin, J. Latché, and T. Nguyen. Consistent explicit staggered schemes for compressible flows part I : the barotropic Euler equations. 2013.
- [124] G. Hernandez-Duenas. A hybrid method to solve shallow water flows with horizontal density gradients. *Journal of Scientific Computing*, 73, 12 2017.
- [125] A. Hildebrand and S. Mishra. Entropy stable shock capturing space-time discontinuous Galerkin schemes for systems of conservation laws. *Numerische Mathematik*, 126(1) :103–151, 2014.
- [126] A. Hildebrand and S. Mishra. Entropy stability and well-balancedness of space-time DG for the shallow water equations with bottom topography. *Networks Heterog. Media*, 11 :145–162, 2016.
- [127] A. Hildebrand, S. Mishra, and C. Parés. Entropy-stable space-time DG schemes for non-conservative hyperbolic systems. *ESAIM : Mathematical Modelling and Numerical Analysis*, 52(3) :995–1022, 2018.
- [128] N. Hiroaki and J. White. An efficient cell-centered finite-volume method with face-averaged nodal-gradients for triangular grids. *Journal of Computational Physics*, 411 :109423, 2020.
- [129] J. Hou, Q. Liang, F. Simons, and R. Hinkelmann. A stable 2D unstructured shallow flow model for simulations of wetting and drying over rough terrains. *Computers and Fluids*, 82 :132–147, 2013.
- [130] C. Hu and C.-W. Shu. Weighted essentially non-oscillatory schemes on triangular meshes. *Journal of Computational Physics*, 150(1) :97–127, 1999.
- [131] W. Huang, R. Li, J. Qiu, and M. Zhang. A well-balanced moving mesh discontinuous Galerkin method for the Ripa model on triangular meshes. 05 2022.
- [132] F. Ismail and P. L. Roe. Affordable, entropy-consistent Euler flux functions II : Entropy production at shocks. *Journal of Computational Physics*, 228(15) :5410–5436, 2009.
- [133] D. I. Ketcheson. Relaxation Runge-Kutta methods : Conservation and stability for inner-product norms. *SIAM J. Numer. Anal.*, 57 :2850–2870, 2019.
- [134] B. Khobalatte and B. Perthame. Maximum principle on the entropy and second-order kinetic schemes. *Math. Comp.*, 62(205) :119–131, 1994.
- [135] J. Koellermeier and E. Pimentel-García. Steady states and well-balanced schemes for shallow water moment equations with topography. *Applied Mathematics and Computation*, 427 :127166, 2022.
- [136] A. Kurganov and E. Tadmor. Solution of two-dimensional Riemann problems for gas dynamics without Riemann problem solvers. *Numer. Methods Partial Differential Equations*, pages 548–608, 2002.



- [137] D. Kuzmin, H. Hajduk, and A. Rupp. Limiter-based entropy stabilization of semi-discrete and fully discrete schemes for nonlinear hyperbolic problems, 2021.
- [138] P. Lax. Shock waves and entropy. In *Contributions to nonlinear functional analysis (Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1971)*, pages 603–634. Academic Press, New York, 1971.
- [139] P. Lax. *Hyperbolic systems of conservation laws and the mathematical theory of shock waves*. Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1973. Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 11.
- [140] P. Lax and B. Wendroff. Systems of conservation laws. *Comm. Pure Appl. Math.*, 13 :217–237, 1960.
- [141] P. D. Lax. Weak solutions of nonlinear hyperbolic equations and their numerical computation. *Communications on Pure and Applied Mathematics*, 7 :159–193, 1954.
- [142] P. D. Lax and R. D. Richtmyer. Survey of the stability of linear finite difference equations. *Communications on Pure and Applied Mathematics*, 9 :267–293, 1956.
- [143] P. LeFloch. *Hyperbolic systems of conservation laws*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2002. The theory of classical and nonclassical shock waves.
- [144] P. LeFloch, J. Mercier, and C. Rohde. Fully discrete, entropy conservative schemes of arbitrary order. *SIAM J. Numerical Analysis*, 40 :1968–1992, 10 2002.
- [145] P. LeFloch and M. Thanh. The Riemann problem for the shallow water equations with discontinuous topography. *Communications in Mathematical Sciences*, 5(4) :865 – 885, 2007.
- [146] P. LeFloch and M. Thanh. A Godunov-type method for the shallow water equations with discontinuous topography in the resonant regime. *J. Comput. Physics.*, 230 :7631–7660, 2011.
- [147] P. G. LeFloch and H. Ranocha. Kinetic functions for nonclassical shocks, entropy stability, and discrete summation by parts. *Journal of Scientific Computing*, 87(2), apr 2021.
- [148] R. LeVeque. *Numerical methods for conservation laws (2. ed.)*. Lectures in mathematics. Birkhäuser, 1992.
- [149] R. LeVeque. *Finite-Volume Methods for Hyperbolic Problems*. Cambridge University Press, 2002.
- [150] J. Li, G. Li, S. Qian, J. Gao, and Q. Niu. A high-order well-balanced discontinuous Galerkin method based on the hydrostatic reconstruction for the Ripa model. *Advances in Applied Mathematics and Mechanics*, 12(6) :1416–1437, 2020.
- [151] M. Li, R. Mu, and H. Dong. A well-balanced discontinuous Galerkin method for the shallow water flows on erodible bottomimage 1. *Computers and Mathematics with Applications*, 119 :13–20, 2022.
- [152] Q. Liang and F. Marche. Numerical resolution of well-balanced shallow water equations with complex source terms. *Advances in Water Resources*, 32(6) :873–884, 2009.
- [153] M. J. Lighthill and G. B. Whitham. On kinematic waves ii. a theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 229 :317 – 345, 1955.
- [154] G. Lissoni. *DDFV method : applications to fluid mechanics and domain decomposition*. Theses, COMUE Université Côte d’Azur (2015 - 2019), Oct. 2019.

- [155] H. Liu and X. Jiao. WLS-ENO : Weighted-least-squares based essentially non-oscillatory schemes for finite volume methods on unstructured meshes. *Journal of Computational Physics*, 314 :749–773, 2016.
- [156] J. G. Liu and C. W. Shu. A high-order discontinuous Galerkin method for 2D incompressible flows. *Journal of Computational Physics*, 160(2) :577 – 596, 2000.
- [157] Y. Liu and Y.-T. Zhang. A robust reconstruction for unstructured WENO schemes. *Journal of Scientific Computing*, 54, 02 2013.
- [158] R. Loubère, M. Dumbser, and S. Michael. A new family of high order unstructured MOOD and ADER finite volume schemes for multidimensional systems of hyperbolic conservation laws. *Communications in Computational Physics*, 16, 09 2013.
- [159] C. Lozano. Entropy production by explicit Runge–Kutta schemes. *Journal of Scientific Computing*, 76 :521–564, 07 2018.
- [160] L. Martaud. Global entropy stability for a class of unlimited second-order schemes for 2D hyperbolic systems of conservation laws on unstructured meshes. *Journal of Computational Physics*, 487 :112176, 2023.
- [161] L. Martaud and C. Berthon. The existence of well-balanced entropy stable numerical scheme for the ripa model with the topography source term. In *Proceedings book. XXVII Congreso de Ecuaciones Diferenciales y XVII Congreso de Matemática Aplicada*, pages pp 89–96, July 2022.
- [162] L. Martaud and C. Berthon. Fully well-balanced entropy stable Godunov numerical schemes for the Shallow Water equations with the topography source term. 2023.
- [163] D. Mavriplis. Revisiting the least squares procedure for gradient reconstruction on unstructured meshes. *National Aeronautics and Space Administration under NASA Grant No. NCC-1-02043*, 12 2003.
- [164] V. Michel-Dansac. *Development of high-order well-balanced schemes for geophysical flows*. Theses, Université de Nantes, Faculté des sciences et des techniques., Sept. 2016.
- [165] V. Michel-Dansac, C. Berthon, S. Clain, and F. Foucher. A well-balanced scheme for the shallow-water equations with topography. *Computers and Mathematics with Applications*, 72, 09 2015.
- [166] V. Michel-Dansac, C. Berthon, S. Clain, and F. Foucher. A well-balanced scheme for the shallow-water equations with topography or manning friction. *Journal of Computational Physics*, 335 :115–154, 2017.
- [167] J. Montagnier. *Etude de schémas numériques d’ordre élevé pour la simulation de dispersion de polluants dans des géométries complexes*. Theses, Université Claude Bernard - Lyon I, July 2010.
- [168] A. Navas-Montilla and J. Murillo. High-order schemes for hyperbolic conservation laws. 04 2016.
- [169] H. Nishikawa. An implicit gradient method for cell centered finite volume solver on unstructured grids. *AIAA SciTech Forum*, 01 2019.
- [170] S. Noelle, Y. Xing, and C. Shu. High-order well-balanced schemes. *Numerical Methods for Balance Laws*, 24, 01 2009.
- [171] S. Noelle, Y. Xing, and C.-W. Shu. High-order well-balanced finite volume WENO schemes for shallow water equation with moving water. *J. Comput. Phys.*, 226(1) :29–58, 2007.
- [172] M. Pelanti, L. Quartapelle, and L. Vigevano. A review of entropy fixes as applied to Roe’s linearization, 2001.

- [173] M. D. Perlman. Jensen’s inequality for a convex vector-valued function on an infinite-dimensional space. *Journal of Multivariate Analysis*, 4(1) :52–65, 1974.
- [174] B. Perthame and Y. Qiu. A variant of van Leer’s method for multidimensional systems of conservation laws. *Journal of Computational Physics*, 112(2) :370–381, 1994.
- [175] A. Prieto-Arranz, L. Ramirez, I. Couceiro, I. Colominas, and X. Nogueira. A well-balanced sph-ale scheme for shallow water applications. *Journal of Scientific Computing*, 88, 09 2021.
- [176] S. Qian, F. Shao, and G. Li. High order well-balanced discontinuous Galerkin methods for shallow water flow under temperature fields. *Computational and Applied Mathematics*, 37, 06 2018.
- [177] H. Ranocha. Shallow water equations : split-form, entropy stable, well-balanced, and positivity preserving numerical methods. *GEM - International Journal on Geomathematics*, 8(1) :85–133, dec 2016.
- [178] H. Ranocha, M. Sayyari, L. Dalcin, M. Parsani, and D. Ketcheson. Relaxation Runge-Kutta methods : Fully discrete explicit entropy-stable schemes for the compressible Euler and Navier-Stokes equations. *SIAM Journal on Scientific Computing*, 42 :A612–A638, 01 2020.
- [179] S. Rathan and G. Naga Raju. A modified fifth-order WENO scheme for hyperbolic conservation laws. *Computers and Mathematics with Applications*, 75(5) :1531 – 1549, 2018.
- [180] D. Ray, P. Chandrashekar, U. Fjordholm, and S. Mishra. Entropy stable scheme on two-dimensional unstructured grids for Euler equations. *Communications in Computational Physics*, 19 :1111–1140, 05 2016.
- [181] A. Rehman, I. Ali, S. Zia, and S. Qamar. Well-balanced finite volume multi-resolution schemes for solving the Ripa models. *Advances in Mechanical Engineering*, 13 :168781402110034, 03 2021.
- [182] F. Renac. Entropy stable DGSEM for nonlinear hyperbolic systems in nonconservative form with application to two-phase flows. *Journal of Computational Physics*, 382, 04 2019.
- [183] P. I. Richards. Shock waves on the highway. *Operations Research*, 4 :42–51, 1956.
- [184] P. L. Roe. Approximate Riemann solvers, parameter vectors, and difference schemes. *J. Comput. Phys.*, 43(2) :357–372, 1981.
- [185] V. Rusanov. The calculation of the interaction of non-stationary shock waves and obstacles. *USSR Computational Mathematics and Mathematical Physics*, 1(2) :304–320, 1962.
- [186] M. Saleem, W. Ashraf, S. Zia, I. Ali, and S. Qamar. A kinetic flux vector splitting scheme for shallow water equations incorporating variable bottom topography and horizontal temperature gradients. *PLOS ONE*, 13 :e0197500, 05 2018.
- [187] M. Saleem, S. Zia, W. Ashraf, I. Ali, and S. Qamar. The space–time CESE scheme for shallow water equations incorporating variable bottom topography and horizontal temperature gradients. *Computers and Mathematics with Applications*, 75, 11 2017.
- [188] G. Schnücker, N. Krais, T. Bolemann, and G. Gassner. Entropy stable discontinuous Galerkin schemes on moving meshes for hyperbolic conservation laws. *Journal of Scientific Computing*, 82, 03 2020.
- [189] S. Schochet. The compressible euler equations in a bounded domain : Existence of solutions and the incompressible limit. *Communications in Mathematical Physics*, 104(1) :49–75, Mar. 1986.
- [190] D. Serre. *Systems of conservation laws. 1*. Cambridge University Press, Cambridge, 1999. Hyperbolicity, entropies, shock waves, Translated from the 1996 French original by I. N. Sneddon.

- [191] J. Shi, Y. Zhang, and C. Shu. Resolution of high order weno schemes for complicated flow structures. *Journal of Computational Physics*, 186(2) :690–696, 2003.
- [192] E. Shima, K. Kitamura, and T. Haga. Green-Gauss/weighted least squares hybrid gradient reconstruction for arbitrary polyhedra unstructured grids. *AIAA Journal*, 51 :2740–2747, 11 2013.
- [193] C. W. Shu. Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws. Technical report, NASA, 1997.
- [194] C. W. Shu. High-order finite difference and finite volume WENO schemes and discontinuous Galerkin methods for CFD. *International Journal of Computational Fluid Dynamics*, 17(2) :107–118, 2003.
- [195] C. W. Shu. Discontinuous Galerkin methods : General approach and stability. *Numerical Solutions of Partial Differential Equations*, 01 2009.
- [196] C. W. Shu. A brief survey on discontinuous Galerkin methods in computational fluid dynamics. *Advances in Mechanics*, 43(6) :541–554, 2013.
- [197] C. W. Shu and G. Jiang. On cell entropy inequality for discontinuous Galerkin methods. *Mathematics of computation*, 62 :531–538, 04 1994.
- [198] G. A. Sod. A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws. *Journal of Computational Physics*, 27(1) :1–31, 1978.
- [199] S. Spiegel, H. Huynh, and J. DeBonis. A survey of the isentropic euler vortex problem using high-order methods. 06 2015.
- [200] M. Svärd. Entropy stable boundary conditions for the Euler equations. *Journal of Computational Physics*, 426 :109947, 2021.
- [201] A. Syrakos, S. Varchanis, Y. Dimakopoulos, A. Goulas, and J. Tsamopoulos. A critical analysis of some popular methods for the discretisation of the gradient operator in finite volume methods. *Physics of Fluids*, 29, 06 2016.
- [202] A. Syrakos, S. Varchanis, Y. Dimakopoulos, A. Goulas, and J. Tsamopoulos. On the order of accuracy of the divergence theorem (Green-Gauss) method for calculating the gradient in finite volume methods. working paper or preprint, June 2017.
- [203] C. Sánchez-Linares, T. M. de Luna, and M. Castro. A HLLC scheme for Ripa model. *Applied Mathematics and Computation*, 272, 06 2015.
- [204] E. Tadmor. Numerical viscosity and the entropy condition for conservative difference schemes. *Mathematics of Computation*, 43 :369–381, oct 1984.
- [205] E. Tadmor. The numerical viscosity of entropy stable schemes for systems of conservation laws. I. *Mathematics of Computation*, 49(179) :91–103, 1987.
- [206] E. Tadmor. Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems. *Acta Numerica*, 12 :451 – 512, 05 2003.
- [207] S. Tann, X. Deng, Y. Shimizu, R. Loubère, and F. Xiao. Solution property preserving reconstruction for finite volume scheme : a BVD+MOOD framework. *International Journal for Numerical Methods in Fluids*, Dec. 2019.
- [208] M. Thanh. The Riemann problem for the shallow water equations with horizontal temperature gradients. *Applied Mathematics and Computation*, 325 :159–178, 05 2018.

- [209] N. Thanh, M. Thanh, and D. H. Cuong. Godunov-type numerical scheme for the shallow water equations with horizontal temperature gradient. *Taiwanese Journal of Mathematics*, 24(1) :179 – 223, 2020.
- [210] E. Toro. High-order ADER schemes for hyperbolic equations : a review. Laboratory of Applied Mathematics University of Trento, Italy.
- [211] E. Toro. *Riemann solvers and numerical methods for fluid dynamics*. Springer-Verlag, Berlin, third edition, 2009. A practical introduction.
- [212] E. Toro, R. C. Millington, and L. A. M. Nejad. *Godunov Method*, chapter Towards very high-order Godunov schemes., pages 907–940. Springer US, 2001.
- [213] E. Toro, M. Spruce, and W. Speares. Restoration of the contact surface in the HLL-Riemann solver. *Shock waves*, 4(1) :25–34, 1994.
- [214] E. Toro and V. Titarev. Solution of the generalized Riemann problem for advection-reaction equations. *Proc. Royal Society of London*, 458 :271–281, 2002.
- [215] E. Toro and V. Titarev. Derivative Riemann solvers for systems of conservation laws and ADER methods. *Journal of Computational Physics*, 212 :150–165, 2006.
- [216] R. Touma and C. Klingenberg. Well-balanced central finite volume methods for the Ripa system. *Applied Numerical Mathematics*, 97 :42–68, 2015.
- [217] P. Tsoutsanis. Stencil selection algorithms for WENO schemes on unstructured meshes. *Journal of Computational Physics : X*, 4 :100037, 2019.
- [218] J. Vaassen, D. Vigneron, and J.-A. Essers. An implicit high order finite volume scheme for the solution of 3d navier–stokes equations with new discretization of diffusive terms. *Journal of Computational and Applied Mathematics*, 215(2) :595–601, 2008.
- [219] B. van Leer. Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov’s method. *Journal of computational Physics*, 32(1) :101–136, 1979.
- [220] F. Vilar, P. Maire, and R. Abgrall. Cell-centered discontinuous Galerkin discretizations for two-dimensional scalar conservation laws on unstructured grids and for one-dimensional lagrangian. *Computers and Fluids*, 46, 07 2011.
- [221] Q. VIVILLE. *A method of hp-adaptation for Residual Distribution schemes*. PhD thesis, Bordeaux, 2016.
- [222] N. Wang, M. Li, R. Ma, and L. Zhang. Accuracy analysis of gradient reconstruction on isotropic unstructured meshes and its effects on inviscid flow simulation. *Advances in Aerodynamics*, pages 1–18, 10 2019.
- [223] X. Wen, W. Don, Z. Gao, and Y. Xing. Entropy stable and well-balanced discontinuous Galerkin methods for the nonlinear shallow water equations. *Journal of Scientific Computing*, 83, 06 2020.
- [224] P. Woodward and P. Colella. The numerical simulation of two-dimensional fluid flow with strong shocks. *Journal of Computational Physics*, 54 :115–173, 04 1984.
- [225] X. Wu, N. Trask, and J. Chan. Entropy stable discontinuous Galerkin methods for the shallow water equations with subcell positivity preservation, 2021.
- [226] Y. Xing, X. Zhang, and C. Shu. Positivity-preserving high-order well-balanced discontinuous Galerkin methods for the shallow water equations. *Advances in Water Resources*, 33(12) :1476–1493, 2010.

- [227] Z. Xu, Y. Liu, H. Du, G. Lin, and C. W. Shu. Point-wise hierarchical reconstruction for discontinuous Galerkin and finite volume methods for solving conservation laws. *Journal of Computational Physics*, 230(17) :6843 – 6865, 2011.
- [228] G. Yan, S. Kaur, J. W. Banks, and J. Hicken. Entropy-stable discontinuous Galerkin difference methods for hyperbolic conservation laws, 2021.
- [229] Y. Zahran. A central WENO-TVD scheme for hyperbolic conservation laws. *Novi Sad J. Math*, 36 :25–42, 01 2006.
- [230] H. Zakerzadeh and U. Fjordholm. High-order accurate, fully discrete entropy stable schemes for scalar conservation laws. *IMA Journal of Numerical Analysis*, 36(2) :633–654, 05 2015.
- [231] Q. Zhang and W. Sheng. A note on the Riemann problem for shallow water equations with discontinuous topography. *Applied Mathematics Letters*, 116 :107042, 2021.
- [232] Z. Zhao and J. Qiu. A hermite WENO scheme with artificial linear weights for hyperbolic conservation laws. *Journal of Computational Physics*, 417 :109583, 2020.
- [233] G. Zwas and S. Abarbanel. Third and fourth order accurate schemes for hyperbolic equations of conservation law form. *Mathematics of Computation*, 25(114), 1971.

**Titre :** Schémas aux volumes finis avec inégalités d'entropie discrètes pour des systèmes hyperboliques non linéaires.

**Mot clés :** Systèmes hyperboliques, schémas volume finis, inégalités d'entropie discrètes, schémas *well-balanced*, schémas d'ordre élevé

**Résumé :** Ce mémoire de thèse concerne le développement de schémas numériques aux volumes finis qui approchent les solutions de systèmes d'équations hyperboliques non linéaires. Ces schémas doivent respecter des critères de stabilité considérés au sens des inégalités d'entropie discrètes et des critères de précisions tels que des propriétés *well-balanced* ou d'ordre élevé. L'obtention d'une inégalité d'entropie discrète locale est proposée sous la forme de conditions suffisantes directement introduites dans la définition des schémas numériques. Cette approche est ap-

pliquée aux cas des systèmes d'Euler, de Saint Venant et de Ripa. Pour ces deux derniers systèmes, les schémas entropiques proposés sont complétés d'une propriété *well-balanced*. Par ailleurs, des schémas d'ordre élevé, sans limiteurs de pente et qui vérifient une inégalité d'entropie discrète globale sont également proposés pour un système hyperbolique quelconque. Ces schémas sont définis en une dimension d'espace et des extensions sur des maillages non structurés bidimensionnels sont également réalisées.

**Title:** Discrete entropy inequalities of finite volume schemes for nonlinear hyperbolic systems

**Keywords:** Hyperbolic systems, finite volume schemes, discrete entropy inequalities, well-balanced schemes, high-order schemes

**Abstract:** This thesis concerns finite volume schemes which approximate the solutions of non linear hyperbolic system of equations. These schemes have to satisfy stability criteria in the sense of discrete entropy inequalities and accuracy criteria such as well-balanced or high-order properties. Sufficient conditions are proposed to ensure a fully discrete local entropy inequality in the definition of numerical schemes. These conditions are used to

design numerical schemes for the Euler, the Shallow Water and the Ripa systems. For the two last systems, the entropic schemes are completed to guarantee a well-balanced property. In the other hand, high-order non limited finite volume schemes are also proposed. These schemes satisfy a fully discrete global entropy inequality in one-dimensional space and their extensions on two-dimensional unstructured meshes are also done.