



Analyse numérique de systèmes hyperboliques-dispersifs

Clémentine Courtès

► To cite this version:

Clémentine Courtès. Analyse numérique de systèmes hyperboliques-dispersifs. Analyse numérique [math.NA]. Université Paris-Saclay, 2017. Français. NNT : 2017SACLS467 . tel-01649841

HAL Id: tel-01649841
<https://theses.hal.science/tel-01649841>

Submitted on 27 Nov 2017

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

L'UNIVERSITÉ PARIS-SACLAY

École doctorale de mathématiques Hadamard (EDMH, ED 574)

Établissement d'inscription : Université Paris-Sud

Laboratoire d'accueil : Laboratoire de mathématiques d'Orsay, UMR 8628 CNRS

Spécialité de doctorat : Mathématiques appliquées

Clémentine COURTÈS

Analyse numérique de systèmes hyperboliques-dispersifs

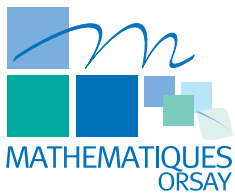
Date de soutenance : 23 novembre 2017 à Orsay

Après avis des rapporteurs : CLAIRE CHAINAIS-HILLAIRET (Université Lille 1)
PASCAL NOBLE (INSA de Toulouse)

Jury de soutenance :

CLAIRE CHAINAIS-HILLAIRET	(Université Lille 1) Rapporteur
RAPHAÈLE HERBIN	(Université d'Aix-Marseille) Présidente du jury
FRÉDÉRIC LAGOUTIÈRE	(Université Lyon 1) Co-Directeur de thèse
BERTRAND MAURY	(Université Paris-Sud) Examinateur
PASCAL NOBLE	(INSA de Toulouse) Rapporteur
FRÉDÉRIC ROUSSET	(Université Paris-Sud) Co-Directeur de thèse

Clémentine Courtès
clementine.courtes@gmail.com
<http://www.math.u-psud.fr/~courtes>



Thèse préparée au
Département de Mathématiques d'Orsay
Laboratoire de Mathématiques (UMR 8628)
Université Paris-Sud
F-91405 Orsay Cedex

Résumé

Analyse numérique de systèmes hyperboliques-dispersifs

Mots Clefs : équations aux dérivées partielles, équation de Korteweg-de Vries, différences finies, estimations d'erreur, convergence numérique, ondes progressives

Résumé : Le but de cette thèse est d'étudier certaines équations aux dérivées partielles hyperboliques-dispersives. Une part importante est consacrée à l'analyse numérique et plus particulièrement à la convergence de schémas aux différences finies pour l'équation de Korteweg-de Vries et les systèmes $abcd$ de Boussinesq. L'étude numérique suit les étapes classiques de consistance et de stabilité. Nous transposons au niveau discret la propriété de stabilité fort-faible des lois de conservations hyperboliques. Nous déterminons l'ordre de convergence des schémas et le quantifions en fonction de la régularité de Sobolev de la donnée initiale. Si nécessaire, nous régularisons la donnée initiale afin de toujours assurer les estimations de consistance. Une étape d'optimisation est alors nécessaire entre cette régularisation et l'ordre de convergence du schéma. Une seconde partie est consacrée à l'existence d'ondes progressives pour l'équation de Korteweg-de Vries-Kuramoto-Sivashinsky. Par des méthodes classiques de systèmes dynamiques : système augmenté, fonction de Lyapunov, intégrale de Melnikov, par exemple, nous démontrons l'existence d'ondes oscillantes de petite amplitude.



Numerical analysis of hyperbolic-dispersive systems

Keys words : partial differential equations, Korteweg-de Vries equation, finite differences, error estimates, numerical convergence, traveling waves

Abstract : The aim of this thesis is to study some hyperbolic-dispersive partial differential equations. A significant part is devoted to the numerical analysis and more precisely to the convergence of some finite difference schemes for the Korteweg-de Vries equation and $abcd$ systems of Boussinesq. The numerical study follows the classical steps of consistency and stability. The main idea is to transpose at the discrete level the weak-strong stability property for hyperbolic conservation laws. We determine the convergence rate and we quantify it according to the Sobolev regularity of the initial datum. If necessary, we regularize the initial datum for the consistency estimates to be always valid. An optimization step is thus necessary between this regularization and the convergence rate of the scheme.

A second part is devoted to the existence of traveling waves for the Korteweg-de Vries-Kuramoto-Sivashinsky equation. By classical methods of dynamical systems : extended systems, Lyapunov function, Melnikov integral, for instance, we prove the existence of oscillating small amplitude traveling waves.

Remerciements

 'est avec un peu d'émotion et beaucoup de reconnaissance que j'écris ces quelques lignes qui achèvent mon aventure orcéenne. Un peu d'émotion car trois années viennent de se terminer sans que je m'en aperçoive. Beaucoup de reconnaissance ensuite envers toutes les personnes qui m'ont aidée. Certaines m'ont suivie depuis le début, de nouvelles têtes sont arrivées en cours de voyage... Ce sont toutes ces personnes que je voudrais remercier ici, toutes celles sans qui ces trois années n'auraient pas été aussi exceptionnelles.

Mes plus sincères remerciements vont à mes directeurs : Frédéric et Frédéric. Que de choses j'ai apprises à vos côtés ! Pour vous être lancés dans cette folle aventure avec moi et y avoir toujours cru (parfois plus fort que moi-même), pour m'avoir fait confiance dès le début et dès mon stage de master, pour toute la patience dont vous avez fait preuve devant mes questions, je ne peux que vous remercier. Vous avez su me remotiver par un simple mot, un geste, quelquefois sans même vous en rendre compte. Je vous remercie pour le sujet passionnant que vous m'avez proposé, pour m'avoir poussée à travailler avec d'autres, pour m'avoir montré que les mathématiques passaient aussi par les discussions, les collaborations et les échanges. Merci de m'avoir toujours incitée à exposer mes travaux, à assister à des conférences, des séminaires, à m'intéresser à des sujets connexes, à participer au CEMRACS, à faire une SEME... Pour tout cela, je vous suis très reconnaissante.

J'adresse ma profonde gratitude à Pascal Noble et Claire Chainais-Hillairet qui ont aimablement accepté de rapporter ma thèse. Je vous remercie sincèrement pour votre relecture attentive et vos remarques constructives qui ont permis d'améliorer mon manuscrit. Je suis aussi très reconnaissante à Raphaële Herbin et Bertrand Maury pour avoir accepté de faire partie de mon jury de soutenance. Merci de l'intérêt que vous avez porté à mes travaux.

Je remercie vivement toutes les personnes avec qui j'ai eu la chance de travailler. Cosmin, je te remercie pour notre collaboration lors de ma dernière année de thèse, travailler avec toi a été très enrichissant et motivant. Je te remercie pour m'avoir toujours accueillie à Paris-Est avec le sourire, pour ta bonne humeur et pour toutes nos conversations. Un grand merci également à Herbert, Emmanuel et Philippe pour notre projet du CEMRACS. Je vous remercie de m'avoir permis de travailler sur un sujet nouveau pour moi et très intéressant, merci pour toute l'aide que vous m'avez apportée et pour tout ce que vous m'avez appris. Je remercie aussi toutes les personnes qui m'ont accompagnée dans mes premiers pas d'enseignante, les enseignants-chercheurs comme les élèves.

J'associe à ces remerciements la direction du laboratoire d'Orsay pour le bon environnement de travail dont j'ai pu bénéficier, tous les membres de l'équipe ANEDP pour votre accueil amical et la bonne ambiance de travail, le personnel de la bibliothèque ainsi que Frédéric Paulin et Stéphane Nonnenmacher pour l'école doctorale. Un grand merci à Catherine Poupon, Estelle Savinien, Marie-Christine Myoupo, Valérie Blandin Lavigne, Florence Rey et Christelle Pires pour votre efficacité, pour votre disponibilité et pour m'avoir grandement facilité les démarches admi-

Remerciements

nistratives. Je me souviendrai longtemps des discussions toujours très agréables avec Catherine et avec Estelle.

Je suis très reconnaissante aux organisateurs des séminaires, conférences, écoles, SEME et CEMRACS auxquels j'ai participé, à l'ANR BoND qui m'a permis d'exposer mes travaux à plusieurs reprises et qui a financé quelques-uns de mes trajets. Merci à Xavier Blanc pour les lettres de recommandation, merci à Philippe et Raphaël de m'accepter à Strasbourg début décembre et merci à toutes les personnes qui ont rendu ce post-doctorat possible auprès du rectorat : mes directeurs bien sûr, mais aussi Frédéric Paulin, Stéphane Nonnenmacher, Patrick Gérard, Denis Merlet et Sylvie Retailleau pour Orsay, Claudine Picaronny et les membres de la présidence pour l'ENS de Cachan et enfin les nombreux soutiens que j'ai eus à Strasbourg de la part de Philippe, Raphaël, Yann Bugeaud et Michel Deneken.

Je tiens à remercier tous les professeurs que j'ai eus au cours de ma scolarité ou qui m'ont encadrée en stage. J'ai une pensée particulière pour Frédéric Pascal, que je ne remercierai jamais assez de m'avoir donné ma chance en m'acceptant à l'ENS de Cachan, pour Claudine Picaronny, qui m'a suivie et encouragée au moment où j'en avais le plus besoin l'année de l'agrégation et pour tous les professeurs d'Orsay qui m'ont tellement appris durant mon master 2. Ce fut un réel plaisir que d'assister à vos cours.

Mon quotidien ces dernières années a été marqué par la bonne ambiance qui règne dans les bureaux 256-258 et je voudrais remercier tous les (ex)-doctorants qui les ont occupés. Merci aux plus anciens qui m'ont accueillie et merci aux plus jeunes de s'être si bien intégrés. J'ai une pensée particulière pour Nina (merci pour tous les précieux conseils que tu m'as donnés depuis mon master et que tu continues encore à me donner) et pour mes co-bureaux Fatima, Anthony et Perla. Fatima, merci pour ta gentillesse et ton calme, je suis contente d'avoir vécu cette aventure thésardesque avec toi, au même rythme. Notre rentrée commune en master 2 me semble encore si proche... Anthony, merci pour tes plaisanteries, pour ton aide si précieuse (quant aux TP de Polytech, quant à ma thèse...), merci de m'avoir fait répéter certains de mes exposés les vendredis soirs après dîner. Perla, merci pour ton naturel, ta bonne humeur et ton rire dans le bureau la dernière année de ma thèse. J'emporte avec moi de très bons souvenirs et je me rappellerai encore longtemps des cours de zumba avec Perla, Fatima et Jana, de ce trajet Massy-Orsay à pied un jeudi matin à 7h00 avec Anthony, du bureau empaqueté de Samer, des discussions très agréables le matin avec Fabien ou Hugo et de celles toujours aussi appréciables de fin de journée avec Yueyuan, d'Alpàr que je connaissais initialement par sa femme Timéa, des taquineries de Loïc et Christèle, des pauses thé/goûter, des biscuits alsaciens de Paul, du barbecue chez Pierre, des habitudes occitanes d'Antonin, de la gentillesse de Tony, de cette soirée pizza improvisée chez Nina (et Jérémie). Je n'oublie pas les autres doctorants et stagiaires qui ont un jour poussé la porte du bureau : Andréa, Jeanne, Xia Bo, Tomas, Caroline...

Je remercie également tous les doctorants que j'ai eu l'occasion de croiser : Hélène (merci pour ces six semaines passées au CIRM, pour m'avoir hébergée à Rennes et à Lyon, merci pour notre amitié tout simplement), Lison et Jeanne (pour vous être occupées du séminaire des doctorants avec moi), Matthieu B., Romain, Alexandre, Ranine, Laura, Charlotte, Florian, Victor, Hala, Olivier, Matthieu A., Matthew, Nathalie et bien sûr tous les (ex)-doctorants croisés au CEMRACS 2015 ou à Strasbourg et dont la liste serait trop longue. Merci à ceux qui m'ont invitée à parler (et invitée tout court) (Olivia, Hélène, Matthieu, Olivier, Fabien, Cosmin, Grégory...). Merci à mes amies plus anciennes : Marine pour avoir gardé le contact depuis la maternelle, Mélanie pour tous ces bons moments passés depuis l'ENS, pour tous nos points communs et pour nos discussions interminables. Un très grand merci à tous ceux que j'ai involontairement oubliés.

J'en viens maintenant aux remerciements plus personnels, à l'égard des personnes qui me sont les plus chères et sans qui rien de tout cela n'aurait pu voir le jour. Un grand merci à ma famille, mes grands-parents, mes oncles, tantes, cousins, cousines, parrain et marraine qui se sont toujours intéressés à ce que je faisais sans forcément tout comprendre (merci à Philippe qui

a bien retenu que tout se résumait à une question de chou-fleur). J'ai une pensée émue pour papy Bernard et Olivier qui, je pense, auraient été ravis de voir l'aboutissement de cette thèse. Merci également à ma petite famille de Sarreguemines que je retrouve toujours avec plaisir certains week-ends.

J'adresse un immense merci à mes deux sœurs : Virginie et Juliette. Merci de n'avoir jamais réussi à prononcer le nom de Korteweg-de Vries convenablement, d'avoir rebaptisé le CEMRACS la «rando-math» de l'été, de voir des «maillages de Descartes» partout (il faudra un jour que vous m'expliquiez ce que c'est) et de trouver le nom «Boussinesq» vraiment trop mignon. Merci tout simplement pour toutes ces blagues que nous ne sommes que trois à comprendre, merci pour nos virées à Rouen, à Bilbao, à Bruxelles, à Vintimille. C'est avec vous que j'ai eu mes plus beaux fous rires et j'espère qu'on en aura encore longtemps. Je suis fière de vous !

Je ne sais comment te remercier, Guillaume. Merci d'être ce que tu es, merci de me soutenir, de m'encourager, de me comprendre si bien (parfois mieux que moi-même). Merci de ta patience infinie à m'écouter (et tu sais combien je peux être pipelette) et à me conseiller. Tu éclaires mon quotidien même à des kilomètres d'écart et tu me permets de donner le meilleur de moi, de m'évader de la recherche, de me faire rêver... Un immense merci. Ma vie est tellement plus belle à tes côtés.

Je voudrais remercier pour finir les personnes qui sont à mes côtés depuis le début : mes parents, Jean-François et Isabelle. Je ne vous remercierai jamais assez pour tout ce que vous avez fait pour moi. Vous m'avez toujours écoutée avec attention parler de mes cours puis de ma recherche. Vous n'avez pas compté votre temps et votre énergie, allant jusqu'à relire certaines parties de cette thèse sans pour autant être vous-même scientifiques (ou commencer à apprendre l'anglais pour papa ?). Merci pour tout votre soutien moral, merci pour vos petites attentions qui me touchent à chaque fois. Sans vous je n'en serais pas arrivée jusque là et s'il ne fallait remercier que deux personnes, ce seraient vous. Cette thèse vous doit tellement...

À vous tous, tout simplement merci.

Glossaire

Notations

$\lfloor \cdot \rfloor$	partie entière.
$\llbracket a, b \rrbracket$	l'ensemble des entiers entre a et b : $\{a, a+1, a+2, \dots, b-2, b-1, b\}$.
$a \lesssim b$	il existe une constante c telle que $a \leq cb$.
$\widehat{u}$	transformée de Fourier de u .
$H^s(\mathbb{R})$	espace de Sobolev d'indice s dont la norme est
	$\ u\ _{H^s(\mathbb{R})} := \left(\int_{\mathbb{R}} (1 + \xi ^2)^s \widehat{u}(\xi) ^2 d\xi \right)^{\frac{1}{2}}.$
$\mathbb{1}_A$	fonction indicatrice de A valant 1 si $x \in A$ et 0 sinon.

Équations étudiées

Airy	l'équation d'Airy $\partial_t u + \partial_x^3 u = 0$.
BBM	l'équation de Benjamin-Bona-Mahony $(I - \partial_x^2) \partial_t u + \partial_x \left(\frac{u^2}{2} \right) = 0$.
Burgers	l'équation de Burgers $\partial_t u + \partial_x \left(\frac{u^2}{2} \right) = 0$.
Fisher-KPP	l'équation de Fisher-Kolmogorov-Petrovsky-Piskunov $\partial_t u = \partial_x^2 u + u(1-u)$.
KdV	l'équation de Korteweg-de Vries $\partial_t u + \partial_x \left(\frac{u^2}{2} \right) + \partial_x^3 u = 0$.
KdV ⁻	l'équation de Korteweg-de Vries avec dispersion négative
	$\partial_t u + \partial_x \left(\frac{u^2}{2} \right) - \partial_x^3 u = 0.$
KdV-4	l'équation de Korteweg-de Vries avec diffusion d'ordre 4
	$\partial_t u + \partial_x \left(\frac{u^2}{2} \right) = \mu \partial_x^3 u - \varsigma \partial_x^4 u, \quad \mu, \varsigma > 0.$
KdV-KS	l'équation de Korteweg-de Vries-Kuramoto-Sivashinsky
	$\partial_t u + \partial_x \left(\frac{u^2}{2} \right) = \lambda \partial_x^2 u + \mu \partial_x^3 u - \varsigma \partial_x^4 u, \quad \lambda, \mu, \varsigma \geq 0.$
KdV-B	l'équation de Korteweg-de Vries-Burgers
	$\partial_t u + \partial_x \left(\frac{u^2}{2} \right) = \lambda \partial_x^2 u + \mu \partial_x^3 u, \quad \lambda, \mu > 0.$
système $abcd$	le système suivant : $(I - b\partial_x^2) \partial_t \eta + (I + a\partial_x^2) \partial_x u + \partial_x(\eta u) = 0$ $(I - d\partial_x^2) \partial_t u + (I + c\partial_x^2) \partial_x \eta + \partial_x \left(\frac{u^2}{2} \right) = 0.$

Schémas numériques

Δx	pas d'espace du schéma.
Δt	pas de temps du schéma.
D_+	opérateur de dérivation discrète décentré à droite, $D_+(a)_j = \frac{a_{j+1} - a_j}{\Delta x}$.
D_-	opérateur de dérivation discrète décentré à gauche, $D_-(a)_j = \frac{a_j - a_{j-1}}{\Delta x}$.
D_c	opérateur de dérivation discrète centré, $D_c(a)_j = \frac{D_+(a)_j + D_-(a)_j}{2}$.
u_Δ	solution exacte moyennée sur une maille en espace (et parfois en temps)
	$(u_\Delta)_j^n = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u(t^n, y) dy$ ou $(u_\Delta)_j^n = \frac{1}{[\inf(T, t^{n+1}) - t^n] \Delta x} \int_{t^n}^{\inf(T, t^{n+1})} \int_{x_j}^{x_{j+1}} u(s, y) dy ds.$
$(v_j^n)_{(j,n)}$	solution numérique.

Glossaire

$\left(\epsilon_j^n\right)_{(j,n)}$	erreur de consistance.
$\left(e_j^n\right)_{(j,n)}$	erreur de convergence $e_j^n = v_j^n - (u_\Delta)_j^n$.
$\ell^\infty(\llbracket 0, N \rrbracket, \ell_\Delta^2(\mathbb{Z}))$	espace fonctionnel dans lequel l'ordre de convergence est étudié, de norme
$ a _{\ell^\infty(\llbracket 0, N \rrbracket, \ell_\Delta^2(\mathbb{Z}))} := \sup_{n \in \llbracket 0, N \rrbracket} \sqrt{\sum_{j \in \mathbb{Z}} \Delta x a_j^n ^2}.$	

Table des matières

Résumé	i
Remerciements	iii
Glossaire	vii
Table des matières	ix
Table des figures	xiii

I	Introduction	1
----------	-------------------------------	----------

1	Introduction générale	3
1.1	Dérivation hydrodynamique des systèmes étudiés	3
1.2	Problèmes de Cauchy	9
1.2.1	Résultats d'existence connus pour l'équation de KdV	9
1.2.2	Résultats d'existence connus pour le système $abcd$	11
1.3	Aspect numérique	12
1.3.1	Principaux résultats numériques sur les équations linéaires	12
1.3.2	Principaux résultats numériques sur l'équation de KdV	17
1.3.3	Principaux résultats numériques sur le système $abcd$	22
1.4	Ondes progressives	24
1.4.1	Orbites hétéroclines, orbites homoclines	24
1.4.2	Principaux résultats d'existence d'ondes progressives sur l'équation KdV-KS	25
1.5	Plan de la thèse	29

II	Étude des équations linéaires de dérivées d'ordre élevé	31
-----------	--	-----------

2	Numerical analysis of high-order linear equations	33
2.1	Introduction (en français)	35
2.2	High-order linear dispersive equations	38

Table des matières

2.2.1	Order of accuracy for an initial datum in $H^{4p+2}(\mathbb{R})$	38
2.2.1.1	A review of divided differences	38
2.2.1.2	Consistency estimate	40
2.2.1.3	Stability	42
2.2.1.4	Error estimates	47
2.2.2	Less smooth initial data	48
2.2.2.1	A regularization of u_0	48
2.2.2.2	Initial datum in $H^s(\mathbb{R})$ with $s \geq 2p + 2$	49
2.2.2.3	Initial datum in $H^s(\mathbb{R})$ with $s \geq 0$	51
2.2.3	Numerical results	52
2.3	High-order polyharmonic heat equations	54
2.3.1	For an initial datum in $H^{4p}(\mathbb{R})$	54
2.3.1.1	Consistency estimate	54
2.3.1.2	Stability	56
2.3.1.3	Convergence error	57
2.3.2	For a less smooth initial datum	57
2.3.3	Numerical results	59
III	Analyse numérique de l'équation KdV et du système $abcd$	63
3	Error estimates for a finite difference scheme for the Korteweg-de Vries equation	65
3.1	Introduction (en français)	67
3.2	Known results on the Cauchy problem for the KdV equation	72
3.3	Consistency error estimate	73
3.4	Stability estimate	75
3.4.1	Preliminary results	75
3.4.2	The discrete relative entropy inequality	77
3.5	Rate of convergence	84
3.5.1	Properties of dispersive operators	84
3.5.2	Simplification of Inequality (3.37)	86
3.5.3	Induction method	90
3.6	Convergence for less smooth initial data	95
3.6.1	Proof of convergence Theorem 3.10	96
3.7	Numerical results	99
3.7.1	Convergence rates	99
3.7.2	Smooth initial data	100
3.7.3	For a regularity between $H^3(\mathbb{R})$ and $H^{10}(\mathbb{R})$	101
3.7.4	For a regularity between $L^2(\mathbb{R})$ and $H^3(\mathbb{R})$	103
3.8	Afterword : go one step further in the Burgers and Airy structure of (KdV)	107
3.8.1	The Airy equation	107
3.8.2	The Burgers equation	107
3.8.2.1	The well-known hyperbolic CFL	108
3.8.2.2	Numerical entropy flux for Burgers equation	109
4	Energy estimates for a finite difference scheme on the $abcd$ system	113
4.1	Introduction (en français)	115
4.2	One step towards the $abcd$ system : the KdV-BBM equation	120
4.2.1	Left and right winded numerical schemes	120
4.2.1.1	The case of a non negative dispersion ($a \geq 0$)	121

4.2.1.2	The case of a negative dispersion ($a < 0$)	124
4.2.2	For a centered scheme	124
4.3	Non zero BBM-dispersive term in both equations : $abcd$ system with $b, d > 0$. . .	125
4.3.1	Discrete energy inequality	126
4.3.2	Convergence of the scheme	130
4.4	KdV equation coupled with BBM equation : $abcd$ system with $bd = 0$	131
4.4.1	Burgers-type estimates	132
4.4.2	The case $b = d = 0$	135
4.4.3	The case $b = 0, d > 0$	137
4.4.4	The case $d = 0, b > 0$	142
4.5	Experimental results	147
4.5.1	Linear case	147
4.5.2	Case $b > 0, d > 0$	148
4.5.3	Case $bd = 0$	151
4.6	Afterword : traveling waves collision	152
IV	Étude de l'équation de Korteweg-de Vries-Kuramoto-Sivashinsky	155
5	Small-amplitude traveling waves for KdV-KS with fourth order diffusion	157
5.1	Introduction (en français)	159
5.2	Speed of propagation and equilibrium points	163
5.3	The extended system and the small amplitude hypothesis	164
5.4	The center manifold reduction	166
5.4.1	Study of the linear part	167
5.4.2	Projection on the center manifold	168
5.4.3	Foliation in α	169
5.4.4	Blowup of the system : a singular change of coordinates	171
5.5	Singular limit system	173
5.6	Perturbed system : a homoclinic bifurcation	174
5.6.1	The Melnikov integral	175
5.6.2	Study of the homoclinic bifurcation	178
5.7	Global «weak» Lyapunov function : elimination of pathogenic cases	179
5.7.1	A global «weak» Lyapunov function	179
5.7.2	The trapping region	181
5.8	Afterword : Traveling waves for KdV-Burgers equation	184
5.8.1	From KdV-Burgers equation to Fisher-KPP equation	185
6	Stability of the traveling wave for the KdV-KS equation	189
6.1	Introduction (en français)	191
6.2	Linear stability and the Evans function : a revue	192
6.2.1	From the PDE towards the ρ -dependent first-order dynamical system <i>via</i> the spectral problem	192
6.2.2	Spectrum of linearized operator $\mathcal{M}$	193
6.2.2.1	Essential spectrum	193
6.2.2.2	Eigenvalues of $\mathcal{M}$ and Evans function	194
6.3	The ρ -dependent first order system for (KdV-4) equation	195
6.3.1	Limit matrices $\mathbb{A}_{\pm}^{KdV-4}(\rho)$	195
6.3.2	The essential spectrum of $\mathcal{M}^{KdV-4}$	196

Table des matières

6.4	Numerical computation of the Evans function	197
6.4.1	In a complex neighborhood of the origin	197
6.4.2	In the real high-frequency limit	200
6.5	Afterword : Stability for the traveling wave of (KdV ⁻ - B) equation	203
Conclusion		207
V	Annexes	209
A	Computations of consistency errors	211
A.1	Consistency error for the KdV equation (Chapter 3)	211
A.2	Consistency error for the $abcd$ system (Chapter 4)	214
A.2.1	Consistency error for the first equation of the $abcd$ system : ϵ_1^n	214
A.2.2	Consistency error for the second equation of the $abcd$ system : ϵ_2^n	218
B	Proofs of technical Lemmas dealing with the operators D_+, D_-, D_c	221
B.1	Proofs of useful Lemmas of Section 3.4.1	221
B.2	Proofs of useful Lemmas of Section 3.8.2.2	227
C	Traveling wave solution for Fisher-KPP equation	229
C.1	Location of the unstable manifold $\mathcal{W}_{(1,0)}^u$	230
C.2	The proof of Proposition C.1	233
Bibliographie		243

Table des figures

1.1	Variables du problème	4
1.2	Variables w_θ et θH pour le système $abcd$	5
1.3	Écoulement d'un fluide sur un plan incliné	8
1.4	Correspondance onde progressive avec $u_- \neq u_+$ /orbite hétérocline et onde solitaire/orbite homocline	25
2.1	Numerical versus theoretical orders	54
2.2	Experimental versus theoretical orders : case $p = 1$ - Continuous line : theoretical orders, Dots : experimental orders, Dashed line : conjectured orders	60
2.3	Experimental versus theoretical orders : case $p = 2$ - Continuous line : theoretical orders, Dots : experimental orders, Dashed line : conjectured orders	60
2.4	Experimental slopes - Red dots : Crank-Nicolson scheme ($\theta = \frac{1}{2}$) and Black crosses : Implicit scheme ($\theta = 1$)	61
3.1	Experimental rate of convergence for cosinusoidal solution	100
3.2	Experimental rate of convergence for cnoidal-wave solution	101
3.3	Experimental rate of convergence for $u_0 \in H^{\frac{7}{2}-}([0, L])$	102
3.4	Experimental rate of convergence for $u_0 \in H^{\frac{9}{2}-}([0, L])$	102
3.5	Experimental rate of convergence for $u_0 \in H^{4-}([0, L])$	103
3.6	Experimental rate of convergence for $u_0 \in H^{5-}([0, L])$	103
3.7	Experimental rate of convergence for $u_0 \in H^{\frac{1}{2}-}([0, L])$	104
3.8	Experimental rate of convergence for $u_0 \in H^{\frac{3}{2}-}([0, L])$	104
3.9	Experimental rate of convergence for $u_0 \in H^{\frac{5}{2}-}([0, L])$	105
3.10	Experimental rate of convergence for $u_0 \in H^{1-}([0, L])$	105
3.11	Experimental rate of convergence for $u_0 \in H^{2-}([0, L])$	106
3.12	Experimental rate of convergence for $u_0 \in H^{3-}([0, L])$	106
3.13	Rates of convergence according to the Sobolev regularity of u_0 . – Rates proved in this chapter (solid line) <i>versus</i> experimental rates (dots)	106
4.1	Energy conservation in the linear case	148
4.2	Case (A) where $a = 0, b = \frac{1}{6}, c = 0, d = \frac{1}{6}$: results for η (left) and u (right)	148
4.3	Case (B) where $a = 0, b = \frac{1}{6}, c = 0, d = \frac{1}{6}$: results for η (left) and u (right)	149
4.4	Case (C) where $a = -\frac{7}{30}, b = \frac{7}{15}, c = -\frac{2}{5}, d = \frac{1}{2}$: results for η (left) and u (right) . . .	150
4.5	Case (D) where $a = 0, b = \frac{1}{3}, c = -\frac{1}{3}, d = \frac{1}{3}$: results for η (left) and u (right) . . .	151
4.6	Case (E) where $a = b = c = 0, d = \frac{1}{6}$: results for η (left) and u (right)	151

Table des figures

4.7	Case (F) where $a = -\frac{1}{6}, b = c = 0, d = \frac{1}{2}$: results for η (left) and u (right)	152
4.8	Explosion for $(a = 0, b = \frac{1}{6}, c = 0, d = \frac{1}{6})$: results for η (left) and u (right)	153
4.9	No collision for $(a = -\frac{7}{30}, b = \frac{7}{15}, c = -\frac{2}{5}, d = \frac{1}{6})$: results for η (left) and u (right)	154
5.1	Onde progressive pour (KdV ⁻ - 4) ($f(x) = \frac{x^2}{2}$) - À gauche : pour $0 < t < 64$, à droite : pour $40 < t < 80$	160
5.2	Heteroclinic orbit : traveling wave in the phase space in (X, Y) -frame	172
5.3	Level curves of the Hamiltonian $\mathcal{H}$ in the (KdV ⁻ - 4) case ($h \equiv \frac{1}{2}$) and for $\Lambda = 1$ - The blue curve corresponds to the level $\frac{2}{3}$ and the red one to the homoclinic orbit	174
5.4	Two possible scenarios from the homoclinic orbit	175
5.5	Pathogenic cases	175
5.6	Definition of points X_θ^u and X_θ^s - The normal vector is represented in blue.	176
5.7	Traveling wave solution for the KdV ⁻ -Burgers equation	185
6.1	Spectrum of the limit operators $\mathcal{M}_\pm^{KdV-4}$ and essential spectrum of $\mathcal{M}^{KdV-4}$	196
6.2	Evans function $D_-(\rho)$ for the (KdV ⁻ - 4) equation on the real axis	199
6.3	Evans function D_- on the imaginary axis, for the (KdV ⁻ - 4) equation	199
6.4	Evans function D_- on the domain Ω , for the (KdV ⁻ - 4) equation	199
6.5	Evans function for the (KdV ⁻ - B) equation	204
6.6	Evans function D_+ for the (KdV ⁻ - B) equation on the imaginary axis	205
6.7	Evans function on the domain Ω , for the (KdV ⁻ - B) equation	205
C.1	Level sets of $\mathcal{L}$ - the blue domain corresponds to the compact Ω and the red line to the level set going through $(1, 0)$	231
C.2	Vector field of the Fisher-KPP autonomous system and trapping region Ω - the black vectors correspond to unstable eigenvector for $(1, 0)^t$ and stable eigenvector for $(0, 0)^t$ which are connected via the traveling wave	232
C.3	Vector field of the Fisher-KPP autonomous system and trapping triangle $\mathcal{T}$ - the black vectors correspond to unstable eigenvector for $(1, 0)^t$ and stable eigenvector for $(0, 0)^t$	235

PREMIÈRE PARTIE

Introduction

CHAPITRE 1

Introduction générale

*L*es équations dont il sera question dans cette thèse permettent de modéliser des phénomènes hydrodynamiques comme les écoulements dans les rivières et d'expliquer l'apparition de chocs dispersifs avec oscillations. Il s'agit d'équations de type hyperbolique-dispersif, avec une dérivée troisième en espace.

Pour de tels systèmes, comme nous ne disposons pas en général de formules explicites pour les solutions du problème de Cauchy, il est très utile d'utiliser des approximations numériques pour obtenir des informations sur le comportement qualitatif des solutions. Il est nécessaire dans ce cas de s'assurer que la solution numérique converge bien vers la solution exacte.

Le but de cette thèse est d'étudier d'un point de vue numérique et théorique certains de ces systèmes hyperboliques-dispersifs.

1.1 Dérivation hydrodynamique des systèmes étudiés

Nous nous intéressons au mouvement des vagues ou plus généralement au mouvement des ondes de gravité à la surface d'un fluide supposé parfait, incompressible, irrotationnel et soumis à la seule force gravitationnelle. Nous nous focaliserons principalement sur un régime particulier : le régime des grandes longueurs d'onde en eau peu profonde.

Dérivation hydrodynamique du système $abcd$. Comme expliqué en préambule, nous considérons un fluide parfait, incompressible, en écoulement irrotationnel dont la surface libre à l'instant t et à la position $\mathbf{X} = (x, y) \in \mathbb{R}^2$ est donnée par $\eta(t, \mathbf{X})$. Nous supposons de plus que la topographie est celle d'un fond plat, *i.e* $z = -H$. L'amplitude maximale de la surface libre est notée A et λ correspond à sa longueur d'onde, conformément à ce qui est représenté à la figure 1.1.

La dynamique d'un tel fluide peut être décrite par les équations d'Euler à surface libre, écrites ici avec la formulation de Bernouilli :

$$\begin{cases} \Delta_{\mathbf{X},z} \phi = 0, & -H \leq z \leq \eta, \\ \partial_z \phi = 0, & z = -H, \\ \partial_t \eta + \nabla_{\mathbf{X}} \eta \cdot \nabla_{\mathbf{X}} \phi = \partial_z \phi, & z = \eta, \\ \partial_t \phi + \frac{1}{2} (|\nabla_{\mathbf{X},z} \phi|)^2 + gz = 0, & z = \eta, \end{cases}$$

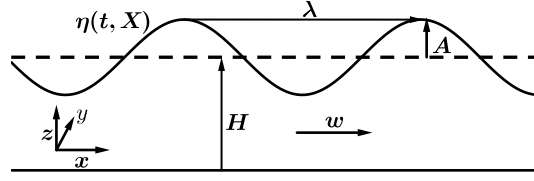


FIGURE 1.1 – Variables du problème

avec $\nabla_{X,z} = (\partial_x, \partial_y, \partial_z)^T$, $\nabla_X = (\partial_x, \partial_y)^T$ et $\Delta_{X,z} = \partial_x^2 + \partial_y^2 + \partial_z^2$. Dans le système précédent, ϕ correspond au potentiel, dont l'existence est assurée par l'hypothèse d'irrotationnalité qui permet d'écrire le champ de vitesse w comme $w = \nabla_{X,z}\phi$. En adimensionnant le système, deux grandeurs caractéristiques apparaissent : le paramètre de non linéarité ϵ ainsi que le paramètre de profondeur μ , définis comme suit

$$\epsilon := \frac{A}{H} \quad \text{et} \quad \mu := \left(\frac{H}{\lambda}\right)^2.$$

La complexité de ce système d'équations conduit à étudier des régimes particuliers liés à des asymptotiques précises sur ϵ et μ . Par exemple, l'hypothèse de faible profondeur d'eau, $\mu \ll 1$, et de grande amplitude, $\epsilon = \mathcal{O}(1)$, permet d'obtenir le système de Saint-Venant¹

$$\begin{cases} \partial_t h + \nabla \cdot (hw) = 0, \\ \partial_t (hw) + \nabla \cdot (hw \otimes w) + \nabla p = 0, \end{cases}$$

où $h(t, X) := \eta(t, X) + H$ représente la hauteur de l'eau, $p(t, X) := \frac{gh^2(t, X)}{2}$ la pression du fluide et g la constante gravitationnelle universelle. Ce système ne prend pas du tout en compte les effets dispersifs.

L'asymptotique qui nous intéressera dans toute la suite correspond au régime des faibles amplitudes en eau peu profonde dans le régime de Boussinesq, qui se traduit par les hypothèses suivantes

$$\epsilon \ll 1, \quad \mu \ll 1 \quad \text{et} \quad \epsilon \approx \mu.$$

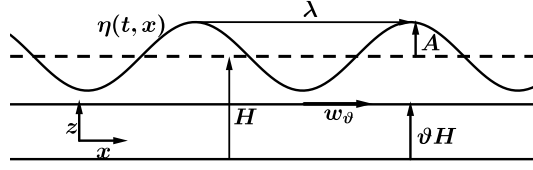
La dernière hypothèse sur l'ordre de grandeur relatif de ϵ et μ s'appelle l'hypothèse de Boussinesq et peut être reformulée en utilisant le nombre de Stokes $S := \frac{\epsilon}{\mu} = \frac{A\lambda^2}{H^3} \approx 1$. Ce régime permet de prendre en compte les termes non-linéaires, dûs aux petites amplitudes mais aussi les termes dispersifs, dûs aux grandes longueurs d'onde et donc d'étudier l'équilibre entre les deux. Supposer que le nombre de Stokes est d'ordre 1 signifie que les effets non linéaires et les effets dispersifs sont du même ordre de grandeur.

Introduisons un degré de liberté supplémentaire et notons $w_\vartheta(t, X)$ la vitesse horizontale à une hauteur ϑH , avec $0 \leq \vartheta \leq 1$ (le fond correspondant à $\vartheta = 0$ et la surface libre à $\vartheta = 1$), voir figure 1.2.

Par abus de notations, nous noterons toujours $X, t, \eta(t, X)$ et ϕ les variables adimensionnées, $\frac{X}{\lambda}, t\frac{\sqrt{gH}}{\lambda}, \frac{\eta(t, X)}{A}$ (la déviation de la surface libre par rapport à sa position au repos) et $\frac{\phi\sqrt{gH}}{gA\lambda}$ [75].

En négligeant les termes du second ordre ou plus (ceux avec ϵ^2 , $\mu\epsilon$ ou μ^2 en facteur), Bona, Chen et Saut [16] obtiennent le système suivant, que nous appellerons système *abcd* (une géné-

1. Ce système a fait l'objet de mon projet du CEMRACS 2015 [35]


 FIGURE 1.2 – Variables w_θ et θH pour le système $abcd$

ralisation à des ondes avec tension de surface se trouve dans [40])

$$\begin{cases} (I - \mu b_\theta \Delta) \partial_t \eta + \operatorname{div} w_\theta + a_\theta \mu \operatorname{div} \Delta w_\theta + \epsilon \operatorname{div} (\eta w_\theta) = 0, \\ (I - \mu d_\theta \Delta) \partial_t w_\theta + \nabla \eta + c_\theta \mu \nabla \Delta \eta + \epsilon w_\theta \cdot \nabla w_\theta = 0, \end{cases} \quad (t, \mathbf{X}) \in [0, T] \times \mathbb{R}^2, \quad (1.1)$$

où I est l'opérateur identité et les paramètres $a_\theta, b_\theta, c_\theta, d_\theta$ vérifient

$$\begin{aligned} a_\theta &= \frac{1}{2} \left(\theta^2 - \frac{1}{3} \right) \kappa_1, & b_\theta &= \frac{1}{2} \left(\theta^2 - \frac{1}{3} \right) (1 - \kappa_1), \\ c_\theta &= \frac{1}{2} (1 - \theta^2) \kappa_2 - Bo, & d_\theta &= \frac{1}{2} (1 - \theta^2) (1 - \kappa_2). \end{aligned}$$

Dans ce qui précède, κ_1 et κ_2 sont des nombres réels sans signification physique, par contre θ représente la hauteur (adimensionnée) de l'eau à laquelle on étudie notre système. La variable Bo représente le nombre de Bond : un nombre sans dimension qui mesure la tension de surface

$$Bo := \frac{\sigma}{\rho g H^2}$$

avec σ la tension de surface (par exemple entre l'eau et l'air), g l'accélération de la pesanteur, ρ la densité du fluide et H la hauteur caractéristique du fluide, déjà introduite précédemment. Notons la relation entre ces quatre coefficients : $a_\theta + b_\theta + c_\theta + d_\theta = \frac{1}{3} - Bo$.

Remarque 1.1. Lorsque la tension de surface est négligée ($Bo = 0$) et que $a_\theta = c_\theta$, alors le système peut être totalement symétrisé, voir [74] avec $\epsilon = \mu$ et le changement de variables $v_\theta := w_\theta (1 + \frac{\epsilon}{2} \eta)$

$$\begin{cases} (I - \epsilon b_\theta \Delta) \partial_t \eta + \operatorname{div} v_\theta + a_\theta \epsilon \operatorname{div} \Delta v_\theta + \frac{\epsilon}{2} \operatorname{div} (\eta v_\theta) = 0, \\ (I - \epsilon d_\theta \Delta) \partial_t v_\theta + \nabla \eta + a_\theta \epsilon \nabla \Delta \eta + \epsilon \left(\frac{1}{2} \eta \nabla \eta + \frac{1}{2} (v_\theta \cdot \nabla) v_\theta + \frac{1}{2} v_\theta \operatorname{div} v_\theta + \frac{1}{4} \nabla |v_\theta|^2 \right) = 0, \end{cases} \quad (1.2)$$

avec la même contrainte sur les paramètres : $2a_\theta + b_\theta + d_\theta = \frac{1}{3}$.

Remarque 1.2. L'erreur entre la solution du système d'Euler à surface libre et la solution du système $abcd$ au temps t est d'ordre $\mathcal{O}(\epsilon^2 t)$, voir [18] ou [2]. Ceci prouve la pertinence de l'approximation des équations d'Euler à surface libre par le système $abcd$.

Remarque 1.3. Nous avons supposé ici un fond plat, cependant de nombreux travaux généralisent les résultats initiaux à des topographies plus variables. Dans [29], deux cas sont étudiés : une faible variation topographique, $\frac{B}{H} = \mathcal{O}(\epsilon)$ avec B l'amplitude maximale du fond, et une variation forte $\frac{B}{H} = \mathcal{O}(1)$. Dans [31], une topographie dépendante du temps est considérée.

Par la suite, nous restreindrons notre étude à un système de dimension 1 ($\mathbf{X} = x$) et omettrons la dépendance en θ de toutes les variables. Le système $abcd$ avec coefficients unitaires s'écrit alors

$$\begin{cases} (I - b\partial_x^2) \partial_t \eta + \partial_x u + a\partial_x^3 u + \partial_x (\eta u) = 0, \\ (I - d\partial_x^2) \partial_t u + \partial_x \eta + c\partial_x^3 \eta + \partial_x \left(\frac{u^2}{2}\right) = 0, \end{cases} \quad (t, x) \in [0, T] \times \mathbb{R}, \quad (\text{abcd})$$

avec la contrainte

$$a + b + c + d = \frac{1}{3} - Bo. \quad (1.3)$$

⋈

Dérivation hydrodynamique de l'équation de KdV. Supposons maintenant une propagation unilatérale de la vague selon les x croissants. Les opérateurs de divergence et de gradient dans le système (1.1) sont donc remplacés par la dérivée partielle ∂_x . Certains choix de paramètres $a_\vartheta, b_\vartheta, c_\vartheta, d_\vartheta$ conduisent à l'équation de Korteweg-de Vries comme par exemple le choix de $a_\vartheta = \frac{1}{3}, b_\vartheta = c_\vartheta = d_\vartheta = 0$ (avec $Bo = 0$) dans [18]. En supposant $\epsilon = \mu$ (en accord avec le régime de Boussinesq), les auteurs diagonalisent une première fois le système (1.1) en définissant les variables $N = w_\vartheta + \eta$ et $V = w_\vartheta - \eta$ puis utilisent un développement WKB de la forme

$$\begin{aligned} N(t, x) &\approx N_0(\epsilon t, x - t) + \epsilon N_1(\epsilon t, t, x), \\ V(t, x) &\approx V_0(\epsilon t, x + t) + \epsilon V_1(\epsilon t, t, x). \end{aligned}$$

À l'ordre le plus bas, ils obtiennent deux équations de Korteweg-de Vries découplées

$$\begin{cases} \partial_t n_0 + \partial_x n_0 + \frac{3}{2} \epsilon \partial_x \left(\frac{n_0^2}{2}\right) + \frac{1}{2} \left(\frac{1}{3} - Bo\right) \epsilon \partial_x^3 n_0 = 0, \\ \partial_t w_0 - \partial_x w_0 + \frac{3}{2} \epsilon \partial_x \left(\frac{w_0^2}{2}\right) - \frac{1}{2} \left(\frac{1}{3} - Bo\right) \epsilon \partial_x^3 w_0 = 0, \end{cases}$$

en posant $n_0(t, x) = N_0(\epsilon t, x - t)$ et $w_0(t, x) = V_0(\epsilon t, x + t)$ (nous avons enlevé la dépendance en ϑ pour plus de clarté).

Remarque 1.4. Il est également possible d'obtenir deux équations découplées de Korteweg-de Vries, en partant du système symétrisé (1.2) avec $a_\vartheta = c_\vartheta = \frac{1}{6} - \frac{1}{2}Bo$ et $b_\vartheta = d_\vartheta = 0$, cf [74] (section 7.1.1).

Remarque 1.5. L'estimation suivante, dont une preuve se trouve dans [74],

$$\|(\eta, w) - (\eta^{KdV}, w^{KdV})\|_{L^\infty([0, t], H^s(\mathbb{R})^2)} = \mathcal{O}(\epsilon t^{\frac{1}{2}}),$$

avec $\eta^{KdV} = \frac{n_0 - w_0}{2}$ et $w^{KdV} = \frac{n_0 + w_0}{2}$ permet de justifier l'approximation des équations d'Euler à surface libre unidirectionnelles par l'équation de Korteweg-de Vries (appelée KdV par la suite).

Remarque 1.6. L'équation de Korteweg-de Vries a été pour la première fois établie par Boussinesq et doit son nom à Korteweg et de Vries qui l'ont redécouverte en 1895 dans [69]. Sa dérivation rigoureuse n'a été envisagée qu'à partir de 1985, voir par exemple [62] et [37] (l'équation de Korteweg-de Vries y est dérivée à partir de l'équation de Boussinesq 1D et non du système abcd comme expliqué ci-dessus).

Comme vu précédemment, l'hypothèse d'une propagation uniforme conduit à l'équation de KdV

$$\partial_t n_0 + \partial_x n_0 + \frac{3}{2} \epsilon n_0 \partial_x n_0 + \frac{1}{2} \left(\frac{1}{3} - Bo\right) \mu \partial_x^3 n_0 = 0. \quad (1.4)$$

En supposant une propagation bi-directionnelle avec des ondes faiblement transverses selon y , nous obtenons l'équation de Kadomtsev-Petviashvili (KP) d'après [60] (une dérivation rigoureuse se trouve dans [2] ou [75])

$$\partial_x \left(\partial_t n_0 + \partial_x n_0 + \frac{3}{2} \epsilon n_0 \partial_x n_0 + \frac{1}{2} \left(\frac{1}{3} - Bo\right) \mu \partial_x^3 n_0 \right) + \frac{\gamma}{2} \partial_y^2 n_0 = 0,$$

avec $\gamma = \frac{\lambda_x^2}{\lambda_y^2}$ le rapport des longueurs d'onde dans les deux directions.

Lorsque la tension de surface est élevée (par exemple dans une eau très peu profonde), $Bo > \frac{1}{3}$, l'équation de KP est appelée équation KP-I. Lorsque la tension de surface est faible voire inexistante, $Bo < \frac{1}{3}$, l'équation KP est appelée équation KP-II.

Remarque 1.7. Pour le cas critique $Bo = \frac{1}{3}$, la dispersion fait intervenir un terme d'ordre cinq, que ce soit pour l'équation de KdV ou l'équation KP [57].

En posant $u(t, x) := n_0(t, x + t)$ (afin d'éliminer le terme de transport) et en s'intéressant aux équations à coefficients unitaires, on trouve les formes (conservatives) plus classiques pour (KdV) et (KP)

$$\partial_t u + \partial_x \left(\frac{u^2}{2} \right) + \partial_x^3 u = 0, \quad (\text{KdV})$$

et

$$\partial_x \left(\partial_t u + \partial_x \left(\frac{u^2}{2} \right) + \partial_x^3 u \right) + \lambda \partial_y^2 u = 0, \quad (\text{KP})$$

avec $\lambda = \pm 1$.

Une étude des différentes asymptotiques de l'équation d'Euler à surface libre est détaillée dans [2], ainsi que l'existence de solutions en temps long.

Remarque 1.8. En linéarisant l'équation (KdV) autour de l'état d'équilibre $u = 0$, nous retrouvons l'équation d'Airy

$$\partial_t u + \partial_x^3 u = 0. \quad (\text{Airy})$$



Dérivation hydrodynamique de l'équation de BBM. Avec les mêmes hypothèses hydrodynamiques que celles faites pour obtenir l'équation de KdV (onde de faible amplitude en eau peu profonde sous l'hypothèse de Boussinesq), il est également possible d'obtenir une variante de KdV, appelée l'équation de Benjamin-Bona-Mahony [10], équation de BBM par la suite :

$$\partial_t n_0 + \partial_x n_0 + \frac{3}{2} \epsilon n_0 \partial_x n_0 - \frac{1}{2} \left(\frac{1}{3} - Bo \right) \mu \partial_t \partial_x^2 n_0 = 0 \quad (1.5)$$

en supposant $Bo < \frac{1}{3}$. En effet, à l'ordre zéro $\partial_x n_0 = -\partial_t n_0 + \mathcal{O}(\epsilon + \mu)$, ce qui permet (au moins formellement) de passer de KdV à l'équation de BBM. L'équation avec des coefficients unitaires prend la forme classique

$$\partial_t (1 - \partial_x^2) u + \partial_x \left(\frac{u^2}{2} \right) = 0. \quad (\text{BBM})$$

Cette équation est aussi appelée équation RLWE pour «regularized long-wave equation», étudiée pour la première fois par Peregrine en 1966 [89] pour l'étude des mascarets de petite amplitude. Cette équation (BBM) possède de meilleures propriétés dispersives que (KdV). En effet, la relation de dispersion pour l'équation (BBM) linéarisée (autour de l'état d'équilibre $u \equiv 1$) est

$$W^{BBM}(\xi) := \frac{\xi}{1 + \xi^2},$$

alors que celle de (KdV) est

$$W^{KdV}(\xi) := \xi - \xi^3.$$

Ainsi, la relation de dispersion n'est pas bornée pour (KdV) et la vitesse de phase $v_\phi(\xi) := \frac{W(\xi)}{\xi} = 1 - \xi^2$ devient négative pour $\xi^2 > 1$ (ce qui contredit l'hypothèse de propagation unidirectionnelle selon le sens des x croissants) [10].

Section 1.1. Dérivation hydrodynamique des systèmes étudiés

Remarque 1.9. Le fait que la relation de dispersion ne soit pas bornée pose aussi problème d'un point de vue plus numérique. En effet, à cause de cette relation non bornée, les phénomènes de petites amplitudes ont tendance à se propager à très grande vitesse dans le sens des x décroissants [10]. Les erreurs numériques sont justement un exemple de phénomènes de petites échelles. Pour avoir un schéma plus stable, il est donc intéressant numériquement d'utiliser l'équation de (BBM).

Remarque 1.10. On peut de même obtenir une équation de type BBM-II en partant de l'équation KP-II, cf [81].



Dérivation hydrodynamique de l'équation KdV-KS. Étudions maintenant le même fluide incompressible, visqueux (de viscosité ν) mais cette fois-ci en mouvement sur un plan incliné d'un angle $\alpha \ll 1$ comme représenté sur la figure 1.3.

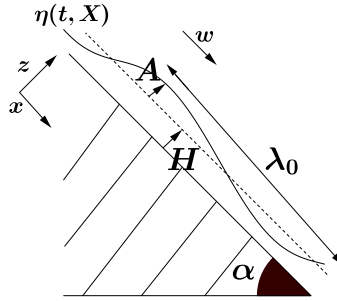


FIGURE 1.3 – Écoulement d'un fluide sur un plan incliné

La hauteur H ne représente plus la hauteur verticale au repos mais la hauteur au repos dans la direction normale au plan incliné (le fluide étant au repos lorsque sa surface est parallèle au plan incliné). L'amplitude maximale A sera choisie normale au plan incliné. Nous supposons le mouvement unidirectionnel et dirigé dans la direction de l'axe des x , seule la variable x sera donc considérée par la suite. La longueur d'onde du fluide dans le sens de la plus grande pente sera notée λ_0 et $w(t, x)$ représentera sa vitesse selon la plus grande pente à l'instant t en position x .

En toute généralité, le fluide vérifie les équations de Navier-Stokes avec condition de non glissement sur le plan incliné. L'asymptotique des grandes longueurs d'onde et des petites amplitudes conduit comme pour l'équation de KdV à introduire les paramètres de non linéarité et de profondeur ϵ et μ . En négligeant les $\mathcal{O}(\mu^{\frac{3}{2}}, \sqrt{\mu}\epsilon)$ dans l'équation, on obtient l'équation de Korteweg-de Vries-Kuramoto-Sivashinsky suivante, équation KdV-KS par la suite, (par abus de notation η sera la variable adimensionnée $\frac{\eta}{A}$), cf [102, 58]

$$\partial_t \eta + 4\eta \partial_x \eta + \delta_1 \partial_x^2 \eta + \delta_2 \partial_x^3 \eta + \delta_3 \partial_x^4 \eta = 0,$$

avec $\delta_1 = \mathcal{O}(\sqrt{\mu}(R - \frac{5}{4}\cot\alpha))$, $\delta_2 = \mathcal{O}(\mu(1 + R(R - \frac{5}{4}\cot\alpha)))$ et $\delta_3 = \mathcal{O}(\mu)$. Le nombre R correspond au nombre de Reynold

$$R := \frac{gH^3 \sin\alpha}{2\nu^2},$$

et représente le rapport entre les forces d'inertie du fluide et ses forces visqueuses. La valeur critique

$$R^* = \frac{5}{4} \cotan \alpha$$

correspond au seuil d'instabilité : pour un nombre de Reynold plus faible les solutions constantes (pour lesquelles la surface du fluide est parallèle au plan incliné) sont stables, pour un nombre de Reynold plus grand, des instabilités se forment [102] et de petites variations autour de l'état d'équilibre croissent exponentiellement. Apparaissent alors des instabilités hydrodynamiques appelées des «roll-waves».

Nous nous intéresserons à l'équation

$$\partial_t u + \partial_x \left(\frac{u^2}{2} \right) = \lambda \partial_x^2 u + \mu \partial_x^3 u - \varsigma \partial_x^4 u, \quad \lambda \geq 0, \mu \geq 0, \varsigma \geq 0, \quad (\text{KdV-KS})$$

et étudierons particulièrement le cas où $R \equiv R^*$ c'est-à-dire l'équation (KdV-4)

$$\partial_t u + \partial_x \left(\frac{u^2}{2} \right) = \mu \partial_x^3 u - \varsigma \partial_x^4 u, \quad \mu > 0, \varsigma > 0. \quad (\text{KdV-4})$$

Attention, les variables λ et μ , dans les équations (KdV-KS) et (KdV-4), sont ici des paramètres de l'équation, ce ne sont plus la longueur d'onde ni le paramètre de faible amplitude.

Remarque 1.11. L'équation de KdV-KS peut aussi être dérivée à partir des équations de Saint-Venant visqueux sur plan incliné

$$\begin{cases} \partial_t h + \partial_x (hw) = 0, \\ \partial_t (hw) + \partial_x \left(hw^2 + \frac{h^2}{2F^2} \right) = h - w^2 + \nu \partial_x (h \partial_x w). \end{cases}$$

Le paramètre F correspond ici au nombre de Froude qui mesure le rapport entre l'énergie cinétique et l'énergie potentielle gravitationnelle du fluide. Une fois les variables adimensionnées pour que la hauteur de la surface libre non perturbée (parallèle au plan incliné) vérifie $H = 1$, le nombre de Froude (adimensionné) vaut

$$F = \sqrt{\frac{\tan(\alpha)}{C}}$$

avec C le coefficient de friction [104]. L'asymptotique formelle autour du nombre de Froude critique ($F \equiv 2$), i.e le passage à la limite $F \xrightarrow{>} 2$ dans les équations de Saint-Venant sur plan incliné, conduit à l'équation de KdV-KS suivante

$$\partial_t \eta + \partial_x \left(\frac{\eta^2}{2} \right) + \delta_4 (\partial_x^2 \eta + \partial_x^4 \eta) + \delta_5 \partial_x^3 \eta = 0,$$

avec $0 < \delta_4 \sim \sqrt{F-2} \ll 1$, cf [9, 105, 86].

1.2 Problèmes de Cauchy

Tout d'abord listons les résultats connus d'existence théorique pour les différents problèmes de Cauchy associés aux équations étudiées.

1.2.1 Résultats d'existence connus pour l'équation de KdV

Résultats d'existence locale. Les premiers résultats d'existence locale dans l'espace de Sobolev $H^s(\mathbb{R})$ ont été prouvés pour $s \geq 2$ dans [91] et pour $s > \frac{3}{2}$ dans [63], par exemple.

Section 1.2. Problèmes de Cauchy

Dans [64], Kenig, Ponce et Vega démontrent l'existence locale dans $H^s(\mathbb{R})$ pour $s > \frac{3}{4}$. Plus précisément, ils démontrent que si $u_0 \in H^s(\mathbb{R})$ avec $s > \frac{3}{4}$, alors, il existe un temps δ ne dépendant que de $\|u_0\|_{H^s(\mathbb{R})}$ tel que

$$u \in \mathcal{C}([0, \delta], H^s(\mathbb{R})) \quad \text{et} \quad \partial_x u \in L^4([0, \delta], L^\infty(\mathbb{R})). \quad (1.6)$$

Pour cela, ils utilisent les effets régularisants suivants sur w , solution de $\partial_t w + \partial_x^{2p+1} w = 0$:

- l'estimation de Strichartz (ou effet régularisant global)

$$\left(\int_0^T \left\| |\partial_x^{\frac{2p-1}{4}} w(t, \cdot)| \right\|_{L^\infty(\mathbb{R})}^4 dt \right)^{\frac{1}{4}} \leq c \|w_0\|_{L^2(\mathbb{R})},$$

- l'effet régularisant local

$$\sup_{x \in \mathbb{R}} \int_0^T |\partial_x^p w(t, x)|^2 dt \leq c \|w_0\|_{L^2(\mathbb{R})}^2.$$

Dans les inégalités précédentes, c est une constante et $|\partial_x^{\frac{2p-1}{4}}|$ est défini par sa transformée de Fourier en espace (notée $\mathcal{F}$ ou $\widehat{\cdot}$) par $\mathcal{F} \left(|\partial_x^{\frac{2p-1}{4}} u \right) (t, \xi) := |\xi|^{\frac{2p-1}{4}} \widehat{u}(t, \xi)$.

Remarque 1.12. À la section 3.2, nous aurons besoin d'une borne $L^4([0, T], L^\infty(\mathbb{R}))$ sur la dérivée de u pour $s \geq \frac{3}{4}$ avec un temps T arbitraire (équation (1.6) mais avec T à la place de δ). Nous l'avons automatiquement pour $s \geq 1$, d'après [64] (voir le paragraphe suivant sur l'existence globale). Pour $s \in [\frac{3}{4}, 1[$, nous aurons besoin de descendre à $L^2(\mathbb{R})$ et d'utiliser l'existence locale dans $L^2(\mathbb{R})$, expliquée ci-dessous.

Les résultats d'existence locale pour des régularités plus faibles font intervenir d'autres outils comme les espaces de Bourgain $X^{s,b}$ [24], dont la norme est

$$\|u\|_{X^{s,b}} := \left(\int_{\mathbb{R}} \int_{\mathbb{R}} (1 + |\xi|)^{2s} (1 + |\tau - \xi^3|)^{2b} |\tilde{u}(\xi, \tau)|^2 d\xi d\tau \right)^{\frac{1}{2}},$$

avec $\tilde{u}$ la transformée de Fourier de u en espace et en temps. Dans [24], la question de l'existence locale dans $L^2(\mathbb{R})$ est prouvée : il existe δ un temps dépendant de la norme $\|u_0\|_{L^2(\mathbb{R})}$ tel que $u \in \mathcal{C}([0, \delta], L^2(\mathbb{R})) \cap X^{0,b}$ pour $b \in]\frac{1}{2}, 1[$.

Remarque 1.13. Dans toute la suite, nous nous restreindrons aux espaces de Sobolev $H^s(\mathbb{R})$ à indice s positif, mais il existe des résultats d'existence locale pour des régularités plus faibles que $L^2(\mathbb{R})$. Dans [66], Kenig, Ponce et Vega démontrent l'existence d'un temps δ ne dépendant que de $\|u_0\|_{H^s(\mathbb{R})}$ tel que $u \in \mathcal{C}([0, \delta], H^s(\mathbb{R})) \cap X^{s,b}$ pour $s \in]-\frac{3}{4}, 0]$ et $b \in]\frac{1}{2}, 1[$.



Résultats d'existence globale. Le but de l'existence globale est de montrer que si $u_0 \in H^s(\mathbb{R})$ alors $u \in \mathcal{C}([0, T], H^s(\mathbb{R}))$ pour un temps T arbitrairement grand.

Beaucoup de résultats d'existence globale se basent sur les existences locales et les quantités conservées. En effet, l'équation de KdV possède une infinité de lois de conservation [64]

$$\int_{\mathbb{R}} u^2(t, x) dx, \quad \int_{\mathbb{R}} \left((\partial_x u)^2 - \frac{1}{3} u^3 \right) (t, x) dx, \quad \int_{\mathbb{R}} \left((\partial_x^2 u)^2 - \frac{5}{3} u (\partial_x u)^2 + \frac{5}{36} u^4 \right) (t, x) dx \quad \dots$$

En utilisant ces lois de conservation et le fait que les temps d'existence locale δ ne dépendent que des normes $\|u_0\|_{H^s(\mathbb{R})}$, on peut itérer les résultats d'existence locale pour obtenir une existence sur tout $[0, T]$, avec T arbitrairement grand. L'existence globale est ainsi prouvée pour $s \geq 0$ dans [24].

Notons également que pour $s \geq \frac{3}{4} : \partial_x u \in L^4([0, T], L^\infty(\mathbb{R}))$, voir [64] et [24].

Remarque 1.14. Le caractère globalement bien posé est aussi connu pour des indices de Sobolev négatifs : pour $s > -\frac{3}{4}$ dans [33], par exemple.

Une synthèse des résultats d'existence locale et globale se trouve dans [99] ou [80].

1.2.2 Résultats d'existence connus pour le système $abcd$

Condition de signe sur les variables a, b, c, d . En plus de la condition (1.3) sur les coefficients a, b, c et d , seules deux configurations possibles assurent au système $abcd$ (linéarisé autour de l'état d'équilibre au repos) d'être linéairement bien posé

$$a \leq 0, c \leq 0, b \geq 0, d \geq 0, \quad (1.7)$$

$$\text{ou } a = c > 0, b \geq 0, d \geq 0, \quad (1.8)$$

voir par exemple Bona, Chen et Saut [17], Anh [4], ou Linares, Pilod et Saut [79]. À partir de maintenant, nous nous focalisons seulement sur la configuration (1.7). Le tableau 1.1 recense les équations classiques de la littérature correspondantes à cette configuration en fonction des termes de dispersion présents (dispersion de type BBM $\partial_x^2 \partial_t$ et de type KdV ∂_x^3).

	$b > 0, d > 0$	$b = 0, d > 0$	$b > 0, d = 0$	$b = 0, d = 0$
$a = c = 0$	BBM-BBM	syst. de Boussinesq	BBM-Burgers	syst. de Saint Venant ¹
$a < 0, c = 0$		KdV-BBM		syst. de Kaup
$a < 0, c < 0$				KdV-KdV
$a = 0, c < 0$	syst. de Bona-Smith		BBM-KdV	Euler-Korteweg ²

TABLE 1.1 – Cas particuliers avec $a \leq 0, c \leq 0, b \geq 0, d \geq 0$

Remarque 1.15. Le système $(abcd)$ n'est pas le seul couplage possible entre KdV et BBM, il existe d'autres «combinaisons» physiquement réalistes à commencer par l'équation de Korteweg-de Vries-Benjamin-Bona-Mahony (KdV-BBM)

$$(I - b\partial_x^2)\partial_t u + \partial_x u + a\partial_x^3 u + \partial_x \left(\frac{u^2}{2} \right) = 0, \quad (\text{KdV-BBM})$$

où $a \in \mathbb{R}$ et $b > 0$. À notre connaissance, cette équation a été introduite pour la première fois par Bona et Smith en 1975 [22] comme une régularisation de KdV. L'équation de (KdV-BBM) peut être réduite à l'équation (BBM) par un changement de variable : si u vérifie (KdV-BBM) avec $a = b$, alors v définie par

$$v(t, x) := u(t, x - t)$$

vérifie $\partial_t v + 2\partial_x v + \partial_x \left(\frac{v^2}{2} \right) - a\partial_x^2 \partial_t v = 0$ (qui est l'équation (BBM) avec terme de transport à vitesse 2).

1. une asymptotique différente $\frac{A}{H} = \mathcal{O}(1)$, pas de dispersion

2. avec un coefficient de capillarité constant (ce qui arrive par exemple en mécanique des fluides classique) [25]



Existence d'une solution au problème de Cauchy. La question de l'existence globale n'a, à notre connaissance, été prouvée que dans le cas classique du système de Boussinesq ($a = b = c = 0, d > 0$), étudié par Amick [3] et Schonbek [93] ainsi que dans le cas $b = d > 0, a \leq 0, c < 0$ (qui inclut le système de Bona et Smith, par exemple) avec données initiales petites [17, 23]. Pour les cas plus généraux, l'existence de solutions en temps long, c'est-à-dire de l'ordre de $\mathcal{O}(\frac{1}{\epsilon})$, a été traitée dans [27] pour la configuration (1.7) en excluant deux cas

$$\begin{cases} a = d = 0, c < 0, b > 0, \\ a = b = 0, c < 0, d > 0, \end{cases} \quad (1.9)$$

à l'aide d'estimations sur l'énergie

$$\begin{aligned} \mathcal{E}_s^{abcd}(\eta, u) := & \|\eta\|_{H^s(\mathbb{R})}^2 + \epsilon(b-c)\|\nabla\eta\|_{H^s(\mathbb{R})}^2 + \epsilon^2(-c)b\|\nabla^2\eta\|_{H^s(\mathbb{R})}^2 \\ & + \|u\|_{H^s(\mathbb{R})}^2 + \epsilon(d-a)\|\nabla u\|_{H^s(\mathbb{R})}^2 + \epsilon^2(-a)d\|\nabla^2u\|_{H^s(\mathbb{R})}^2, \end{aligned} \quad (1.10)$$

pour $s \in \mathbb{R}$. Nous rappelons ci-dessous le théorème d'existence de [27]. Soit $s \in \mathbb{R}$, considérons les indices suivants :

$$\begin{cases} s_{bc} = s + \operatorname{sgn}(b) - \operatorname{sgn}(c), \\ s_{ad} = s + \operatorname{sgn}(d) - \operatorname{sgn}(a), \end{cases} \quad (1.11)$$

où la fonction sgn est définie par :

$$\operatorname{sgn}(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & \text{if } x < 0. \end{cases}$$

Théorème 1.16. Soient $a, c \leq 0$ and $b, d \geq 0$ en excluant les deux cas (1.9), s un entier tel que

$$s > \frac{5}{2} - \operatorname{sgn}(b+d),$$

et s_{bc}, s_{ad} définis par (1.11). Supposons que $(\eta_0, u_0) \in H^{s_{bc}}(\mathbb{R}) \times H^{s_{ad}}(\mathbb{R})$ alors il existe un temps strictement positif $T = \mathcal{O}(\frac{1}{\epsilon})$ et une solution unique

$$(\eta, u) \in \mathcal{C}([0, T[, H^{s_{bc}}(\mathbb{R}) \times H^{s_{ad}}(\mathbb{R}))$$

au système (abcd).

1.3 Aspect numérique

Les trois premiers chapitres de cette thèse sont consacrés au calcul des solutions approchées des équations décrites ci-dessus.

1.3.1 Principaux résultats numériques sur les équations linéaires

Notations générales. Dans la suite, nous nous focaliserons essentiellement sur les méthodes de différences finies à un pas. Le pas de temps $\Delta t > 0$ et le pas d'espace $\Delta x > 0$ sont pris uniformes pour simplifier l'analyse. Nous notons $t^n = n\Delta t$ pour tout $n \in \llbracket 0, N \rrbracket$ avec $N = \lfloor \frac{T}{\Delta t} \rfloor$, où $\lfloor \cdot \rfloor$ représente la partie entière et $x_j = j\Delta x$ pour tout $j \in \mathbb{Z}$. La discrétisation du modèle consiste à calculer une solution numérique approchée v_j^n à chaque point du maillage $(t^n, x_j) := (n\Delta t, j\Delta x)$.

Définition 1.17. Soit u la solution exacte de l'EDP, et $\mathcal{F}_{\Delta t, \Delta x} v = 0$ avec $v = (v_j^n)_{(n,j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}}$, un schéma numérique discrétisant cette EDP.

- L'erreur de consistance ϵ_j^n est définie par

$$\epsilon_j^n := (\mathcal{F}_{\Delta t, \Delta x} u_\Delta)_j^n, \quad (n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}.$$

Ici, u_Δ représente une solution constante par morceaux sur chaque maille, construite à partir de u (sa construction sera détaillée à la définition 1.18).

- Le schéma est stable dans $\ell^2(\mathbb{Z})$ s'il existe une constante K , indépendante de Δt et Δx , telle que $\|v^n\|_{\ell^2(\mathbb{Z})} \leq K \|v^0\|_{\ell^2(\mathbb{Z})}, \forall n \in \llbracket 0, N \rrbracket$.
- La différence $v_j^n - (u_\Delta)_j^n$ est appelée l'erreur de convergence et nous la noterons e_j^n par la suite.
- Un schéma est convergent en norme $\ell^\infty(\llbracket 0, N \rrbracket, \ell^2(\mathbb{Z}))$ s'il existe p et q tels que

$$\|e^n\|_{\ell^\infty(\llbracket 0, N \rrbracket, \ell^2(\mathbb{Z}))} = \mathcal{O}(\Delta t^p + \Delta x^q).$$

Les puissances p et q sont appelées vitesses de convergence (ou ordre de convergence) du schéma.

La condition de Courant-Friedrichs-Lewy (condition CFL)

La stabilité n'est parfois assurée que sous une relation liant Δt et Δx . Une telle relation est appelée *condition de type Courant-Friedrichs-Lewy* et sera abrégée par la suite en condition de type CFL.

Nous utiliserons certaines simplifications dans les notations et définissons pour cela les opérateurs discrets décentré à droite, à gauche et centré suivants

$$(D_+ v)_j^n := \frac{v_{j+1}^n - v_j^n}{\Delta x}, \quad (D_- v)_j^n := \frac{v_j^n - v_{j-1}^n}{\Delta x}, \quad (D_c v)_j^n := \frac{v_{j+1}^n - v_{j-1}^n}{2\Delta x}. \quad (1.12)$$

Nous les appellerons respectivement *right winded scheme*, *left winded scheme* et *central scheme* en anglais. Ils constituent les « briques élémentaires » pour toute autre dérivation d'ordre élevée, cf remarque 1.22.



Objectifs et buts suivis. Notre étude numérique porte essentiellement sur l'équation de (KdV) (et ses variantes suivantes : Airy, (KdV-BBM)) ainsi que sur le système (abcd). Son objectif est double :

- déterminer l'ordre de convergence en espace et en temps de certains schémas aux différences finies pour ces équations
- quantifier cet ordre de convergence en fonction de la régularité de Sobolev de la donnée initiale.

Afin d'étudier d'un point de vue numérique nos schémas, nous nous plaçons dans un espace discret qui imite l'espace $L^2(\mathbb{R})$, l'espace noté par la suite $\ell_\Delta^2(\mathbb{Z})$ et dont la norme est définie par

$$\|v\|_{\ell_\Delta^2} := \left(\sum_{j \in \mathbb{Z}} \Delta x (v_j)^2 \right)^{\frac{1}{2}}. \quad (1.13)$$

Pour les suites numériques dépendant du temps et de l'espace, nous utilisons l'espace suivant $\ell^\infty(\llbracket 0, N \rrbracket; \ell_\Delta^2(\mathbb{Z}))$ dont la norme est

$$\|v\|_{\ell^\infty(\llbracket 0, N \rrbracket; \ell_\Delta^2(\mathbb{Z}))} := \sup_{n \in \llbracket 0, N \rrbracket} \left(\sum_{j \in \mathbb{Z}} \Delta x (v_j^n)^2 \right)^{\frac{1}{2}}. \quad (1.14)$$

Section 1.3. Aspect numérique

Grâce à ce poids Δx rajouté dans la norme discrète, nous avons

$$\|e\|_{\ell^\infty([0,N],\ell_\Delta^2(\mathbb{Z}))} = \|v - u_\Delta\|_{L^\infty([0,T],L^2(\mathbb{R}))}.$$

Il nous reste à préciser cette fonction u_Δ (déjà mentionnée par exemple à la définition 1.17) construite à partir de u et qui représente une fonction constante par morceaux sur chaque maille que nous comparerons à la solution numérique $(v_j^n)_{(n,j)}$. Nous construisons cette fonction u_Δ comme suit.

Définition 1.18. Soit u la solution exacte de l'équation considérée telle que $u|_{t=0} = u_0$, alors nous définissons la suite

$$\begin{cases} (u_\Delta)_j^n = \frac{1}{\Delta x [\inf(t^{n+1}, T) - t^n]} \int_{t^n}^{\inf(t^{n+1}, T)} \int_{x_j}^{x_{j+1}} u(s, y) dy ds, & \text{si } (n, j) \in \llbracket 1, N \rrbracket \times \mathbb{Z}, \\ (u_\Delta)_j^0 = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u_0(y) dy, & \text{si } j \in \mathbb{Z}. \end{cases} \quad (1.15)$$

À partir de la suite moyennée $([u_\Delta]_j^n)_{(n,j)}$, nous construisons la fonction constante par morceaux u_Δ telle que si $(t, x) \in [t^n, \inf(t^{n+1}, T)] \times [x_j, x_{j+1}[$, alors $u_\Delta(t, x) = (u_\Delta)_j^n$.

Remarque 1.19. • La moyenne en espace dans u_Δ facilite l'étude de l'erreur de consistance. En effet, le calcul de $\|e^n\|_{\ell_\Delta^2(\mathbb{Z})}$, à n fixé, fait intervenir $\sum_{j \in \mathbb{Z}} \Delta x |[u_\Delta]_j^n|^2$. Le fait d'avoir une intégrale en espace dans u_Δ assure la sommabilité de cette somme en faisant apparaître la norme de Sobolev $H^s(\mathbb{R})$ pour $s \in \mathbb{R}$ de u

$$\|u\|_{H^s(\mathbb{R})} = \left(\int_{\mathbb{R}} (1 + |\xi|^2)^s |\widehat{u}(\xi)|^2 d\xi \right)^{\frac{1}{2}}, \quad (1.16)$$

avec $\widehat{u}$ la transformée de Fourier de u .

• L'intégrale en temps ne sert que pour les équations non linéaires : elle permet au coefficient K de la définition de stabilité 1.17 de ne dépendre que de u_0 et non des u_Δ^n . Nous reviendrons sur cette explication à la remarque 3.45. Pour les équations linéaires du chapitre 2, nous n'intégrerons pas en temps dans les définitions u_Δ .

Pour répondre à notre deuxième objectif (quantifier l'ordre de convergence en fonction de la régularité), nous avons initialisé nos schémas avec des données u_0 parfois très peu régulières. Cependant, l'étude de la consistance se fait par développements de Taylor jusqu'à un certain ordre voulu, parfois plus élevé que la régularité de la donnée initiale u_0 . Lorsque la régularité de la donnée initiale est insuffisante pour mener à bien nos calculs, il est intéressant de rajouter une étape intermédiaire de régularisation de cette donnée initiale. Nous procèderons toujours de la même manière pour cette étape et garderons les mêmes notations. Lorsque u_0 n'a pas la régularité souhaitée, nous la régularisons par produit de convolution avec une suite régularisante $(\varphi^\delta)_{\delta > 0}$, construite comme suit.

Soit χ une fonction C^∞ telle que

- $0 \leq \chi \leq 1$,
- $\chi \equiv 1$ dans $[-\frac{1}{2}, \frac{1}{2}]$ et à support compris dans $[-1, 1]$,
- $\chi(-\xi) = \chi(\xi)$, $\forall \xi \in [-1, 1]$.

Soit φ tel que $\widehat{\varphi}(\xi) = \chi(\xi)$ (nous rappelons que $\widehat{\varphi}$ représente la transformée de Fourier de φ) et pour tout $\delta > 0$, nous définissons φ^δ tel que $\widehat{\varphi^\delta}(\xi) = \chi(\delta\xi)$, ce qui implique $\varphi^\delta = \frac{1}{\delta} \varphi(\frac{\cdot}{\delta})$.

Notations : Étapes de régularisation

- Nous notons u la solution exacte de l'EDP étudiée partant de la donnée initiale peu régulière u_0 .
- La solution numérique du schéma choisi sera notée $(v_j^n)_{(n,j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}}$.
- Soit u^δ la solution exacte de l'EDP avec $u_0^\delta := u_0 \star \varphi^\delta$ comme donnée initiale, où $\star$ représente le produit de convolution.
- Nous notons ensuite $((v^\delta)_j^n)_{(n,j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}}$ la solution numérique obtenue en appliquant le schéma à la donnée initiale $(u_0^\delta)_\Delta$:

$$v_j^0 = (u_0^\delta)_\Delta = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u_0 \star \varphi^\delta(y) dy.$$

Remarque 1.20. Cette régularisation est un argument classique déjà présent dans [43] par exemple.

La méthode consiste alors à régulariser la donnée initiale, à faire l'étude de convergence et ensuite à déterminer l'écart entre la solution régularisée et la solution partant d'une donnée peu régulière. Ce qui revient par inégalité triangulaire à

$$\begin{aligned} \|e\|_{\ell^\infty(\llbracket 0, N \rrbracket, \ell_\Delta^2(\mathbb{Z}))} &= \|u_\Delta - v\|_{\ell^\infty(\llbracket 0, N \rrbracket, \ell_\Delta^2(\mathbb{Z}))} \\ &\leq \|u_\Delta - (u^\delta)_\Delta\|_{\ell^\infty(\llbracket 0, N \rrbracket, \ell_\Delta^2(\mathbb{Z}))} + \|(u^\delta)_\Delta - v^\delta\|_{\ell^\infty(\llbracket 0, N \rrbracket, \ell_\Delta^2(\mathbb{Z}))} + \|v^\delta - v\|_{\ell^\infty(\llbracket 0, N \rrbracket, \ell_\Delta^2(\mathbb{Z}))}. \end{aligned}$$

Le premier et le dernier terme reviennent à étudier l'écart entre la solution (continue ou numérique) régularisée et la solution partant d'une donnée peu régulière. Ils se traitent respectivement par stabilité de l'EDP et stabilité du schéma numérique. Le terme central correspondant à l'erreur de convergence pour une donnée très régulière. Pour ce terme, les développements de Taylor de la partie sur la consistance ne sont pas problématiques.

Remarque 1.21. Notons que dans le cadre de la résolution numérique d'une équation, nous ne pouvons travailler qu'en domaine borné. Cette remarque sera par exemple à garder en mémoire au chapitre 3 où l'analyse théorique de convergence se fait sur tout $\mathbb{R}$ mais où les illustrations numériques ne concernent que des solutions périodiques.



Résultats obtenus sur les équations dispersives linéaires. Illustrons tout d'abord notre étude sur une généralisation de l'équation d'Airy : l'équation linéaire dispersive d'ordre $2p + 1$

$$\partial_t u(t, x) + \partial_x^{2p+1} u(t, x) = 0, \quad (t, x) \in [0, T] \times \mathbb{R},$$

avec $p \in \mathbb{N}$. Le cas $p = 1$ correspond à l'équation d'Airy et le cas $p = 0$ à l'équation de transport.

Le schéma numérique choisi consiste en un θ -schéma à $2p + 2$ ou $2p + 3$ points, selon que nous utilisons les schémas décentrés à droite, à gauche ou le schéma centré

$$\begin{cases} v_j^{n+1} + \theta \Delta t \left(D_\bullet^{2p+1} v \right)_j^{n+1} = v_j^n - (1 - \theta) \Delta t \left(D_\bullet^{2p+1} v \right)_j^n, & (j, n) \in \mathbb{Z} \times \llbracket 0, N \rrbracket, \end{cases} \quad (1.17a)$$

$$\begin{cases} v_j^0 = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u_0(y) dy, & j \in \mathbb{Z}, \end{cases} \quad (1.17b)$$

avec

$$\left(D_{\bullet}^{2p+1}v\right)_j^n = \left(D_+^{2p+1}v\right)_j^n = \sum_{k=0}^{2p+1} \frac{\binom{2p+1}{k}(-1)^k}{\Delta x^{2p+1}} v_{p-k+j+1}^n \text{ (schéma décentré à droite)} \quad (1.18a)$$

$$\text{ou} \quad \left(D_{\bullet}^{2p+1}v\right)_j^n = \left(D_-^{2p+1}v\right)_j^n = \sum_{k=0}^{2p+1} \frac{\binom{2p+1}{k}(-1)^k}{\Delta x^{2p+1}} v_{p-k+j}^n \text{ (schéma décentré à gauche)} \quad (1.18b)$$

$$\text{ou} \quad \left(D_{\bullet}^{2p+1}v\right)_j^n = \left(D_c^{2p+1}v\right)_j^n = \frac{1}{2} \left(D_+^{2p+1}v + D_-^{2p+1}v\right)_j^n \text{ (schéma centré)}. \quad (1.18c)$$

Remarque 1.22. Ces opérateurs de dérivation d'ordre élevé $D_{\bullet}^s$ peuvent se réécrire comme composition des opérateurs élémentaires $D_{\bullet}$ introduits en (1.12) :

$$\left\{ \begin{array}{l} D_+^{2p+1} := \underbrace{D_+ D_+ \dots D_+}_{p+1 \text{ fois}} \underbrace{D_- D_- \dots D_-}_{p \text{ fois}}, \\ D_-^{2p+1} := \underbrace{D_+ D_+ \dots D_+}_{p \text{ fois}} \underbrace{D_- D_- \dots D_-}_{p+1 \text{ fois}}, \\ D_c^{2p+1} := \underbrace{D_+ D_+ \dots D_+}_{p \text{ fois}} D_c \underbrace{D_- D_- \dots D_-}_{p \text{ fois}}, \\ D_c^{2p} := \underbrace{D_+ D_+ \dots D_+}_{p \text{ fois}} \underbrace{D_- D_- \dots D_-}_{p \text{ fois}}. \end{array} \right.$$

Nous montrons que la stabilité dépend de la parité de p .

Proposition : Stabilité des schémas décentrés

Les schémas décentrés à droite (1.17a) avec (1.18a) sont stables dès que p est impair sous condition CFL de type dispersif

$$\Delta t(1 - 2\theta) < \frac{\Delta x^{2p+1}}{2^{2p}}.$$

Alors que les schémas décentrés à gauche (1.17a) avec (1.18b) sont stables dès que p est pair sous la même condition CFL.

Dans les cas contraires, la stabilité est assurée sous une condition non classique

$$(p+1) \left(\frac{p^p}{(2C)^p (2p+1)^{2p+1}} \right)^{\frac{1}{p+1}} \Delta x^{\frac{(2p+1)}{p+1}} < \Delta t [2\theta - 1 + 2\theta^2 C \Delta t],$$

avec $C > 0$ une constante indépendante de Δt et Δx .

La condition $(p+1) \left(\frac{p^p}{(2C)^p (2p+1)^{2p+1}} \right)^{\frac{1}{p+1}} \Delta x^{\frac{(2p+1)}{p+1}} < \Delta t [2\theta - 1 + 2\theta^2 C \Delta t]$ est non classique dans le sens où elle impose une restriction sur Δx et non Δt . Par la suite, nous nous focaliserons sur les schémas stables avec une condition de type CFL. Lorsqu'un schéma décentré est utilisé, l'équation de transport sera donc discrétisée à gauche, l'équation d'Airy à droite, l'équation dispersive linéaire d'ordre 5 à gauche, et ainsi de suite. Ce résultat est à la base du schéma décentré choisi pour discrétiser l'équation de Korteweg-de Vries.

La stabilité ℓ^2 met aussi en lumière une condition de type Courant-Friedrichs-Lewy (ou CFL) dispersive très restrictive, nous avons besoin de pas de temps très petits difficilement réalisables en pratique. C'est pour contourner cette condition que nous utiliserons dans toutes les simulations numériques, des schémas implicites (donc sans condition de type CFL) ou des schémas de Crank-Nicolson.

Pour montrer la convergence, nous utilisons le théorème de Lax et Richtmyer [76]

Théorème 1.23. *Un schéma aux différences finies linéaire, stable et tel que son erreur de consistance vérifie la majoration $\|\epsilon^n\|_{\ell^2_\Delta(\mathbb{Z})} = \mathcal{O}(\Delta t^p + \Delta x^q)$, pour p et q deux entiers, est convergent dans $\ell^\infty(\llbracket 0, N \rrbracket, \ell^2_\Delta(\mathbb{Z}))$ au sens où*

$$\|v^n - (u_\Delta)^n\|_{\ell^\infty(\llbracket 0, N \rrbracket, \ell^2_\Delta(\mathbb{Z}))} = \mathcal{O}(\Delta t^p + \Delta x^q).$$

Remarque 1.24. *Le théorème de Lax-Richtmyer suppose le caractère linéaire du schéma numérique (et de l'EDP à résoudre), or l'équation de KdV et le système abcd ne le sont pas. Nous ne pourrions donc pas utiliser ce théorème pour (KdV) et (abcd).*

Une part importante de l'étude est accordée à la détermination de l'ordre de convergence en fonction de la régularité de Sobolev de la donnée initiale u_0 . Pour les équations linéaires dispersives d'ordre élevé, la régularité de Sobolev de la donnée initiale est transmise à la solution en tout temps, sans perte ni gain de régularité globale, cf lemme 2.5. Cela nous permet d'exprimer les estimations de consistance et de stabilité seulement en fonction des normes de Sobolev de u_0 (et non de la solution $u(t, \cdot)$ pour $t \neq 0$). Nous pouvons alors étudier l'évolution de ces estimations lorsque la donnée initiale est plus ou moins régulière.

Théorème : Ordre de convergence de (1.17a)-(1.17b) avec (1.18)

- Lorsque $u_0 \in H^s(\mathbb{R})$, les schémas sont convergents en temps à l'ordre $\frac{\min(s, 4p+2)}{4p+2}$ (ou $2 \frac{\min(s, 6p+3)}{6p+3}$ dans le cas du schéma de Crank-Nicolson).
- Pour l'ordre en espace, le schéma centré est convergent à l'ordre $2 \frac{\min(s, 2p+3)}{2p+3}$. Les schémas décentrés, s'ils sont stables, sont convergents à l'ordre $\frac{\min(s, 2p+2)}{2p+2}$ en espace.

Nous utilisons des formules de consistance pour les dérivées temporelles et spatiales précises à l'ordre 1 (voir 2 pour les schémas centrés ou de Crank-Nicolson) en espace et en temps. Il est donc naturel de retrouver un ordre de convergence égal à 1 (ou 2) pour le schéma numérique, lorsque les estimations de consistance ont un sens.

Pour des données initiales de faible régularité, il nous faut d'abord régulariser u_0 comme expliqué précédemment. Il y a alors une compétition entre la régularisation de u_0 et le coût de cette régularisation au niveau des dérivées. Il est donc nécessaire de trouver le juste milieu qui nous permettra de régulariser assez pour que les estimations de consistance aient un sens mais pas trop pour ne pas avoir de dérivées de la suite régularisante trop raides (et donc trop coûteuses). Cela aboutit à des ordres de convergences fractionnaires. Ces ordres semblent être les ordres optimaux de convergence d'après les résultats expérimentaux de la section 2.2.3.

1.3.2 Principaux résultats numériques sur l'équation de KdV

État de l'art sur l'aspect numérique. La méthode des différences finies est une approche très classique de résolution, et nous noterons comme précédemment Δx et Δt la grille discrète spatio-temporelle et $(v_j^n)_{(n,j)}$ la solution numérique. Un des premiers schémas aux différences finies utilisés pour résoudre (KdV) est le schéma saute-mouton explicite proposé par Zabusky et Kruskal dans [106]

$$\frac{v_j^{n+1} - v_j^{n-1}}{2\Delta t} + \frac{(v_{j+1}^n + v_j^n + v_{j-1}^n)(v_{j+1}^n - v_{j-1}^n)}{6\Delta x} + \frac{v_{j+2}^n - 2v_{j+1}^n + 2v_{j-1}^n - v_{j-2}^n}{2\Delta x^3} = 0,$$

avec des conditions aux bords périodiques. Ce schéma à trois niveaux en temps préserve la norme ℓ^1 discrète et est convergent à l'ordre deux en temps et en espace. Cependant, il nécessite une condition CFL de stabilité assez restrictive, cf [101]

$$\frac{\Delta t}{\Delta x} \left[|u_{\max}| + \frac{4}{\Delta x^2} \right] \leq 1.$$

Section 1.3. Aspect numérique

Notons qu'il est nécessaire d'initialiser ce schéma par une autre discrétisation : pour le premier pas de temps, nous pouvons par exemple remplacer la discrétisation centrée temporelle par $\frac{v_j^1 - v_j^0}{\Delta t}$.

Dans [101], Vliegenthart propose différents schémas numériques comme le schéma de Lax-Friedrichs

$$\frac{v_j^{n+1} - \frac{v_{j+1}^n + v_{j-1}^n}{2}}{\Delta t} + \frac{(v_{j+1}^n)^2 - (v_{j-1}^n)^2}{4\Delta x} + \frac{v_{j+2}^n - 2v_{j+1}^n + 2v_{j-1}^n - v_{j-2}^n}{2\Delta x^3} = 0.$$

Ce schéma dissipatif ne nécessite pas d'initialisation particulière comme celui de Zabusky et Kruskal. Néanmoins, il est convergent à l'ordre un en temps et deux en espace et requiert toujours une condition de stabilité restrictive $\Delta t = \mathcal{O}(\Delta x^3)$.

Greig et Morris dans [50] utilisent le schéma de Hopscotch suivant

$$\begin{cases} \frac{v_j^{n+1} - v_j^n}{\Delta t} + \frac{(v_{j+1}^n)^2 - (v_{j-1}^n)^2}{4\Delta x} + \frac{v_{j+2}^n - 2v_{j+1}^n + 2v_{j-1}^n - v_{j-2}^n}{2\Delta x^3} = 0, & \text{si } (j+n) \text{ est pair,} \\ \frac{v_j^{n+1} - v_j^n}{\Delta t} + \frac{(v_{j+1}^{n+1})^2 - (v_{j-1}^{n+1})^2}{4\Delta x} + \frac{v_{j+2}^{n+1} - 2v_{j+1}^{n+1} + 2v_{j-1}^{n+1} - v_{j-2}^{n+1}}{2\Delta x^3} = 0, & \text{si } (j+n) \text{ est impair.} \end{cases}$$

Comme nous pouvons le voir, dans un tel schéma un point sur deux au temps $n+1$ est déterminé de manière explicite et un point sur deux est déterminé implicitement. Ce schéma conservatif à deux niveaux en temps a une erreur de convergence en $\mathcal{O}(\Delta t^2 + \Delta x^2)$. La condition CFL de stabilité est ici $\frac{\Delta t}{\Delta x} \|u_{\max}\| - \frac{2}{\Delta x^2} \leq 1$.

Dans tous les exemples précédents, le pas de temps Δt doit respecter la condition dispersive restrictive $\Delta t = \mathcal{O}(\Delta x^3)$. Le prix à payer pour utiliser un schéma sans condition CFL trop forte est d'impliciter la dispersion, comme expliqué précédemment. Winther dans [103], par exemple, propose un schéma aux différences finies implicite à douze-points avec trois niveaux en temps, d'ordre deux en espace et en temps. Son schéma est basé sur une reformulation de l'équation de KdV en le système suivant

$$\begin{cases} \partial_t u - \partial_x w = 0, \\ w = -\partial_x^2 u - \frac{u^2}{2}. \end{cases}$$

Dans [98], Taha et Ablowitz utilisent le schéma suivant

$$\frac{v_j^{n+1} - v_j^n}{\Delta t} + \frac{v_{j+1}^{n+1}(v_j^n + v_{j+1}^n) - v_{j-1}^{n+1}(v_j^n + v_{j-1}^n)}{6\Delta x} + \frac{v_{j+2}^{n+1} - 2v_{j+1}^{n+1} + 2v_{j-1}^{n+1} - v_{j-2}^{n+1}}{2\Delta x^3} = 0,$$

qui est du premier ordre en temps et d'ordre deux en espace. Aucune condition restrictive de stabilité n'est nécessaire, mais ils doivent résoudre un système quasi-pentagonal à chaque pas de temps.

Plus récemment, Holden, Koley et Risebro dans [55] prouvent la convergence d'un schéma de Lax-Friedrichs avec la dispersion implicite sous la CFL $\Delta t = \mathcal{O}(\Delta x^{\frac{3}{2}})$. Ils s'affranchissent de la condition CFL dispersive en implicitant la dispersion mais n'exhibent pas d'ordre de convergence pour leur schéma. Ils prouvent également la convergence forte de la solution numérique vers la solution classique si $u_0 \in H^3(\mathbb{R})$ et une convergence faible dans $L^2([0, T]; L_{\text{loc}}^2(\mathbb{R}))$ si $u_0 \in L^2(\mathbb{R})$.

Remarque 1.25. Il est intéressant de remarquer qu'une des premières preuves d'existence et d'unicité de solutions pour (KdV) repose sur un schéma proche des différences finies, cf [96],

$$\frac{dv_j(t)}{dt} = v_j(t) \frac{v_{j+1}(t) - v_{j-1}(t)}{6\Delta x} + \frac{(v_{j+1}(t))^2 - (v_{j-1}(t))^2}{6\Delta x} + \delta \frac{v_{j+1}(t) - 3v_j(t) + 3v_{j-1}(t) - v_{j-2}(t)}{\Delta x^3}.$$

Mentionnons également d'autres méthodes numériques pour résoudre (KdV). Kanazawa, Matsuo et Yaguchi [61] utilisent le schéma aux différences finies compact suivant, combiné à une méthode discrète variationnelle (visant à imiter la formulation variationnelle dans un cadre discret) :

$$\frac{v_j^{n+1} - v_j^n}{\Delta t} = \delta_c^{<1>} \left(-\frac{(v_j^{n+1})^2 + v_j^{n+1}v_j^n + (v_j^n)^2}{6} - \left(\delta_c^{<1>}\right)^2 \left(\frac{v_j^{n+1} + v_j^n}{2}\right) \right).$$

L'opérateur $\delta_c^{<1>}$ correspond à l'équivalent discret compact de ∂_x et est déterminé par

$$\delta_c^{<1>} f_j + \alpha(\delta_c^{<1>} f_{j+1} + \delta_c^{<1>} f_{j-1}) + \beta(\delta_c^{<1>} f_{j+2} + \delta_c^{<1>} f_{j-2}) = a \frac{f_{j+1} - f_{j-1}}{2\Delta x} + b \frac{f_{j+2} - f_{j-2}}{4\Delta x} + c \frac{f_{j+3} - f_{j-3}}{6\Delta x},$$

pour tout vecteur $(f_j)_j$. Un jeu de données possible pour les paramètres sont $\alpha = \frac{1}{6}$, $\beta = 0$, $a = \frac{14}{9}$, $b = \frac{1}{9}$ et $c = 0$. Ce schéma conserve l'énergie discrète et est précis à l'ordre six. Li et Visbal dans [78] combinent, quant à eux, un schéma compact en espace à une méthode de Runge-Kutta d'ordre quatre en temps.

Comme expliqué précédemment, deux effets sont présents dans l'équation de KdV : le terme non linéaire de Burgers et le terme dispersif d'Airy. Holden, Karlsen, Risebro et Tao dans [54] proposent une méthode de *splitting d'opérateurs*. Cette méthode se base sur la forme particulière de (KdV) : $\partial_t u = \mathcal{L}_B u + \mathcal{L}_A u$, avec $\mathcal{L}_B$ le terme différentiel de type «Burgers» et $\mathcal{L}_A$ le terme différentiel de type «Airy». En notant $\Phi_B(t)u_0$, la solution de l'équation de Burgers partant de u_0 et $\Phi_A(t)u_0$ celle de l'équation d'Airy partant de u_0 , ils décomposent la solution de (KdV) partant de u_0 en

$$\begin{cases} u(n\Delta t) \approx (\Phi_B(\Delta t)\Phi_A(\Delta t))^n u_0, & \text{pour le splitting de Godunov,} \\ u(n\Delta t) \approx \left(\left(\Phi_B\left(\frac{\Delta t}{2}\right)\Phi_A\left(\frac{\Delta t}{2}\right) \right) \left(\Phi_A\left(\frac{\Delta t}{2}\right)\Phi_B\left(\frac{\Delta t}{2}\right) \right) \right)^n u_0, & \text{pour le splitting de Strang.} \end{cases}$$

Les auteurs prouvent que la formule de Godunov est précise à l'ordre 1 en temps pourvu que $u_0 \in H^s(\mathbb{R})$, avec $s \geq 5$ et que la formule de Strang est précise à l'ordre deux en temps pourvu que $u_0 \in H^s(\mathbb{R})$, avec $s \geq 17$ (ou pour une donnée initiale dans $H^{s+5}(\mathbb{R})$ avec $s \geq 1$ dans [56]). La difficulté pour ce schéma est de prouver que la solution de l'équation de Burgers n'explose pas en temps fini. En effet, la solution d'Airy est prise comme donnée initiale pour l'équation de Burgers, mais elle n'est pas forcément «bien préparée» et les oscillations de la solution d'Airy pourraient amener à la formation de chocs en temps fini. Le point essentiel repose sur le fait que la solution de (KdV) elle-même est bornée dans $H^s(\mathbb{R})$ en tout temps, à ce niveau de régularité.

Ascher et McLachlan dans [7] combinent un splitting de Strang à un schéma symplectique pour imiter d'un point de vue discret la structure hamiltonienne de l'équation de KdV. En effet, (KdV) est une équation hamiltonnienne d'Hamiltonien $\mathcal{H} = \int_{\mathbb{R}} (\partial_x u)^2(t, x) - \frac{u^3}{3}(t, x) dx$. Ils obtiennent un schéma qui nécessite seulement une faible CFL $\Delta t < \frac{\Delta x}{2|u_{\max}|}$ mais requiert la résolution d'un gros système à chaque pas de temps. Notons que le schéma compact de Kanazawa, Matsuo et Yaguchi [61] est lui aussi symplectique.

Pour finir, il existe une grande richesse de schémas de type *éléments finis*. Dans [8], Baker, Dougalis et Karakashian étudient l'approche classique de Galerkin pour l'équation de KdV avec condition initiale 1-périodique. La méthode de Galerkin consiste à créer une approximation de la solution en utilisant des fonctions discontinues, polynomiales par morceaux ou splines. À la fois la solution reconstruite et la fonction test sont donc discontinues. Ils utilisent des splines de degré $r - 1$ et combinent un schéma de Newton (pour la partie non-linéaire) à un schéma de Crank-Nicolson (pour la dérivée en temps). L'ordre de convergence qu'ils obtiennent est $\mathcal{O}(\Delta t^2 + \Delta x^r)$ à condition que la CFL suivante soit respectée $\Delta t < c\Delta x^{\frac{3}{4}}$ avec c suffisamment petit pour assurer la convergence de l'algorithme de Newton.

Pour obtenir des schémas avec un ordre plus élevé, Dougalis et Karakashian, cf [45], utilisent

Section 1.3. Aspect numérique

une méthode de Galerkin discrète (toujours pour l'équation de KdV avec donnée initiale périodique) basée sur une méthode de Runge-Kutta d'ordre trois ou quatre en temps et une méthode de Newton en espace. Leur erreur de convergence est en $\mathcal{O}(\Delta t^p + \Delta x^r)$ avec $p = 3$ ou 4 (selon la méthode de Runge-Kutta utilisée) et avec des splines de degré $r - 1$. La CFL est dans ce cas $\Delta t = \mathcal{O}(\Delta x)$ et u_0 doit être assez régulier.

Récemment, Bona *et al.*, dans [13], ont établi une méthode de Galerkin pour l'équation généralisée de KdV 1-périodique (dans cette équation, la non-linéarité est remplacée par u^{p+1}). Leur but est de proposer un schéma non dissipatif pour simuler de manière plus précise les ondes progressives (*traveling waves* en anglais) qui sont des solutions à énergie constante, et qui sont mal capturées par des schémas trop dissipatifs. Leur schéma est basé sur une projection non locale de la solution, consistante avec le terme dispersif de (KdV).



Résultats obtenus sur (KdV). Nous avons choisi de discrétiser l'équation de KdV par un schéma de Rusanov pour la partie non linéaire et un θ -schéma décentré à droite pour la partie linéaire ce qui donne le schéma numérique suivant, avec $\theta \in [0, 1]$,

$$\left\{ \begin{array}{l} \frac{v_j^{n+1} - v_j^n}{\Delta t} + D_c \left(\frac{v^2}{2} \right)_j^n + \theta D_+ D_+ D_- (v)_j^{n+1} + (1 - \theta) D_+ D_+ D_- (v)_j^n = \frac{\tau \Delta x}{2} D_+ D_- (v)_j^n, \\ (n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}, \end{array} \right. \quad (1.19a)$$

$$\left\{ \begin{array}{l} v_j^0 = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u_0(y) dy, \quad j \in \mathbb{Z} \text{ si } u_0 \in H^s(\mathbb{R}), \text{ avec } s \geq 6, \end{array} \right. \quad (1.19b)$$

$$\left\{ \begin{array}{l} v_j^0 = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u_0 \star \varphi^\delta(y) dy, \quad j \in \mathbb{Z} \text{ si } u_0 \in H^s(\mathbb{R}), \text{ avec } \frac{3}{4} \leq s < 6. \end{array} \right. \quad (1.19c)$$

La fonction φ^δ est la suite régularisante introduite page 14. Nous énonçons ci-dessous le principal résultat de convergence numérique obtenu sur l'équation de KdV.

Théorème 1.26. *Convergence pour une donnée initiale quelconque*

Supposons $u_0 \in H^s(\mathbb{R})$ avec $s \geq \frac{3}{4}$ et $T \geq 0$. Choisissons $\alpha_0 > 0$ et $\tau > 0$ tels que

$$\| (u_\Delta)^n \|_{\ell^\infty} + \alpha_0 = \sup_{j \in \mathbb{Z}} | (u_\Delta)_j^n | + \alpha_0 \leq \tau, \quad n \in \mathbb{N}. \quad (1.20)$$

Soit $\beta_0 \in]0, 1[$, il existe $\omega_0 > 0$ tel que

$$\Delta t \leq \omega_0, \quad \Delta x \leq \omega_0,$$

et

$$\left\{ \begin{array}{l} 4(1 - 2\theta) \frac{\Delta t}{\Delta x^3} \leq 1 - \beta_0, \\ \left[\tau + \frac{1}{2} \right] \frac{\Delta t}{\Delta x} \leq 1 - \beta_0. \end{array} \right. \quad (1.21)$$

Alors, le schéma aux différences finies (1.19a)-(1.19b) (si $s \geq 6$) ou (1.19a)-(1.19c) (sinon) avec le coefficient de Rusanov valant la constante τ choisie précédemment et les pas de temps et d'espace Δt et Δx satisfaisant la condition de Courant-Friedrichs-Lewy (condition CFL) (1.21), vérifie, pour tout $\eta \in]0, s - \frac{1}{2}]$,

$$\| e^n \|_{\ell^\infty(\llbracket 0, N \rrbracket, \ell_\Delta^2(\mathbb{Z}))} \leq \Theta_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}} \left(1 + \|u_0\|_{H^{\frac{1}{2} + \eta}}^2 \right) \Delta x^\eta, \quad (1.22)$$

- $q = \frac{s}{12 - 2s}$, si $\frac{3}{4} \leq s \leq 3$,
- $q = \frac{\min(s, 6)}{6}$, si $3 < s$.

La quantité $\Theta_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}}$ est définie par

$$\Theta_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}} = \begin{cases} \Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}} \left(\frac{\|u_0\|_{H^6}}{\tau + \frac{1}{2}} + \|u_0\|_{H^4} + \|u_0\|_{H^{\frac{3}{2}+\eta}} \|u_0\|_{H^1} \right), & \text{si } s \geq 6, \\ C \left[\Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}} + \exp \left(\frac{T^{\frac{3}{4}} C e^{\kappa T}}{2} \|u_0\|_{H^{\frac{3}{4}}} \right) \right] \left(\frac{1}{\tau + \frac{1}{2}} + 1 + \|u_0\|_{H^{\min(1,s)}} \right) \|u_0\|_{H^s}, & \text{si } \frac{3}{4} \leq s < 6, \end{cases}$$

où C est une constante, κ ne dépend que de $\|u_0\|_{L^2(\mathbb{R})}$ et $\Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}}$ est défini par

$$\Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}} = \exp \left(\frac{C}{2} (1 + \tau^2) \left(1 + \frac{\Delta t^2}{\Delta x^2} \right) \left(T + (T^{\frac{3}{4}} + T^{\frac{1}{2}}) \|u_0\|_{H^{\frac{3}{4}}} e^{\kappa T} \right) \right) C e^{\kappa T} \sqrt{T \left\{ 1 + \frac{\Delta t}{\Delta x} \right\}}.$$

Remarque 1.27. Dans la définition précédente de $\Theta_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}}$, le terme $\left(\frac{\|u_0\|_{H^6}}{\tau + \frac{1}{2}} + \|u_0\|_{H^4} + \|u_0\|_{H^{\frac{3}{2}+\eta}} \|u_0\|_{H^1} \right)$ pour $s \geq 6$ (et $\left(\frac{1}{\tau + \frac{1}{2}} + 1 + \|u_0\|_{H^{\min(1,s)}} \right) \|u_0\|_{H^s}$ pour $\frac{3}{4} \leq s < 6$) vient de la contribution de l'erreur de consistance, alors que $\Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}}$ est principalement dû à la stabilité du schéma. Le terme supplémentaire $\exp \left(\frac{T^{\frac{3}{4}} C e^{\kappa T}}{2} \|u_0\|_{H^{\frac{3}{4}}} \right)$ qui apparaît pour $\frac{3}{4} \leq s < 6$ provient de l'approximation de u par u^δ comme expliqué page 14.

Remarque 1.28. Les résultats précédents sont encore valables pour un pas de temps et un coefficient de Rusanov dépendants du temps Δt^n et τ^n .

Remarque 1.29. -Nous retrouvons donc dans le précédent théorème, les conditions CFL classiques de type hyperbolique pour un schéma de Crank-Nicolson ou un schéma implicite, ainsi que la condition dispersive de type Airy pour un schéma explicite.

-La condition classique $\tau \Delta t \leq \Delta x$ est appelée condition hyperbolique puisqu'elle implique le même scaling que celui utilisé dans les équations hyperboliques $\Delta t \leftrightarrow \Delta x$. En comparaison, la condition CFL $4[1 - 2\theta] \frac{\Delta t}{\Delta x^3} \leq 1 - \beta_0$ est appelée condition CFL dispersive à cause du scaling $\Delta t \leftrightarrow \Delta x^3$.

-L'hypothèse $[\tau + \frac{1}{2}] \frac{\Delta t}{\Delta x} \leq 1 - \beta_0$ semble être une hypothèse technique, qui semble pouvoir être remplacée par la condition CFL hyperbolique classique $\tau \Delta t \leq \Delta x$.

Pour démontrer ce théorème, nous suivons la méthode classique d'étude de consistance et de stabilité, cette méthode sera détaillée dans l'introduction 3.1 du chapitre 3. La partie stabilité semble être la plus délicate puisqu'il faut tenir compte simultanément des termes dispersifs et des termes hyperboliques de l'équation. Pour la traiter, nous transposons d'un point de vue discret l'inégalité de stabilité fort-faible des lois de conservation hyperbolique, pour plus de détails voir [39]. Cette inégalité s'écrit, pour r et s deux constantes

$$\|w(T) - u(T)\|_{L^2([-r, r])} \leq C(T, u) \|w_0 - u_0\|_{L^2([-r-sT, r+sT])},$$

avec w une solution faible et u la solution forte. L'idée est de remplacer la solution faible par la solution numérique.

Une part importante de l'étude porte aussi sur la quantification de l'ordre de convergence en fonction de la régularité de Sobolev de la donnée initiale. Lorsque cette donnée initiale n'est pas suffisamment régulière pour assurer les estimations de consistance, nous la régularisons au moyen d'une suite régularisante (comme expliqué page 14). Néanmoins, nous supposons toujours u_0 au moins dans $H^{\frac{3}{4}}(\mathbb{R})$. Il faut s'assurer que l'étape de stabilité est toujours valide, même avec l'étape de régularisation. Pour cela, nous tirons profit de la régularité minimale requise : puisque u_0 est au moins dans $H^{\frac{3}{4}}(\mathbb{R})$, la solution exacte sera globale sur $[0, T]$, restera en tout

Section 1.3. Aspect numérique

temps dans le même espace de Sobolev que u_0 et surtout $\partial_x u$ sera dans $L^4([0, T], L^\infty(\mathbb{R}))$, ce qui assure au coefficient $C(T, u)$ de l'unicité fort-faible d'être indépendant du paramètre de régularisation. La nécessité d'avoir un facteur $C(T, u)$ indépendant de la régularisation et intégrable en temps nous empêche d'initialiser le schéma par des données de régularités plus faibles que $H^{\frac{3}{4}}(\mathbb{R})$.

Remarque 1.30. Nous venons de voir que la partie stabilité du schéma utilisé pour (KdV) nécessitait une certaine régularité ($u_0 \in H^{\frac{3}{4}}(\mathbb{R})$). Lorsque nous majorons l'erreur, par inégalité triangulaire, comme détaillé page 14, il faudrait donc avoir une certaine régularité sur v^δ pour pouvoir majorer $\|v^\delta - v\|_{\ell^\infty([0, N], \ell_\Delta^2(\mathbb{Z}))}$, ce qui paraît complexe à démontrer. Nous contournons la difficulté en changeant l'initialisation du schéma et définissons l'erreur de convergence comme $\|e\|_{\ell^\infty([0, N], \ell_\Delta^2(\mathbb{Z}))} = \|u_\Delta - v^\delta\|_{\ell^\infty([0, N], \ell_\Delta^2(\mathbb{Z}))}$ (seulement dans les cas où u_0 est peu régulier) (voir théorème 1.26).

Numériquement, les ordres de convergence expérimentaux semblent coïncider avec les résultats théoriques (au moins jusqu'à une régularité $H^3(\mathbb{R})$, voir la section 3.7), il semblerait donc que l'ordre optimal du schéma soit bien $\Delta x^{\frac{\min(s, 6)}{6}}$. Pour des régularités plus faibles que $H^3(\mathbb{R})$, les résultats numériques semblent montrer que l'ordre théorique n'est pas optimal et que la régularisation qu'on a introduite tend à sous estimer l'ordre de convergence du schéma. Au vu de ces résultats expérimentaux, l'ordre optimal serait toujours $\Delta x^{\frac{\min(s, 6)}{6}}$ alors que nous trouvons théoriquement un ordre en $\Delta x^{\frac{s}{12-2s}}$.

1.3.3 Principaux résultats numériques sur le système $abcd$

État de l'art numérique. La résolution numérique du système $abcd$ a fait l'objet de nombreuses études, la première, à notre connaissance, étant le travail de Peregrine [89] sur le système classique de Boussinesq $a = b = c = 0$, $d = \frac{1}{3}$. Depuis, d'autres cas particuliers ont été traités. Par exemple, Bona et Chen dans [14] implémentent un algorithme numérique précis à l'ordre quatre en temps et en espace du système BBM-BBM (qui correspond aux paramètres $a = c = 0$, $b = d = 1/6$). Bona, Dougalis et Mitsotakis dans [19] et [20] proposent un schéma basé sur une discrétisation temporelle de type Runge-Kutta implicite et une discrétisation spatiale de type méthode de Galerkin avec splines périodiques pour le système KdV-KdV, $b = d = 0$, $a = c = 1/6$, avec conditions aux bords périodiques.

Ce choix de schéma basé sur une approximation de type Galerkin combinée à une méthode de Runge-Kutta a aussi été étudié pour une gamme plus large de choix des paramètres a, b, c, d par Antonopoulos, Dougalis et Mitsotakis dans [5]. Dans [6], les auteurs l'étendent à des conditions aux bords de type Dirichlet homogènes et des conditions aux bords réfléchissantes sur les systèmes de type Bona-Smith $a = 0$, $b = d > 0$, $c \leq 0$. Ils utilisent leur schéma pour étudier les interactions entre deux solitons.

Plus récemment, Bona et Chen [15] fournissent une preuve numérique d'explosion en temps fini pour le système BBM-BBM. Ce phénomène d'explosion semble se produire à la collision frontale de deux solitons particuliers de l'équation de BBM.



Résultats numériques obtenus sur (abcd). Nous avons suivi la même discrétisation temporelle et spatiale que pour l'équation de KdV. La partie non linéaire est ainsi discrétisée au moyen d'un schéma de Rusanov, (nous noterons par la suite τ_1 et τ_2 les deux coefficients de viscosité de Rusanov). Par contre, une discrétisation centrée pour la partie dispersive a été préférée à la discrétisation décentrée de l'équation de (KdV). Pour ne conserver qu'une condition CFL hyperbolique, nous avons implicité les termes de dispersion et de transport, ou utilisé un schéma de Crank-Nicolson pour certains cas.

Nous noterons $(\eta_j^n)_{(n,j)}$ et $(u_j^n)_{(n,j)}$ les solutions numériques du schéma suivant

$$\begin{cases} \frac{1}{\Delta t} (I - bD_+D_-) (\eta_j^{n+1} - \eta_j^n) + (I + aD_+D_-) \left(\theta D_c(u)_j^{n+1} + (1 - \theta) D_c(u)_j^n \right) + D_c(\eta u)_j^n \\ \quad = \frac{1}{2} (1 - \operatorname{sgn}(b)) \tau_1 \Delta x D_+D_- (\eta)_j^n, \\ \frac{1}{\Delta t} (I - dD_+D_-) (u_j^{n+1} - u_j^n) + (I + cD_+D_-) \left(\theta D_c(\eta)_j^{n+1} + (1 - \theta) D_c(\eta)_j^n \right) + \frac{1}{2} D_c(u^2)_j^n \\ \quad = \frac{1}{2} (1 - \operatorname{sgn}(d)) \tau_2 \Delta x D_+D_- (u)_j^n, \end{cases} \quad (1.23)$$

avec $n \geq 0$ et $j \in \mathbb{Z}$. Nous choisissons $\theta = 1$ si $bd = 0$ et $\theta = \frac{1}{2}$ si $b > 0$ et $d > 0$. Nous initialisons le schéma par la donnée initiale moyennée sur une maille en espace, comme précédemment,

$$\begin{cases} \eta_j^0 = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} \eta_0(y) dy, \\ u_j^0 = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u_0(y) dy, \end{cases} \quad j \in \mathbb{Z}. \quad (1.24)$$

Nous noterons

$$e_j^n = \eta_j^n - (\eta_\Delta)_j^n \quad \text{et} \quad f_j^n = u_j^n - (u_\Delta)_j^n$$

les deux erreurs de convergence, avec η_Δ et u_Δ définies par la définition 1.18.

Remarque 1.31. Une distinction se fera ici entre les valeurs de b et d . Deux cas se distinguent : si $b > 0$ et $d > 0$, la dispersion de type BBM assure, même au niveau discret une certaine régularité. La convergence aura lieu sans nécessiter de CFL ni de viscosités τ_1 et τ_2 . Par contre, dès que $bd = 0$, il faudra supposer une CFL hyperbolique et des viscosités de Rusanov τ_i non nulles.

Nous restreignons notre étude aux paramètres

$$a \leq 0, c \leq 0, b \geq 0, d \geq 0,$$

en excluant les cinq cas suivants :

$$\begin{cases} a = b = 0, d > 0, c < 0, \\ a = d = 0, b > 0, c < 0, \\ b = d = 0, c < 0, a < 0, \\ a = b = c = d = 0, \\ a = b = d = 0, c < 0. \end{cases} \quad (1.25)$$

Remarque 1.32. Cette restriction au niveau des paramètres permet d'assurer le caractère bien posé au niveau continu des équations, comme précisé en (1.7)-(1.8).

Théorème 1.33. Convergence du schéma (1.23)

Soit $a \leq 0, b \geq 0, c \leq 0$ et $d \geq 0$ en excluant les cas (1.25), considérons $s > 9$ si $bd = 0$ et $s > 6$ si $b > 0$ et $d > 0$, considérons $(\eta_0, u_0) \in H^{s_{bc}}(\mathbb{R}) \times H^{s_{ad}}(\mathbb{R})$ et $T > 0$ tel que $(\eta, u) \in \mathcal{C}([0, T], H^{s_{bc}}(\mathbb{R}) \times H^{s_{ad}}(\mathbb{R}))$ soit la solution sur $[0, T]$ du système (abcd) partant de la donnée initiale (η_0, u_0) , avec

$$\begin{cases} s_{bc} = s + \operatorname{sgn}(b) - \operatorname{sgn}(c), \\ s_{ad} = s + \operatorname{sgn}(d) - \operatorname{sgn}(a). \end{cases}$$

Choisissons $\alpha_0 > 0, \tau_1 > 0$ et $\tau_2 > 0$ tels que

$$(1 - \operatorname{sgn}(bd)) (\|u_\Delta^n\|_{\ell^\infty} + \alpha_0) \leq \tau_1 \quad \text{et} \quad (1 - \operatorname{sgn}(bd)) (\|u_\Delta^n\|_{\ell^\infty} + \alpha_0) \leq \tau_2, \quad n \in \llbracket 0, N \rrbracket.$$

Soit $\beta_0 \in]0, 1[$, il existe une constante $\omega_0 > 0$ (dépendant de $\alpha_0, \beta_0, \tau_1, \tau_2$ et $\sup_{t \in [0, T]} \|(\eta(t), u(t))\|_{H^{s_{bc}} \times H^{s_{ad}}}$)

telle que si le nombre de pas de temps $N \in \mathbb{N}$ et le pas d'espace Δx sont choisis tels que

$$\Delta t = T/N \leq \omega_0 \quad \text{et} \quad \Delta x \leq \omega_0$$

et

$$\max\{(1 - \operatorname{sgn}(b)) \tau_1, (1 - \operatorname{sgn}(d)) \tau_2\} \Delta t \leq (1 - \beta_0) \Delta x. \quad (1.26)$$

et si nous considérons le schéma numérique (1.23) avec les viscosités numériques τ_1 et τ_2 choisies précédemment, si nécessaire, et l'initialisation (1.24), alors ce schéma numérique est d'ordre un en temps et en espace. Plus précisément, les erreurs de convergence satisfont l'inégalité suivante

$$\sup_{n \in \llbracket 0, N \rrbracket} \mathcal{E}(e^n, f^n) \leq C_{abcd} \left\{ \Delta t^2 + \Delta x^2 \right\},$$

où C_{abcd} dépend des paramètres a, b, c, d , de $\sup_{t \in [0, T]} \|(\eta(t), u(t))\|_{H^{s_{bc}} \times H^{s_{ad}}}$ et de T et où $\mathcal{E}$ est l'énergie suivante

$$\begin{aligned} \mathcal{E}(e, f) := & \|e\|_{\ell_\Delta^2}^2 + (b - c) \|D_+ e\|_{\ell_\Delta^2}^2 + b(-c) \|D_+ D_- e\|_{\ell_\Delta^2}^2 \\ & + \|f\|_{\ell_\Delta^2}^2 + (d - a) \|D_+ f\|_{\ell_\Delta^2}^2 + d(-a) \|D_+ D_- f\|_{\ell_\Delta^2}^2. \end{aligned} \quad (1.27)$$

Remarque 1.34. Puisque nous avons choisi le schéma de Crank-Nicolson ($\theta = \frac{1}{2}$) lorsque $b > 0$ et $d > 0$, nous retrouvons même un ordre de convergence en espace égal à deux dans ce cas

$$\sup_{n \in \llbracket 0, N \rrbracket} \mathcal{E}(e^n, f^n) \leq C_{abcd} \left\{ \Delta t^2 + \Delta x^4 \right\}.$$

La méthode suivie pour déterminer ce résultat de convergence sur (abcd) consiste à transposer au cas d'un système la méthode suivie pour l'équation de (KdV), voir l'introduction 4.1 du chapitre 4. La stabilité est toujours déterminée à partir d'une inégalité de type inégalité de stabilité fort-faible obtenue par intégrations par parties discrètes. Nous établissons les estimations de stabilité directement sur un équivalent discret de l'énergie continue (1.10) : l'énergie (1.27). Nous retrouvons bien expérimentalement des ordres de convergence égaux à un lorsque les schémas sont discrétisés par des estimations consistantes à l'ordre un (voir la section 4.5).

1.4 Ondes progressives

Une technique classique pour analyser les phénomènes propagatifs est la recherche de certaines solutions particulières appelées *ondes progressives* (ou *traveling wave* en anglais), représentées par exemple en figures 1.4(a) et 1.4(b).

Définition 1.35. On dit que $v(t, x)$ est une onde progressive de vitesse $c \in \mathbb{R}$ joignant les états u_- et u_+ s'il existe un profil u tel que

$$v(t, x) := u(x - ct), \quad (t, x) \in \mathbb{R}^2$$

satisfaisant $\lim_{\xi \rightarrow \pm\infty} u(\xi) = u_\pm$ et $\lim_{\xi \rightarrow \pm\infty} u^{(i)}(\xi) = 0$ pour $i \in \{1, 2, 3, \dots\}$.

Remarque 1.36. Si la vitesse de propagation c est positive, l'onde se propage vers la droite, alors que pour une vitesse de propagation c négative, l'onde se propage vers la gauche.

La partie IV a pour but d'étudier l'existence d'ondes progressives pour l'équation (KdV-4), cas particulier de l'équation (KdV-KS).

1.4.1 Orbites hétéroclines, orbites homoclines

Lorsqu'une onde progressive v est solution d'une EDP (de dérivée en espace d'ordre n au plus), son profil u vérifie nécessairement une EDO (d'ordre n), que l'on peut réduire à un système

d'ordre 1 dans l'espace des phases $\{X = (u, u', u'', \dots, u^{(n-1)})^t, u \in \mathbb{R}\}$

$$\frac{d}{d\xi} X(\xi) = F(X(\xi)), \quad X \in \mathbb{R}^n, \quad (1.28)$$

la fonction F dépendant, entre autre, de la vitesse c et de la non linéarité.

Nous pouvons donc décrire une onde progressive à travers son profil, en prenant le point de vue d'un système dynamique. Nous rappelons tout d'abord l'orbite de X_0 : c'est l'ensemble

$$\{x \in \mathbb{R}^n, \text{ tel qu'il existe } \xi \in \mathbb{R}, \text{ tel que } X(\xi) = x, \text{ avec } X(0) = X_0\}.$$

Il existe des orbites particulières parmi lesquelles les orbites hétéroclines et homoclines.

Définition 1.37. Soit X la solution du système (1.28) partant de $X(0) = X_0$. S'il existe $X_{\pm}$, deux points d'équilibre du système (1.28) (i.e $F(X_{\pm}) = 0$) vérifiant $\lim_{\xi \rightarrow -\infty} X(\xi) = X_-$ et $\lim_{\xi \rightarrow +\infty} X(\xi) = X_+$, alors l'orbite de X_0 est dite

- hétérocline si $X_- \neq X_+$,
- homocline si $X_- = X_+$.

Une orbite hétérocline correspond donc à une onde progressive avec des états initial et final différents et une orbite homocline correspond à une onde solitaire (ou soliton). Nous avons représenté schématiquement cette correspondance à la figure 1.4, lorsque le plan des phases correspond à $\{(u, u'), u \in \mathbb{R}\}$.

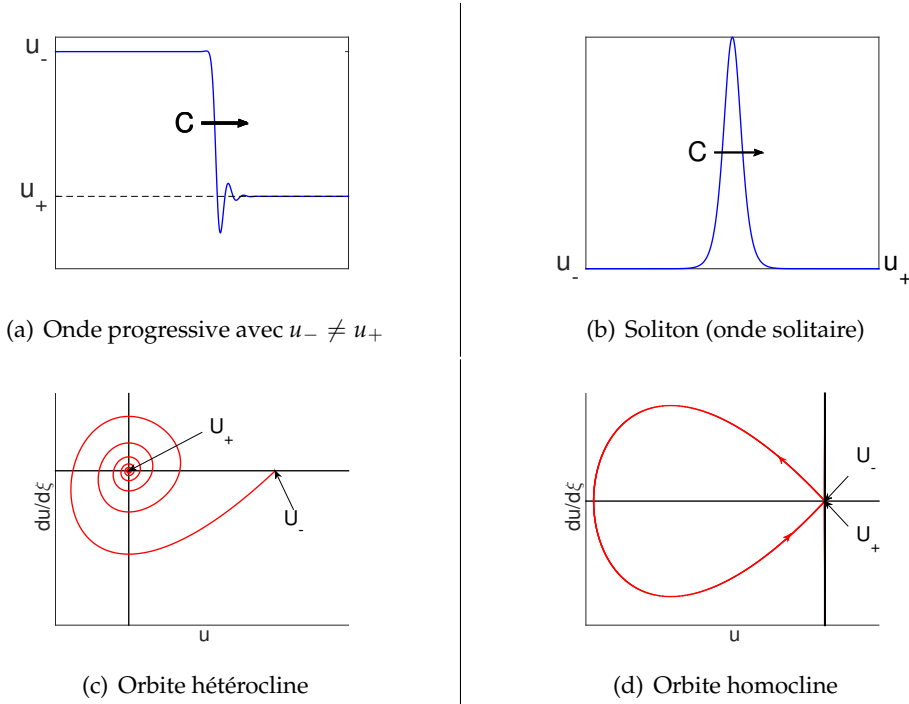


FIGURE 1.4 – Correspondance onde progressive avec $u_- \neq u_+$ / orbite hétérocline et onde solitaire / orbite homocline

1.4.2 Principaux résultats d'existence d'ondes progressives sur l'équation KdV-KS

Équations diffusives-dispersives. L'équation de KdV-KS tient son nom des deux équations sous-jacentes : l'équation dispersive (avec poids μ) de KdV

$$\partial_t v + \partial_x \left(\frac{v^2}{2} \right) = \mu \partial_x^3 v, \quad \mu \geq 0$$

Section 1.4. Ondes progressives

et l'équation diffusive de Kuramoto-Sivashinsky (KS)

$$\partial_t v + \partial_x \left(\frac{v^2}{2} \right) = \lambda \partial_x^2 v - \varsigma \partial_x^4 v, \quad \lambda \geq 0, \varsigma \geq 0. \quad (\text{KS})$$

L'équation (KS) apparaît dans de nombreux phénomènes physiques : elle a été établie initialement pour l'étude des petites instabilités thermiques diffusives dans les fronts de flammes [95] ou l'étude des systèmes de réaction-diffusion [70].

Remarque 1.38. Nous noterons KdV^- l'équation avec dispersion négative :

$$\partial_t v + \partial_x \left(\frac{v^2}{2} \right) = \mu \partial_x^3 v, \quad \mu > 0 \quad (\text{KdV}^-)$$

afin de la différencier de l'équation classique de KdV avec dispersion positive :

$$\partial_t v + \partial_x \left(\frac{v^2}{2} \right) + \mu \partial_x^3 v = 0, \quad \mu > 0.$$

Qu'il y ait ou non un poids μ sur la dispersion, nous appellerons l'équation KdV ou KdV^- , par abus de langage.

Selon les paramètres λ , μ et ς considérés dans l'équation, nous retrouvons d'autres équations classiques des écoulements de fluides. Un résumé des équations obtenues en fonction des paramètres est donné par le tableau 1.2.

- Ainsi, lorsqu'il n'y a ni dispersion, ni diffusion ($\lambda = \mu = \varsigma = 0$), il s'agit de l'équation de Burgers.
- Il est parfois intéressant de prendre en compte la viscosité du fluide considéré. Dans ce cas, nous obtenons l'équation de Burgers visqueux ($\lambda \neq 0, \mu = \varsigma = 0$).
- L'équation de KdV^- -Burgers, ou (KdV^- -B) par la suite, correspond à l'équation de KdV^- -KS sans diffusion d'ordre 4 ($\lambda, \mu > 0, \varsigma = 0$).

		dispersion	
		non ($\mu = 0$)	oui ($\mu \neq 0$)
diffusion	non ($\lambda=0$) ($\varsigma = 0$)	Burgers non visqueux	KdV^-
	ordre 2 ($\lambda \neq 0$) ($\varsigma = 0$)	Burgers visqueux	KdV^- -Burgers ¹
		KS	KdV^- -KS
	ordre 4 ($\lambda=0$) ($\varsigma \neq 0$)	Burgers avec diffusion d'ordre 4	KdV^- avec diffusion d'ordre 4 ¹

TABLE 1.2 – Équations diffusives-dispersives en fonction des valeurs des paramètres λ , μ et ς



1. À la partie IV, nous nous focaliserons exclusivement sur ces deux équations.

État de l'art sur les ondes progressives pour (KdV-KS). Certaines équations répertoriées dans le tableau 1.2 sont déjà connues pour posséder de telles ondes progressives. Par exemple,

$$v(t, x) := c - \frac{(u_- - u_+)}{2} \tanh \left(\frac{(u_- - u_+)}{4\lambda} (x - ct) \right), \quad \text{pour } (t, x) \in \mathbb{R}^2$$

est une onde progressive pour l'équation de Burgers visqueux ($\lambda > 0$ et $\mu = \zeta = 0$). Elle se déplace à vitesse $c = \frac{u_+ + u_-}{2}$ et permet de relier les états d'équilibre u_- et $u_+ < u_-$.

Pour l'équation de (KdV) ($\lambda = \zeta = 0$, $\mu = -1$), John Scott Russel observe en 1834 un phénomène de propagation de vague dans un canal à vitesse et amplitude constantes

«Cette onde continua sa marche dans le canal sans que sa forme et sa vitesse parussent s'altérer en rien.»²

Il s'agit de la première observation connue d'un *soliton* ou *onde solitaire*. L'expression du soliton de KdV est la suivante

$$u(t, x) = 3c \operatorname{sech}^2 \left(\frac{\sqrt{c}}{2} (x - ct) \right)$$

avec $c > 0$, la vitesse de propagation. On voit que la vitesse de déplacement du soliton dépend directement de son amplitude : plus l'amplitude est grande et plus la vitesse sera importante.

Remarque 1.39. Pour ramener (KdV⁻) avec dispersion à poids, à la forme classique de l'équation de KdV : $\partial_t u + \partial_x \left(\frac{u^2}{2} \right) + \partial_x^3 u = 0$, il suffit de poser le changement de variable

$$u(t, x) = -\frac{1}{\mu^{\frac{1}{5}}} v(-\mu^{\frac{1}{5}} t, \mu^{\frac{2}{5}} x).$$

Ainsi, le soliton de (KdV⁻) s'écrit

$$v(t, x) = -3\sigma \operatorname{sech}^2 \left(\frac{\sqrt{\sigma}}{2\sqrt{\mu}} (x + \sigma t) \right)$$

avec $\sigma = \mu^{\frac{1}{5}} c > 0$, et on retrouve un soliton se propageant vers la gauche.

Remarque 1.40. La notion de soliton s'est, depuis, généralisée à d'autres domaines que l'hydrodynamique comme l'optique (dans les fibres optiques), la physique des plasmas, la chimie des matériaux, l'acoustique... Les propriétés d'un soliton sont intéressantes, par exemple lorsque deux solitons se croisent, il n'y a aucun changement de leur direction, de leur vitesse et de leur amplitude respectives. L'apparition de soliton ne peut avoir lieu que par l'équilibre qui se réalise entre terme(s) non-linéaire(s) et terme(s) dispersif(s). L'équation de KdV apparaît donc comme un «modèle jouet» afin d'étudier et de mieux comprendre ces phénomènes non linéaires localisés.

De nombreuses études ont été menées afin de montrer l'existence d'ondes progressives pour certaines valeurs des paramètres λ , μ et ζ . Par exemple, une des premières études, menée par Mock en 1976 [84], concerne l'équation

$$\partial_t u + \partial_x (f(u)) = -\zeta \partial_x^4 u, \quad (1.29)$$

avec f lipschitz. Mock s'est inspiré de l'EDO (1.28) dans l'espace des phases pour démontrer par une méthode de tir que l'on peut joindre les deux états u_- et u_+ prescrits, sous certaines conditions de signe sur F .

2. M.H. Darcy and M.H. Bazin. *Recherches Hydrauliques*. Deuxième partie, Paris, Imprimerie Impériale, p.9, 1865

Section 1.4. Ondes progressives

Remarque 1.41. Notons que le résultat principal de [84] est, en fait, de démontrer un résultat d'existence de solutions faibles pour le système hyperbolique de loi de conservation $\partial_t u + \partial_x(f(u)) = 0$. Ces solutions faibles seront obtenues par passage à la limite $\zeta \rightarrow 0$ sur les solutions de (1.29). Cette technique de viscosité évanescence est classique, voir [94] par exemple.

L'existence d'ondes progressives pour l'équation de KdV-Burgers : $\partial_t u + \partial_x \left(\frac{u^2}{2} \right) = \lambda \partial_x^2 u - \mu \partial_x^3 u$ dans le cas particulier $\mu = 1$ a été étudiée dans [59]. Il y est démontré que l'onde solitaire (pour $\lambda = 0$) devient progressivement une onde progressive oscillante (pour de petits λ)

Bona et Schonbek, dans [21], généralisent l'étude précédente au cas λ et μ quelconques. En étudiant «géométriquement» le champ de vecteur associé à l'EDO de KdV-Burgers, ils démontrent l'existence d'ondes progressives oscillantes si la dispersion est plus importante que la diffusion $\lambda^2 < 4\mu(u_- - c)$ et d'ondes progressives monotones pour une forte diffusion $\lambda^2 \geq 4\mu(u_- - c)$. Ils exhibent alors le bon scaling diffusif-dispersif : $\lambda^2 = \mathcal{O}(\mu)$. Le comportement des ondes progressives est aussi détaillé asymptotiquement lorsque $\lambda \rightarrow 0$ ou $\mu \rightarrow 0$: sans dispersion l'onde progressive tend vers l'onde progressive de l'équation de Burgers visqueuse alors que sans diffusion, l'onde progressive tend vers le soliton de (KdV), et nous retrouvons dans ce dernier cas les résultats de [59].

Mentionnons également les travaux de Levy et Shearer [77] qui s'intéressent aux ondes progressives pour un système formé de l'équation (KdV-KS) couplée avec une équation sur la concentration d'un tensioactif ainsi que les travaux de Mansour [82] qui démontre la persistance de l'onde solitaire de KdV pour une équation (KdV-KS) avec des coefficients $\lambda, \mu, \zeta \ll 1$.



Résultats obtenus sur (KdV-KS). Nous avons démontré l'existence d'ondes progressives de faible amplitude pour l'équation (KdV-KS) avec non linéarité générale lorsque le nombre de Reynold converge vers le nombre de Reynold critique R^* . Pour cette asymptotique, l'équation est dégénérée (*i.e* $\lambda = 0$) et correspond à l'équation de (KdV-4) avec non linéarité générale :

$$\partial_t u + \partial_x(f(u)) = \mu \partial_x^3 u - \zeta \partial_x^4 u, \quad (1.30)$$

avec f une fonction convexe telle que $f(0) = f'(0) = 0$ et $f''(0) > 0$. Lorsque $f(\xi) = \frac{\xi^2}{2}$, nous retrouvons (KdV-4).

Dans notre cas particulier de l'équation (1.30), nous introduisons des profils d'onde remis à l'échelle afin que les coefficients de l'EDO du plan de phase soient tous de l'ordre de l'unité

$$v(t, x) = u \left(\frac{x - ct}{\zeta^{\frac{1}{3}}} \right).$$

Ceci nous permet de connaître les bonnes échelles entre les paramètres : $\mu = \mathcal{O}(\zeta^{\frac{2}{3}})$.

Nous nous focalisons surtout sur des ondes progressives avec états initial et final différents, ce qui revient à prouver l'existence d'orbites hétéroclines dans l'espace des phases.

Théorème 1.42. Onde progressive pour (1.30)

Supposons f convexe telle que $f(0) = f'(0) = 0$ et $f''(0) > 0$. dans (1.30). Il existe $\eta > 0$ tel que pour $u_+ \in]-\eta, 0[$, il existe une onde progressive (de petite amplitude) joignant 0 à u_+ à vitesse $c = \frac{f(u_+)}{u_+}$, solution de (1.30).

Remarque 1.43. Nous avons fixé $u_- = 0$ dans le précédent théorème sans perte de généralité, puisqu'il suffit de considérer $f(\xi) = g(\xi + u_-) - g(u_-) - g'(u_-)\xi$ avec une non linéarité g quelconque pour se ramener au cas où $u_- = 0$, cf remarque 5.2.

1.5 Plan de la thèse

Le manuscrit se découpe en cinq chapitres détaillés ci-dessous.

Chapitre 2. Nous amorçons l'étude numérique de l'équation de KdV en étudiant tout d'abord l'équation dispersive linéaire sous jacente : l'équation d'Airy

$$\partial_t u + \partial_x^3 u = 0$$

qui est une linéarisation de l'équation de KdV autour de zéro. Nous étudions dans ce chapitre la stabilité et la convergence de schémas aux différences finies pour l'équation d'Airy mais aussi pour une généralisation de cette équation : les équations dispersives linéaires de dérivée spatiale d'ordre impair

$$\partial_t u + \partial_x^{2p+1} u = 0, \quad p \in \mathbb{N}.$$

Cette classe d'équations regroupe l'équation de transport à vitesse un ($p = 0$) par exemple et bien sûr l'équation d'Airy ($p = 1$). Nous avons établi que la stabilité est assurée par un schéma décentré à gauche pour p pair (l'équation de transport par exemple) et à droite pour p impair (l'équation d'Airy par exemple).

L'étude est généralisée au cas des équations aux dérivées d'ordre pair en espace quelconque : les équations polyharmoniques

$$\partial_t u + (-1)^p \partial_x^{2p} u = 0, \quad p \in \mathbb{N}.$$

Ce chapitre est partiellement issu de l'acte de conférence [34].



Chapitre 3. Dans ce chapitre, nous discrétisons l'équation (KdV) par le schéma aux différences finies (1.19a)-(1.19b) ou (1.19a)-(1.19c). La convergence du schéma par rapport à la régularité de Sobolev de la donnée initiale (théorème 1.26) est démontrée et des simulations numériques sont effectuées afin de montrer que ces ordres de convergence semblent optimaux si $u_0 \in H^s(\mathbb{R})$, avec $s \geq 3$).

À cause de la présence des termes non linéaire et dispersif, la stabilité du schéma est prouvée par une méthode directe reposant sur des intégrations par parties discrètes. Une grande connaissance des opérateurs de dérivations élémentaires D_+ , D_- et D_c définis en (1.12) est donc nécessaire et une synthèse des différents calculs de normes discrètes, intégrations par parties discrètes et dérivations de produits discrets sont résumés dans ce chapitre.

Nous remarquons que cette connaissance des opérateurs élémentaires nous permet d'exhiber explicitement le flux numérique associé à l'entropie quadratique $\frac{u^2}{2}$ pour l'équation de Burgers à la fin de ce chapitre.

Certains résultats de ce chapitre se trouvent dans [36].



Chapitre 4. Ce chapitre est consacré à l'étude numérique du système (abcd). Nous choisissons d'étudier au préalable l'équation mixte de KdV-BBM

$$(I - b\partial_x^2) \partial_t u + \partial_x \left(\frac{u^2}{2} \right) + (I + a\partial_x^2) \partial_x u = 0.$$

Pour $a = 0$, nous retrouvons l'équation de (BBM) (avec un terme de transport) et pour $b = 0$, nous retrouvons l'équation de (KdV) (avec un terme de transport). Nous avons adapté l'étude de

Section 1.5. Plan de la thèse

stabilité et de convergence faite sur l'équation de KdV à cette équation mixte. Nous retrouvons des conditions de type CFL hyperboliques ou dispersives selon les valeurs des paramètres a et b , en utilisant un schéma décentré à droite ou à gauche selon le signe de a .

Nous généralisons l'étude numérique au système (abcd) et démontrons la décroissance de certaines énergies dans des espaces $H^1(\mathbb{R})$ ou $H^2(\mathbb{R})$ discrets selon les valeurs des paramètres a, b, c et d (cf théorème 1.33). Les simulations numériques semblent valider les résultats théoriques.

Il s'agit d'un travail en collaboration avec Cosmin Burtea [28].



Chapitre 5. L'existence d'ondes progressives pour l'équation de (KdV-KS) fait l'objet de ce chapitre. En utilisant les outils des systèmes dynamiques (théorème de la variété centrale, fonction de Lyapunov, intégrale de Melnikov...), nous montrons l'existence d'ondes progressives de petite amplitude pour l'équation de (KdV-4) généralisée, c'est-à-dire le théorème 1.42 pour l'équation (1.30).

Nous retrouvons aussi les résultats de Bona et Schonbek [21] pour l'équation de KdV-Burgers $\varsigma = 0$ (onde monotone lorsque la dispersion est plus faible que la diffusion et onde oscillante sinon) par une analogie intéressante avec l'équation de Fisher-KPP.



Chapitre 6. Ce chapitre, très prospectif, constitue une introduction préliminaire à l'étude de stabilité des ondes progressives pour l'équation de (KdV-KS). En calculant numériquement la fonction d'Evans sur l'axe réel $\mathbb{R}^+$, sur l'axe imaginaire pur $i\mathbb{R}$ ainsi que sur un voisinage de 0 dans $\mathbb{C}$, nous conjecturons la stabilité spectrale de telles ondes progressives.



DEUXIÈME PARTIE

Étude des équations linéaires de dérivées d'ordre élevé

CHAPITRE 2

Numerical analysis of high-order linear equations



We compute the rate of convergence of right-winded, left-winded and central finite difference θ -schemes for linear PDEs with an arbitrary odd order spatial derivative term

$$\partial_t u(t, x) + \partial_x^{2p+1} u(t, x) = 0, \quad (t, x) \in [0, T] \times \mathbb{R},$$

with $p \in \mathbb{N}$. We prove convergence of the first or second order for smooth initial data, and fractional order for any initial data. The numerical results match the theoretical orders proved. We finally mimic the convergence study for high-order polyharmonic heat equations

$$\partial_t u(t, x) + (-1)^p \partial_x^{2p} u(t, x) = 0, \quad (t, x) \in [0, T] \times \mathbb{R}.$$

In contrast, we consider only a central numerical scheme.

Summary

2.1	Introduction (en français)	35
2.2	High-order linear dispersive equations	38
2.2.1	Order of accuracy for an initial datum in $H^{4p+2}(\mathbb{R})$	38
2.2.1.1	A review of divided differences	38
2.2.1.2	Consistency estimate	40
2.2.1.3	Stability	42
2.2.1.4	Error estimates	47
2.2.2	Less smooth initial data	48
2.2.2.1	A regularization of u_0	48
2.2.2.2	Initial datum in $H^s(\mathbb{R})$ with $s \geq 2p + 2$	49
2.2.2.3	Initial datum in $H^s(\mathbb{R})$ with $s \geq 0$	51
2.2.3	Numerical results	52
2.3	High-order polyharmonic heat equations	54
2.3.1	For an initial datum in $H^{4p}(\mathbb{R})$	54
2.3.1.1	Consistency estimate	54
2.3.1.2	Stability	56
2.3.1.3	Convergence error	57
2.3.2	For a less smooth initial datum	57
2.3.3	Numerical results	59

2.1 Introduction (en français)



ce chapitre est consacré à l'analyse numérique des équations aux dérivées partielles linéaires d'ordre $\mathcal{P}$ quelconque en espace. Nous restreignons l'étude à la dimension 1 d'espace sans terme de bord, le problème de Cauchy associé à ces EDP s'écrit alors sous la forme

$$\begin{cases} \partial_t u(t, x) + (-1)^{s(\mathcal{P})} \partial_x^{\mathcal{P}} u(t, x) = 0, & (t, x) \in [0, T] \times \mathbb{R}, \\ u|_{t=0} = u_0, & x \in \mathbb{R}, \end{cases} \quad (2.1)$$

avec $\mathcal{P} \in \mathbb{N}$ et

$$s(\mathcal{P}) = \left(1 - \left(\mathcal{P} - 2 \left\lfloor \frac{\mathcal{P}}{2} \right\rfloor \right) \right) \left\lfloor \frac{\mathcal{P}}{2} \right\rfloor$$

où $\lfloor \cdot \rfloor$ est la partie entière. La forme générique de ces équations permet, en réalité, de traiter des EDP ayant des caractéristiques, des solutions et des comportements très différents. Citons quelques cas particuliers connus :

- équation de transport $\partial_t u(t, x) + \partial_x u(t, x) = 0$, $(t, x) \in [0, T] \times \mathbb{R}$,
- équation de la chaleur $\partial_t u(t, x) - \partial_x^2 u(t, x) = 0$, $(t, x) \in [0, T] \times \mathbb{R}$,
- équation d'Airy $\partial_t u(t, x) + \partial_x^3 u(t, x) = 0$, $(t, x) \in [0, T] \times \mathbb{R}$,
- équation diffusive d'ordre quatre $\partial_t u(t, x) + \partial_x^4 u(t, x) = 0$, $(t, x) \in [0, T] \times \mathbb{R}$.



Parité du paramètre $\mathcal{P}$. Selon la parité du paramètre $\mathcal{P}$, nous pouvons regrouper l'ensemble de ces équations en deux classes aux propriétés distinctes. Lorsque $\mathcal{P}$ est impair, $\mathcal{P} = 2p + 1$, l'équation (2.1) prend la forme suivante

$$\partial_t u(t, x) + \partial_x^{2p+1} u(t, x) = 0, \quad (t, x) \in [0, T] \times \mathbb{R}. \quad (2.2)$$

Nous l'appellerons équation linéaire dispersive d'ordre $2p + 1$. Le cas $p = 0$ est particulier puisqu'il n'a pas d'effet dispersif.

Lorsque $\mathcal{P}$ est pair, $\mathcal{P} = 2p$, les équations sont paraboliques et de type diffusif

$$\partial_t u(t, x) + (-1)^p \partial_x^{2p} u(t, x) = 0, \quad (t, x) \in [0, T] \times \mathbb{R}. \quad (2.3)$$

Dans toute la suite, nous les appellerons équations de la chaleur polyharmonique d'ordre $2p$. Ces équations ont la propriété de régulariser instantanément les solutions, comme l'illustre la proposition suivante.

Proposition 2.1. Soit u solution de (2.3). Alors, pour toute donnée initiale $u_0 \in L^2(\mathbb{R})$, u est dans $C^\infty([t_0, +\infty[, \mathbb{R})$, pour tout $t_0 > 0$.

La distinction des équations selon la parité de $\mathcal{P}$ est aussi justifiée d'un point de vue plus numérique. En effet, lorsque l'équation est discrétisée spatialement par des différences finies d'ordre 1 ou 2, les dérivées d'ordre pair ∂_x^{2p} sont généralement discrétisées en utilisant des formules centrées, alors que les dérivées spatiales d'ordre impair ∂_x^{2p+1} peuvent faire appel à des formules décentrées à gauche, à droite ou centrées.

Section 2.1. Introduction (en français)

Remarque 2.2. La présence du facteur $(-1)^p$ dans les équations où $\mathcal{P}$ est pair, équations (2.3), est primordiale pour assurer l'existence ou la non-explosion des solutions. En effet, sans ce coefficient, ou avec $(-1)^{p+1}$ au lieu de $(-1)^p$, l'EDP (2.3) est équivalente au problème de l'équation de la chaleur polyharmonique rétrograde, qui est un problème mal posé dû au caractère irréversible de la diffusion thermique. Il n'en est pas de même pour les équations où $\mathcal{P}$ est impair. Pour l'équation de transport par exemple, le signe de la vitesse de transport n'a aucune influence sur l'existence ou non d'une solution. Il s'agit d'une équation réversible (et le paramètre $(-1)^p$ est, dans ce cas, inutile).



Schémas numériques utilisés. Nous rappelons ici le schéma numérique utilisé, avec $t^n = n\Delta t$ pour tout $n \in \llbracket 0, N \rrbracket$ où $N = \lfloor \frac{T}{\Delta t} \rfloor$, $x_j = j\Delta x$ pour tout $j \in \mathbb{Z}$ et $(v_j^n)_{(j,n) \in \llbracket 0, N \rrbracket \times \mathbb{Z}}$ la solution numérique. Dans le cas $\mathcal{P}$ impair, nous avons testé les schémas décentrés à droite et à gauche à $2p+2$ points, ainsi que le schéma centré à $2p+3$ points et dans le cas $\mathcal{P}$ pair, nous avons testé le schéma centré à $2p+1$ points

$$\text{si } \mathcal{P} = 2p+1 \begin{cases} v_j^{n+1} + \theta \Delta t \left(D_{\bullet}^{2p+1} v \right)_j^{n+1} = v_j^n - (1-\theta) \Delta t \left(D_{\bullet}^{2p+1} v \right)_j^n, & (j, n) \in \mathbb{Z} \times \llbracket 0, N \rrbracket, \\ v_j^0 = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u_0(y) dy, & j \in \mathbb{Z}, \end{cases} \quad (2.4a)$$

avec

$$\left(D_{\bullet}^{2p+1} v \right)_j^n = \left(D_+^{2p+1} v \right)_j^n = \sum_{k=0}^{2p+1} \frac{\binom{2p+1}{k} (-1)^k}{\Delta x^{2p+1}} v_{p-k+j+1}^n \quad (\text{schéma décentré à droite}) \quad (2.5a)$$

$$\text{ou} \quad \left(D_{\bullet}^{2p+1} v \right)_j^n = \left(D_-^{2p+1} v \right)_j^n = \sum_{k=0}^{2p+1} \frac{\binom{2p+1}{k} (-1)^k}{\Delta x^{2p+1}} v_{p-k+j}^n \quad (\text{schéma décentré à gauche}), \quad (2.5b)$$

$$\text{ou} \quad \left(D_{\bullet}^{2p+1} v \right)_j^n = \left(D_c^{2p+1} v \right)_j^n = \frac{1}{2} \left(D_+^{2p+1} v + D_-^{2p+1} v \right)_j^n \quad (\text{schéma centré}), \quad (2.5c)$$

et

$$\text{si } \mathcal{P} = 2p \begin{cases} v_j^{n+1} + \theta (-1)^p \Delta t \left(D_c^{2p} v \right)_j^{n+1} = v_j^n - (1-\theta) (-1)^p \Delta t \left(D_c^{2p} v \right)_j^n, & (j, n) \in \mathbb{Z} \times \llbracket 0, N \rrbracket, \\ v_j^0 = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u_0(y) dy, & j \in \mathbb{Z}, \end{cases} \quad (2.6a)$$

avec

$$\left(D_c^{2p} v \right)_j^n = \sum_{k=0}^{2p} \frac{\binom{2p}{k} (-1)^k}{\Delta x^{2p}} v_{p-k+j}^n \quad (\text{schéma centré}). \quad (2.7)$$

Dans les deux cas, nous définissons l'erreur de consistance de la manière suivante.

Définition 2.3. • Pour tout $(n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}$, l'erreur de consistance ϵ_j^n est définie par

$$\epsilon_j^n = \frac{(u_{\Delta})_j^{n+1} - (u_{\Delta})_j^n}{\Delta t} + \theta (D_{\bullet}^s u_{\Delta})_j^{n+1} + (1-\theta) (D_{\bullet}^s u_{\Delta})_j^n,$$

avec $D_{\bullet}^s = D_{\bullet}^{2p+1}$ donnés par (2.5a)–(2.5c) ou $D_{\bullet}^s = (-1)^p D_c^{2p}$ donné par (2.7), et

$$(u_{\Delta})_j^n = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u(t^n, y) dy,$$

pour u solution exacte du problème de Cauchy (2.1).

• L'erreur de convergence est, quant à elle, définie par

$$e_j^n = v_j^n - (u_{\Delta})_j^n, \quad (n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}. \quad (2.8)$$

Il s'agit, dans la définition précédente, de comparer la solution numérique à la solution exacte moyennée sur une maille en espace. Le choix d'une telle moyenne est motivé par l'utilisation de la norme ℓ_Δ^2 (cf introduction générale) qui fait intervenir une somme sur tous les entiers relatifs j . La moyenne sur chaque maille permet, en effet, de retrouver l'intégrale sur tout $\mathbb{R}$ et ainsi, d'avoir une somme finie sans hypothèse supplémentaire. *A contrario*, il est relativement plus compliqué de montrer que la somme des $|u(t^n, x_j)|^2$ est bien finie (cf remarque 2.16). Ce choix de moyennner la solution exacte justifie aussi l'initialisation du schéma numérique par une moyenne sur chaque maille $[x_j, x_{j+1}]$: cette initialisation permet d'avoir $v_j^0 = (u_\Delta)_j^0$ pour tout $j \in \mathbb{Z}$ ce qui implique

$$e_j^0 = 0, \quad \forall j \in \mathbb{Z}.$$



Principaux résultats. Il est intéressant de noter que dans le cas où $\mathcal{P}$ est impair, la stabilité dépend de la parité de $\lfloor \frac{\mathcal{P}}{2} \rfloor$. Plus précisément, nous démontrons le résultat suivant (cf proposition 2.19)

Résultat : Schémas décentrés et $\mathcal{P}$ impair

Les schémas décentrés à droite (2.4a) avec (2.5a) sont stables sous la condition de Courant-Friedrichs-Lewy (condition CFL) $\Delta t(1 - 2\theta) \leq \frac{\Delta x^{2p+1}}{2^{2p}}$ dès que $\lfloor \frac{\mathcal{P}}{2} \rfloor$ est impair alors que les schémas décentrés à gauche (2.4a) avec (2.5b) sont stables sous la même condition dès que $\lfloor \frac{\mathcal{P}}{2} \rfloor$ est pair.

Les résultats de convergence sont les suivants (cf théorèmes 2.39 pour $\mathcal{P}$ impair ou 2.50 pour $\mathcal{P}$ pair).

Résultat : Ordre de convergence pour le cas $\mathcal{P}$ impair

- Lorsque $u_0 \in H^s(\mathbb{R})$, les schémas (2.4a) sont convergents en temps à l'ordre $\frac{\min(s, 2\mathcal{P})}{2\mathcal{P}}$ ou $2\frac{\min(s, 3\mathcal{P})}{3\mathcal{P}}$ dans le cas du schéma de Crank-Nicolson ($\theta = \frac{1}{2}$).
 - Pour ce qui est de l'ordre en espace, le schéma centré (2.4a) avec (2.5c) est convergent à l'ordre $2\frac{\min(s, \mathcal{P}+2)}{\mathcal{P}+2}$. Les schémas décentrés (2.4a) avec (2.5a) ou (2.5b), s'ils sont stables, sont convergents à l'ordre $\frac{\min(s, \mathcal{P}+1)}{\mathcal{P}+1}$ en espace.
-

Résultat : Ordre de convergence pour le cas $\mathcal{P}$ pair

Lorsque $u_0 \in H^s(\mathbb{R})$, le schéma (2.6a) avec (2.7) est convergent à l'ordre $\frac{\min(s, 2\mathcal{P})}{2\mathcal{P}}$ en temps et $2\frac{\min(s, \mathcal{P}+2)}{\mathcal{P}+2}$ en espace.

Le schéma de Crank-Nicolson ($\theta = \frac{1}{2}$) est convergent en temps à l'ordre $2\frac{\min(s, 3\mathcal{P})}{3\mathcal{P}}$.



Organisation du chapitre. Pour les raisons évoquées précédemment, nous traitons séparément le cas $\mathcal{P}$ impair (cf section 2.2) du cas $\mathcal{P}$ pair (cf section 2.3). Néanmoins, la même méthode sera appliquée aux deux cas : une étude de consistance tout d'abord par développements de Taylor et différences divisées (sous-sections 2.2.1.2 pour le cas impair et 2.3.1.1 pour le cas pair), suivie d'une étude de stabilité par l'utilisation de la transformée de Fourier (sous-sections 2.2.1.3 et 2.3.1.2). Pour finir, l'ordre de convergence sera quantifié en fonction de la régularité de la donnée initiale (sections 2.2.2 et 2.3.2). Lorsque cela est nécessaire, nous régularisons les données initiales par produit de convolution avec une suite régularisante.



2.2 High-order linear dispersive equations

For the particular case $p = 1$, that is Airy equation, our schemes (2.4a) with (2.5a) or (2.5b) or (2.5c) are very similar to those studied by Mengzhao in [83]. He reviews some classical finite difference schemes named rightside/leftside/central explicit schemes (which correspond to our explicit right-winded/left-winded/central schemes), Crank-Nicolson case (which corresponds to $\theta = \frac{1}{2}$ in our scheme). He even details the leap-frog scheme, when the time derivative is approximated by $\frac{v_j^{n+1} - v_j^{n-1}}{2\Delta t}$, not by $\frac{v_j^{n+1} - v_j^n}{\Delta t}$. Because of the strong similarities between Mengzhao's schemes and ours, we recover the same stability conditions in Subsection 2.2.1.3.

Remark 2.4. For more simplicity, we detail proofs on right-winded difference scheme (2.5a) and explain in some remarks the differences that occur in the left-winded and central cases, respectively (2.5b) and (2.5c).

2.2.1 Order of accuracy for an initial datum in $H^{4p+2}(\mathbb{R})$

We hereafter find some conditions on θ , Δt and Δx for the schemes to be consistent and stable, to conclude the convergence study according to the Lax-Richtmyer Theorem 1.23. In this section, initial data is supposed to be smooth enough for all computations to be valid, without needed any regularization tool. Indeed, supposing u_0 regular is sufficient to ensure the same regularity for $u(t, \cdot)$ for all $t \in [0, T]$ because of the following result.

Lemma 2.5. Let u be a solution of (2.2). If $u_0 \in H^s(\mathbb{R})$, for any $s > 0$ then $u(t, \cdot) \in H^s(\mathbb{R})$, $\forall t \in [0, T]$.

Proof. If u is a solution of (2.2), then by linearity of the equation, all the derivatives of u verify (2.2) too and by Fourier transform, the L^2 -norm of all its derivatives are conserved :

$$\|\partial_x^k u(t, \cdot)\|_{L^2(\mathbb{R})} = \|\partial_x^k u_0\|_{L^2(\mathbb{R})},$$

for all $k \in \mathbb{N}$. □

2.2.1.1 A review of divided differences

We recall in this subsection the basic properties of divided differences of a function f , which are useful to simplify calculus of the consistency error.

Definition 2.6. Given a function f and $k + 1$ points $(x_0, \dots, x_k)$, the divided difference of order k of f is defined recursively by :

$$\begin{cases} f[x_j] = f(x_j), \forall j \in \mathbb{Z}, \\ f[x_0, x_1, \dots, x_k] = \frac{f[x_1, \dots, x_k] - f[x_0, \dots, x_{k-1}]}{x_k - x_0}, \forall k \geq 1. \end{cases}$$

The first main property reads as follows and is proved for example in [42] by using Rolle's theorem.

Proposition 2.7. Let $(x_0, \dots, x_J)$ be a sequence of $J + 1$ distinct points and assume that $f \in C^J(\mathbb{R})$. Then there exists $\xi \in (\min(x_0, \dots, x_J), \max(x_0, \dots, x_J))$ such as

$$f[x_0, \dots, x_J] = \frac{f^{(J)}(\xi)}{J!}.$$

Moreover, if we assume an additional hypothesis on the distribution of $(x_0, \dots, x_k)$, then $f[x_0, \dots, x_k]$ has an explicit formula. Indeed, for all $(J, \ell) \in \mathbb{N}^2$, and $(x_\ell, \dots, x_{\ell+J})$, $J+1$ points regularly spaced of h , we have, by induction on J :

$$f[x_\ell, \dots, x_{\ell+J}] = \frac{\sum_{k=0}^J \binom{J}{k} (-1)^{k+J} f(x_{k+\ell})}{J!h^J}. \quad (2.9)$$

Thanks to the previous equality, the divided differences enjoy the following lemma.

Lemma 2.8. *Let $j \in \mathbb{Z}$, $p \in \mathbb{N}$ and $(x_{j-p-1}, \dots, x_{j+p+1})$ $2p+3$ points regularly spaced of h , then the following equations hold*

$$\begin{aligned} f[x_{j-p}, \dots, x_{j+p}] &= \frac{\sum_{k=0}^{2p} \binom{2p}{k} (-1)^k f(x_{k+j-p})}{(2p)!h^{2p}} = \frac{\sum_{k=0}^{2p} \binom{2p}{k} (-1)^k f(x_{p-k+j})}{(2p)!h^{2p}}, \\ f[x_{j-p}, \dots, x_{j+p}, x_{j+p+1}] &= \frac{\sum_{k=0}^{2p+1} \binom{2p+1}{k} (-1)^{k+1} f(x_{k+j-p})}{(2p+1)!h^{2p+1}} = \frac{\sum_{k=0}^{2p+1} \binom{2p+1}{k} (-1)^k f(x_{p-k+j+1})}{(2p+1)!h^{2p+1}}, \\ f[x_{j-p-1}, x_{j-p}, \dots, x_{j+p}] &= \frac{\sum_{k=0}^{2p+1} \binom{2p+1}{k} (-1)^{k+1} f(x_{k+j-p-1})}{(2p+1)!h^{2p+1}} = \frac{\sum_{k=0}^{2p+1} \binom{2p+1}{k} (-1)^k f(x_{p-k+j})}{(2p+1)!h^{2p+1}}. \end{aligned}$$

Proof. We use Equation (2.9) respectively with $(J, \ell) = (2p, j-p)$, $(J, \ell) = (2p+1, j-p)$ and $(J, \ell) = (2p+1, j-p-1)$. $\square$

Combine Proposition 2.7 and Lemma 2.8. Then, under the same hypothesis of Lemma 2.8 and if f is smooth enough, the divided difference of order $2p+3$ of f is equal to :

$$\frac{\sum_{k=0}^{2p} \binom{2p}{k} (-1)^k f(x_{p-k+j})}{h^{2p}} = (2p)!f[x_{j-p}, \dots, x_{j+p}] = f^{(2p)}(\xi), \quad (2.10)$$

with $\xi \in (x_{j-p}, x_{j+p})$,

$$\frac{\sum_{k=0}^{2p+1} \binom{2p+1}{k} (-1)^k f(x_{p-k+j+1})}{h^{2p+1}} = (2p+1)!f[x_{j-p}, \dots, x_{j+p+1}] = f^{(2p+1)}(\xi), \quad (2.11)$$

with $\xi \in (x_{j-p}, x_{j+p+1})$,

$$\frac{\sum_{k=0}^{2p+1} \binom{2p+1}{k} (-1)^k f(x_{p-k+j})}{h^{2p+1}} = (2p+1)!f[x_{j-p-1}, \dots, x_{j+p}] = f^{(2p+1)}(\xi), \quad (2.12)$$

with $\xi \in (x_{j-p-1}, x_{j+p})$.

Sums are considerably simplified thanks to the previous relations, as shown in the following lemma.

Lemma 2.9. *For all ℓ and p in $\mathbb{N}$, there exists $y \in (-p, p)$ such that*

$$\sum_{k=0}^{2p} \binom{2p}{k} (-1)^k (p-k)^\ell = \begin{cases} 0 & \text{if } \ell < 2p, \\ \ell! & \text{if } \ell = 2p, \\ \frac{\ell!}{(\ell-2p)!} y^{\ell-2p} & \text{if } \ell > 2p. \end{cases}$$

Section 2.2. High-order linear dispersive equations

Proof. We set $g_\ell : y \mapsto y^\ell$, and we use Equation (2.10) with $f = g_\ell$, $h = 1$, $j = 0$ and $x_i = i$ for $i \in \mathbb{Z}$.

In that case,

$$\sum_{k=0}^{2p} \binom{2p}{k} (-1)^k (p-k)^\ell = g_\ell^{(2p)}(\xi),$$

with $\xi \in (x_{-p}, x_p) = (-p, p)$. We compute the derivative function of g_ℓ to conclude. $\square$

If we apply this proof with Equation (2.11) (respectively (2.12)) instead of (2.10), we have similarly the following results.

Lemma 2.10. *For all $\ell \in \mathbb{N}$ and $p \in \mathbb{N}$, there exists $y \in (-p, p+1)$ such that*

$$\sum_{k=0}^{2p+1} \binom{2p+1}{k} (-1)^k (p-k+1)^\ell = \begin{cases} 0 & \text{if } \ell < 2p+1, \\ \ell! & \text{if } \ell = 2p+1, \\ \frac{\ell!}{(\ell-2p-1)!} y^{\ell-2p-1} & \text{if } \ell > 2p+1. \end{cases} \quad (2.13)$$

Lemma 2.11. *For all $\ell \in \mathbb{N}$ and $p \in \mathbb{N}$, there exists $y \in (-p-1, p)$ such that*

$$\sum_{k=0}^{2p+1} \binom{2p+1}{k} (-1)^k (p-k)^\ell = \begin{cases} 0 & \text{if } \ell < 2p+1, \\ \ell! & \text{if } \ell = 2p+1, \\ \frac{\ell!}{(\ell-2p-1)!} y^{\ell-2p-1} & \text{if } \ell > 2p+1. \end{cases}$$

2.2.1.2 Consistency estimate

Proposition 2.12. *Assume $u_0 \in H^{4p+2}(\mathbb{R})$ (and $u_0 \in H^{6p+3}(\mathbb{R})$ if $\theta = \frac{1}{2}$) then, for the right-winded or left-winded finite difference schemes (2.5a) and (2.5b), there exists a constant $C > 0$ independent of $\Delta x, \Delta t$ and u verifying the following consistency inequality*

$$\|\epsilon\|_{\ell^\infty([0,N];\ell_\Delta^2(\mathbb{Z}))} \leq C \left(\Delta t \left| \frac{1}{2} - \theta \right| \|\partial_x^{4p+2} u_0\|_{L^2(\mathbb{R})} + \Delta x \|\partial_x^{2p+2} u_0\|_{L^2(\mathbb{R})} + \Delta t^2 \|\partial_x^{6p+3} u_0\|_{L^2(\mathbb{R})} \right).$$

For the central finite difference scheme (2.5c), the consistency inequality is as follows

$$\|\epsilon\|_{\ell^\infty([0,N];\ell_\Delta^2(\mathbb{Z}))} \leq C \left(\Delta t \left| \frac{1}{2} - \theta \right| \|\partial_x^{4p+2} u_0\|_{L^2(\mathbb{R})} + \Delta x^2 \|\partial_x^{2p+3} u_0\|_{L^2(\mathbb{R})} + \Delta t^2 \|\partial_x^{6p+3} u_0\|_{L^2(\mathbb{R})} \right).$$

Remark 2.13. *In Section 2.2.1.2, we suppose the initial datum regular enough to compute all the needed derivatives and the Taylor expansion up to the desired order.*

Proof of Proposition 2.12. For $u_0 \in H^{4p+2}(\mathbb{R})$ and for the right-winded finite difference scheme, one has

$$\left(D_{\bullet}^{2p+1} u_\Delta \right)_j^n = \left(D_+^{2p+1} u_\Delta \right)_j^n = \sum_{k=0}^{2p+1} \frac{\binom{2p+1}{k} (-1)^k}{\Delta x^{2p+1}} \left(\frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u(t^n, y + (p-k+1)\Delta x) dy \right).$$

Using a Taylor expansion (in space) up to order $2p+2$ and exchanging the two sums inside leads to

$$\left(D_+^{2p+1} u_\Delta \right)_j^n = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} \sum_{\ell=0}^{2p+1} \frac{\partial_x^\ell u(t^n, y)}{\ell!} \frac{\left(\sum_{k=0}^{2p+1} \binom{2p+1}{k} (-1)^k (p-k+1)^\ell \Delta x^\ell \right)}{\Delta x^{2p+1}} dy + (R_+)_j^n, \quad (2.14)$$

where

$$(R_+)_j^n = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} \int_y^{y+(p-k+1)\Delta x} \frac{\partial_x^{2p+2} u(t^n, z)}{(2p+1)!} \left(\frac{\sum_{k=0}^{2p+1} \binom{2p+1}{k} (-1)^k (y + (p-k+1)\Delta x - z)^{2p+1}}{\Delta x^{2p+1}} \right) dz dy.$$

For simplicity, we will only use $\|R_+^n\|_{\ell_\Delta^2} \leq C\Delta x \|\partial_x^{2p+2} u(t^n, \cdot)\|_{L^2(\mathbb{R})}$, with $C > 0$ a constant independent of Δx and u . Equation (2.14) is simplified thanks to Lemma 2.10. Eventually, we obtain $(D_+^{2p+1} u_\Delta)_j^n = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} \partial_x^{2p+1} u(t^n, y) dy + (R_+)_j^n$. Similarly, by adapting the previous computation, one has, if $\theta \neq \frac{1}{2}$

$$(D_+^{2p+1} u_\Delta)_j^{n+1} = \frac{1}{\Delta x} \left(\int_{x_j}^{x_{j+1}} \partial_x^{2p+1} u(t^n, y) dy + \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \partial_t \partial_x^{2p+1} u(s, y) ds dy \right) + (R_+)_j^{n+1}$$

and, if $\theta = \frac{1}{2}$

$$(D_+^{2p+1} u_\Delta)_j^{n+1} = \frac{1}{\Delta x} \left(\int_{x_j}^{x_{j+1}} \partial_x^{2p+1} u(t^n, y) dy + \int_{x_j}^{x_{j+1}} \Delta t \partial_t \partial_x^{2p+1} u(t^n, y) dy + \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \partial_t^2 \partial_x^{2p+1} u(s, y) (t^{n+1} - s) ds dy \right) + (R_+)_j^{n+1}.$$

In order to compute the difference $\frac{(u_\Delta)_j^{n+1} - (u_\Delta)_j^n}{\Delta t}$ that appears in Definition 2.3, we perform a Taylor expansion (in time) up to order 2 (order 3 if $\theta = \frac{1}{2}$). Gathering all those results together yields (if $\theta \neq \frac{1}{2}$)

$$\epsilon_j^n = \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \partial_t^2 u(s, x) \frac{(t^{n+1} - s)}{\Delta t \Delta x} ds dx + \frac{\theta}{\Delta x} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \partial_t \partial_x^{2p+1} u(s, y) ds dy + \theta (R_+)_j^{n+1} + (1 - \theta) (R_+)_j^n$$

and, if $\theta = \frac{1}{2}$

$$\epsilon_j^n = \frac{1}{\Delta t \Delta x} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \partial_t^3 u(s, y) \frac{(t^{n+1} - s)^2}{2} ds dy + \frac{\theta}{\Delta x} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \partial_t^2 \partial_x^{2p+1} u(s, y) (t^{n+1} - s) ds dy + \theta (R_+)_j^{n+1} + (1 - \theta) (R_+)_j^n.$$

The conclusion comes from the relation $\partial_t u(t, x) = -\partial_x^{2p+1} u(t, x)$ and the Cauchy-Schwarz inequality (if $\theta \neq \frac{1}{2}$)

$$\|\epsilon^n\|_{\ell_\Delta^2}^2 \leq C \left(\int_{t^n}^{t^{n+1}} \Delta t (1 + \theta^2) \sum_{j \in \mathbb{Z}} \int_{x_j}^{x_{j+1}} |\partial_x^{4p+2} u(s, x)|^2 dx ds + \theta^2 \|R_+^{n+1}\|_{\ell_\Delta^2}^2 + (1 - \theta)^2 \|R_+^n\|_{\ell_\Delta^2}^2 \right), \quad (2.15)$$

and, if $\theta = \frac{1}{2}$

$$\|\epsilon^n\|_{\ell_\Delta^2}^2 \leq C \left(\int_{t^n}^{t^{n+1}} \Delta t^3 (1 + \theta^2) \sum_{j \in \mathbb{Z}} \int_{x_j}^{x_{j+1}} |\partial_x^{6p+3} u(s, y)|^2 ds dy + \theta^2 \|R_+^{n+1}\|_{\ell_\Delta^2}^2 + (1 - \theta)^2 \|R_+^n\|_{\ell_\Delta^2}^2 \right). \quad (2.16)$$

The conservation of the L^2 -norm (cf Lemma 2.5) concludes the proof. $\square$

Remark 2.14. The regularity $H^{4p+2}(\mathbb{R})$ (or $H^{6p+3}(\mathbb{R})$ if $\theta = \frac{1}{2}$) is essential in this proof and comes from the Taylor expansion in time, which involves $\partial_t^2 u(t^n, x)$ (or $\partial_t^2 \partial_x^{2p+1} u(s, y)$ if $\theta = \frac{1}{2}$).

Remark 2.15. We follow exactly the same guidelines for the left-winded finite difference scheme. For the central finite difference scheme, we need to perform a Taylor expansion in space up to order $2p + 3$ to obtain $(D_c^{2p+1} u_\Delta)_j^n = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} \partial_x^{2p+1} u(t^n, y) dy + (R_c)_j^n$, with $\|R_c^n\|_{\ell_\Delta^2} \leq C\Delta x^2 \|\partial_x^{2p+3} u_0\|_{L^2(\mathbb{R})}$.

Section 2.2. High-order linear dispersive equations

Remark 2.16. The previous proof explains our choice of $\frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u(t^n, y) dy$ instead of $u(t^n, x_j)$ in the consistency error (cf Definition 2.3). Indeed, the ℓ_Δ^2 -norm implies, among other terms, the presence of

$$\sum_{j \in \mathbb{Z}} \int_{x_j}^{x_{j+1}} |\partial_x^k u(s, y)|^2 ds dy$$

in Equations (2.15) and (2.16), and this sum is well defined for $u_0 \in H^k(\mathbb{R})$. Without integrating on each cell $[x_j, x_{j+1}]$, the upper bound would be

$$\sum_{j \in \mathbb{Z}} |\partial_x^k u(t^n, x_j)|^2$$

and the summability would be more difficult to prove.

2.2.1.3 Stability

We note, for all $(v_j^n)_{j \in \mathbb{Z}}$ and $\xi \in [0, 1]$,

$$\widehat{v}^n(\xi) := \sum_{k \in \mathbb{Z}} v_k^n e^{2i\pi k \xi} \quad \text{in } L^2([0, 1])$$

(the sequence $(v_j^n)_{(n,j)}$ is seen as Fourier-sequence of the function $\widehat{v}^n$). Parseval inequality yields the equivalence of the norms :

$$\sum_{j \in \mathbb{Z}} \Delta x |v_j^n|^2 = \Delta x \int_0^1 |\widehat{v}^n(\xi)|^2 d\xi. \quad (2.17)$$

Eventually, we define the shift operator $\mathcal{S}^\ell$ by $\mathcal{S}^\ell v^n = (v_{j+\ell}^n)_{j \in \mathbb{Z}}$ thus, $\widehat{\mathcal{S}^\ell v^n} = e^{-2i\pi \ell \xi} \widehat{v}^n$.

Definition 2.17. If the scheme is linear with constant coefficients, we define the amplification factor $A(\xi)$ of the scheme by

$$\widehat{v^{n+1}}(\xi) = A(\xi) \widehat{v}^n(\xi), \quad \xi \in [0, 1].$$

This Fourier description enables us to set the following proposition about the stability.

Proposition 2.18. A linear scheme with constant coefficient is stable as defined in Definition 1.17 if and only if, there exists a constant $C > 0$ independent of Δt and Δx such that, for all $\Delta t, \Delta x \leq 1$ (possibly satisfying a CFL condition)

$$|A(\xi)| \leq (1 + C\Delta t), \quad \xi \in [0, 1].$$

We use the previous proposition to find the following CFL conditions of stability.

Proposition 2.19. • Asymptotically, when Δx and Δt decrease towards 0, the right-winded scheme (2.5a) is stable under the CFL condition

$$\Delta t(1 - 2\theta) < \frac{\Delta x^{2p+1}}{2^{2p}},$$

for p odd, and stable provided if there exists a constant $C > 0$ (independent of Δx and Δt) such that

$$(p+1) \left(\frac{p^p}{(2C)^p (2p+1)^{2p+1}} \right)^{\frac{1}{p+1}} \Delta x^{\frac{(2p+1)}{p+1}} < \Delta t [2\theta - 1 + 2\theta^2 C \Delta t],$$

for p even. The reverse occurs for the left-winded scheme (2.5b).

• The central scheme (2.5c) is always stable if there exists a constant $C > 0$ independent of Δx and Δt such that the following CFL condition occurs : $\Delta t(1 - 2\theta) < 2C \frac{\Delta x^{4p+2}}{2^{4p}}$.

Remark 2.20. For the right-winded scheme with p odd and the left-winded scheme with p even, the asymptotic Proposition 2.19 gives the same results than taking C equal to zero in Proposition 2.18.

The results of Proposition 2.19 on the Courant-Friedrich-Lewy conditions of stability is summarized in Table 2.1.

	p even		p odd
	$p = 0$ (Advection)	$p \neq 0$	
Right-winded scheme	stable provided that $\Delta x < \Delta t(2\theta - 1 + 2\theta^2 C \Delta t)$	stable provided that $B \Delta x^{\frac{2p+1}{p+1}} < \Delta t(2\theta - 1 + 2\theta^2 C \Delta t)$ $B = (p+1) \left(\frac{p^p}{(2C)^p (2p+1)^{2p+1}} \right)^{\frac{1}{p+1}}$	stable under the CFL cond. $\Delta t(1 - 2\theta) < \frac{\Delta x^{2p+1}}{2^{2p}}$
Left-winded scheme	stable under the CFL cond. $\Delta t(1 - 2\theta) < \Delta x$	stable under the CFL cond. $\Delta t(1 - 2\theta) < \frac{\Delta x^{2p+1}}{2^{2p}}$	stable provided that $B \Delta x^{\frac{2p+1}{p+1}} < \Delta t(2\theta - 1 + 2\theta^2 C \Delta t)$ $B = (p+1) \left(\frac{p^p}{(2C)^p (2p+1)^{2p+1}} \right)^{\frac{1}{p+1}}$
Central scheme	stable under the CFL cond. $\Delta t(1 - 2\theta) < 2C \Delta x^2$	stable under the CFL cond. $\Delta t(1 - 2\theta) < 2C \frac{\Delta x^{4p+2}}{2^{4p}}$	stable under the CFL cond. $\Delta t(1 - 2\theta) < 2C \frac{\Delta x^{4p+2}}{2^{4p}}$

TABLE 2.1 – Stability results for finite difference θ -schemes

Remark 2.21. For p even and the right-winded scheme (respectively for p odd and the left-winded scheme), the stability conditions are unconventional in the sense that they impose a restriction on Δx with respect to Δt . Hereafter, we deal only with "classical" CFL condition and consider the right-winded scheme for p odd and the left-winded scheme for p even.

The following computation will simplify the proof of Proposition 2.19.

Lemma 2.22. One has, for all $\xi \in [0, 1]$,

$$\sum_{k=0}^{2p+1} \binom{2p+1}{k} (-1)^k e^{-2i\pi(p-k+1)\xi} = e^{-i\pi\xi} (-2i\sin(\pi\xi))^{2p+1}.$$

Proof of Lemma 2.22. We have to develop the left hand side, which holds

$$\begin{aligned} \sum_{k=0}^{2p+1} \binom{2p+1}{k} (-1)^k e^{-2i\pi(p-k+1)\xi} &= e^{2i\pi p\xi} \sum_{k=0}^{2p+1} \binom{2p+1}{k} (-1)^k (e^{-2i\pi\xi})^{2p-k+1} \\ &= e^{2i\pi p\xi} (e^{-2i\pi\xi} - 1)^{2p+1} \\ &= e^{-i\pi\xi} (-2i\sin(\pi\xi))^{2p+1}. \end{aligned}$$

□

Proof of Proposition 2.19. The right-winded finite difference scheme (2.5a) leads to

$$\begin{aligned} \widehat{v^{n+1}}(\xi) &\left(1 + \frac{\theta \Delta t}{(\Delta x)^{2p+1}} \sum_{k=0}^{2p+1} \binom{2p+1}{k} (-1)^k e^{-2i\pi(p-k+1)\xi} \right) \\ &= \widehat{v^n}(\xi) \left(1 - \frac{(1-\theta)\Delta t}{(\Delta x)^{2p+1}} \sum_{k=0}^{2p+1} \binom{2p+1}{k} (-1)^k e^{-2i\pi(p-k+1)\xi} \right), \text{ for any } \xi \text{ in } [0, 1]. \end{aligned}$$

Section 2.2. High-order linear dispersive equations

The two sums are simplified thanks to Lemma 2.22. We finally obtain $\widehat{v^{n+1}}(\xi) = A_+^{2p+1}(\xi) \widehat{v^n}(\xi)$, with A_+^{2p+1} the amplification coefficient defined by, $\forall \xi \in [0, 1]$

$$A_+^{2p+1}(\xi) = \frac{\left(1 - \frac{(1-\theta)\Delta t}{(\Delta x)^{2p+1}} e^{-i\pi\xi} (-i)^{2p+1} (2\sin(\pi\xi))^{2p+1}\right)}{\left(1 + \frac{\theta\Delta t}{(\Delta x)^{2p+1}} e^{-i\pi\xi} (-i)^{2p+1} (2\sin(\pi\xi))^{2p+1}\right)}. \quad (2.18)$$

According to Proposition 2.18, we are looking for a condition ensuring

$$\left|A_+^{2p+1}(\xi)\right|^2 \leq (1 + C\Delta t)^2 \quad \text{for any } \xi \text{ in } [0, 1]. \quad (2.19)$$

CASE p EVEN : Assume that the parameter p of the spatial derivative is even. Then we get

$$(-i)^{2p+1} e^{-i\pi\xi} = -ie^{-i\pi\xi} = -i\cos(\pi\xi) - \sin(\pi\xi).$$

Condition (2.19) leads to, for any ξ in $[0, 1]$,

$$\begin{aligned} & 1 - \frac{2(1-\theta)\Delta t}{(\Delta x)^{2p+1}} (2\sin(\pi\xi))^{2p+1} (-\sin(\pi\xi)) + \frac{(1-\theta)^2(\Delta t)^2}{(\Delta x)^{4p+2}} (2\sin(\pi\xi))^{4p+2} (-\sin(\pi\xi))^2 \\ & \quad + \frac{(1-\theta)^2(\Delta t)^2}{(\Delta x)^{4p+2}} (2\sin(\pi\xi))^{4p+2} \cos(\pi\xi)^2 \\ & \leq (1 + C\Delta t)^2 \left(1 + \frac{2\theta\Delta t}{(\Delta x)^{2p+1}} (2\sin(\pi\xi))^{2p+1} (-\sin(\pi\xi)) + \frac{\theta^2(\Delta t)^2}{(\Delta x)^{4p+2}} (2\sin(\pi\xi))^{4p+2} (-\sin(\pi\xi))^2 \right. \\ & \quad \left. + \frac{\theta^2(\Delta t)^2}{(\Delta x)^{4p+2}} (2\sin(\pi\xi))^{4p+2} \cos(\pi\xi)^2 \right). \end{aligned}$$

Hence, it simplifies into, for any ξ in $[0, 1]$,

$$\begin{aligned} & \frac{2^{2p+2}\Delta t (\sin(\pi\xi))^{2p+2}}{\Delta x^{2p+1}} [1 + 2C\theta\Delta t + \theta C^2\Delta t^2 \\ & \quad + \frac{(2\sin(\pi\xi))^{2p}\Delta t}{\Delta x^{2p+1}} (1 - 2\theta - 2\theta^2C\Delta t - \theta^2C^2\Delta t^2)] \leq \Delta t [2C + C^2\Delta t]. \end{aligned}$$

We rewrite the previous inequality into

$$\begin{aligned} & 2^{2p+2} (\sin(\pi\xi))^{2p+2} [1 + 2C\theta\Delta t + \theta C^2\Delta t^2 \\ & \quad + \frac{(2\sin(\pi\xi))^{2p}\Delta t}{\Delta x^{2p+1}} (1 - 2\theta - 2\theta^2C\Delta t - \theta^2C^2\Delta t^2)] \leq \Delta x^{2p+1} [2C + C^2\Delta t], \quad \forall \xi \in [0, 1]. \quad (2.20) \end{aligned}$$

Remark 2.23. Asymptotically when Δt and Δx decrease towards 0, Inequality (2.20) implies that $1 - 2\theta \leq 0$. Otherwise, the left hand side is positive and can not be upper bounded by the right hand side (which is as small as possible). Hereafter, for the case p even, we focus only for $\theta \geq \frac{1}{2}$, for this reason.

By defining $X := 2\sin(\pi\xi) \in [-2, 2]$, we are looking for a relation between Δt and Δx which allows to have, for all $X \in [-2, 2]$ and for all $\theta \in [\frac{1}{2}, 1]$,

$$X^{2p+2} [1 + 2C\theta\Delta t + \theta^2C^2\Delta t^2] + X^{4p+2} \frac{\Delta t}{\Delta x^{2p+1}} [1 - 2\theta - 2\theta^2C\Delta t - \theta^2C^2\Delta t^2] \leq \Delta x^{2p+1} [2C + C^2\Delta t].$$

Let f_{even} be the even function defined by

$$f_{\text{even}}(X) := X^{2p+2} [1 + 2C\theta\Delta t + \theta^2C^2\Delta t^2] + X^{4p+2} \frac{\Delta t}{\Delta x^{2p+1}} [1 - 2\theta - 2\theta^2C\Delta t - \theta^2C^2\Delta t^2].$$

Proving that the maximum of f_{even} is less or equal to $\Delta x^{2p+1} [2C + C^2 \Delta t]$ is a sufficient condition to obtain the stability of the scheme. The maximum is obtained for X_{max} , satisfying

$$X_{\text{max}}^{2p} = -\frac{(p+1)\Delta x^{2p+1} [1 + 2C\theta\Delta t + \theta^2 C^2 \Delta t^2]}{(2p+1)\Delta t [1 - 2\theta - 2\theta^2 C\Delta t - \theta^2 C^2 \Delta t^2]}. \quad (2.21)$$

Remark 2.24. Since $\theta \geq \frac{1}{2}$, the right hand side of Equation (2.21) is positive as expected.

Therefore, one has

$$\begin{aligned} f_{\text{even}}(X_{\text{max}}) &= X_{\text{max}}^{2p+2} \left[1 + 2C\theta\Delta t + \theta^2 C^2 \Delta t^2 + X_{\text{max}}^{2p} \frac{\Delta t}{\Delta x^{2p+1}} (1 - 2\theta - 2\theta^2 C\Delta t - \theta^2 C^2 \Delta t^2) \right] \\ &= X_{\text{max}}^{2p+2} [1 + 2C\theta\Delta t + \theta^2 C^2 \Delta t^2] \frac{p}{2p+1} > 0. \end{aligned}$$

A condition which allows $f_{\text{even}}(X_{\text{max}})^{2p} \leq \Delta x^{2p(2p+1)} [2C + C^2 \Delta t]^{2p}$ would be a sufficient condition to obtain stability of the scheme. One has

$$f_{\text{even}}(X_{\text{max}})^{2p} = \frac{p^{2p}(p+1)^{2p+2}\Delta x^{(2p+1)(2p+2)} [1 + 2C\theta\Delta t + \theta^2 C^2 \Delta t^2]^{4p+2}}{(2p+1)^{4p+2}\Delta t^{2p+2} [1 - 2\theta - 2\theta^2 C\Delta t - \theta^2 C^2 \Delta t^2]^{2p+2}}.$$

Imposing

$$\begin{aligned} p^{2p}(p+1)^{2p+2}\Delta x^{2(2p+1)} [1 + 2C\theta\Delta t + \theta^2 C^2 \Delta t^2]^{4p+2} \\ \leq [2C + C^2 \Delta t]^{2p} (2p+1)^{4p+2}\Delta t^{2p+2} [2\theta - 1 + 2\theta^2 C\Delta t + \theta^2 C^2 \Delta t^2]^{2p+2} \end{aligned} \quad (2.22)$$

is thus sufficient to obtain the stability of the right-winded finite difference scheme (2.5a) with p even. Asymptotically, when Δt and Δx decrease towards 0, Condition (2.22) becomes

$$p^{\frac{p}{p+1}}(p+1)\Delta x^{\frac{(2p+1)}{p+1}} < (2C)^{\frac{p}{p+1}} (2p+1)^{\frac{2p+1}{p+1}} \Delta t [2\theta - 1 + 2\theta^2 C\Delta t].$$

For $\theta \neq \frac{1}{2}$, a sufficient condition for the stability is thus

$$(p+1) \left(\frac{p^p}{(2C)^p (2p+1)^{2p+1}} \right)^{\frac{1}{p+1}} \Delta x^{\frac{(2p+1)}{p+1}} < \Delta t [2\theta - 1]. \quad (2.23)$$

For the Crank-Nicolson case ($\theta = \frac{1}{2}$), a sufficient condition is

$$\frac{2(p+1)}{C} \left(\frac{p^p}{(2C)^p (2p+1)^{2p+1}} \right)^{\frac{1}{p+1}} \Delta x^{\frac{(2p+1)}{p+1}} < \Delta t^2. \quad (2.24)$$

Remark 2.25. • Again, the parameter θ has to verify $\theta \geq \frac{1}{2}$ for Relation (2.23) to be possible.

• For $p = 0$, we recognize the classical condition $\Delta x < \Delta t(2\theta - 1)$, which is the stability condition for the right-winded scheme to be stable for the advection equation.

Thus, the right-winded finite difference scheme with p even is stable under the condition (2.23) for $\theta > \frac{1}{2}$ and the condition (2.24) for the Crank-Nicolson case, for Δt and Δx small enough.

CASE p ODD : In this case, the parameter p of the spatial derivative is odd, then

$$(-i)^{2p+1} e^{-i\pi\tilde{\xi}} = i e^{-i\pi\tilde{\xi}} = i \cos(\pi\tilde{\xi}) + \sin(\pi\tilde{\xi}).$$

Section 2.2. High-order linear dispersive equations

Condition (2.19) becomes, for any ξ in $[0, 1]$,

$$\begin{aligned} & 1 - \frac{2(1-\theta)\Delta t}{(\Delta x)^{2p+1}} (2\sin(\pi\xi))^{2p+1} \sin(\pi\xi) + \frac{(1-\theta)^2(\Delta t)^2}{(\Delta x)^{4p+2}} (2\sin(\pi\xi))^{4p+2} \sin(\pi\xi)^2 \\ & \quad + \frac{(1-\theta)^2(\Delta t)^2}{(\Delta x)^{4p+2}} (2\sin(\pi\xi))^{4p+2} \cos(\pi\xi)^2 \\ & \leq (1 + C\Delta t)^2 \left(1 + \frac{2\theta\Delta t}{(\Delta x)^{2p+1}} (2\sin(\pi\xi))^{2p+1} \sin(\pi\xi) + \frac{\theta^2(\Delta t)^2}{(\Delta x)^{4p+2}} (2\sin(\pi\xi))^{4p+2} \sin(\pi\xi)^2 \right. \\ & \quad \left. + \frac{\theta^2(\Delta t)^2}{(\Delta x)^{4p+2}} (2\sin(\pi\xi))^{4p+2} \cos(\pi\xi)^2 \right). \end{aligned}$$

We simplify the previous inequality to obtain, for any ξ in $[0, 1]$

$$\begin{aligned} & (2\sin(\pi\xi))^{2p+2} \left[-1 - 2C\theta\Delta t - \theta C^2\Delta t^2 \right. \\ & \quad \left. + \frac{(2\sin(\pi\xi))^{2p}\Delta t}{\Delta x^{2p+1}} (1 - 2\theta - 2\theta^2C\Delta t - \theta^2C^2\Delta t^2) \right] \leq \Delta x^{2p+1} [2C + C^2\Delta t]. \quad (2.25) \end{aligned}$$

Remark 2.26. If $1 - 2\theta \leq 0$, Inequality (2.25) is automatically satisfied because the left hand side is negative so no CFL condition is needed.

As for the even case, we are searching for a condition to impose that the maximum of the even function f_{odd} is less or equal to $\Delta x^{2p+1} [2C + C^2\Delta t]$ with

$$f_{\text{odd}}(X) := X^{2p+2} \left[-1 - 2C\theta\Delta t - \theta C^2\Delta t^2 + X^{2p} \frac{\Delta t}{\Delta x^{2p+1}} (1 - 2\theta - 2\theta^2C\Delta t - \theta^2C^2\Delta t^2) \right], \quad X \in [-2, 2].$$

This time, the maximum of f_{odd} is obtained for $X_{\text{max}} = 2$ (which would mean choosing $\xi = \frac{1}{2}$). A sufficient condition to obtain the stability condition is as follows

$$2^{2p+2} \left[-1 - 2C\theta\Delta t - \theta C^2\Delta t^2 + 2^{2p} \frac{\Delta t}{\Delta x^{2p+1}} (1 - 2\theta - 2\theta^2C\Delta t - \theta^2C^2\Delta t^2) \right] \leq \Delta x^{2p+1} [2C + C^2\Delta t].$$

Asymptotically, when Δt and Δx decrease towards 0, the previous CFL condition becomes

$$\Delta t(1 - 2\theta) < \frac{\Delta x^{2p+1}}{2^{2p}}, \quad (2.26)$$

which is sufficient to ensure Inequality (2.25), for Δx and Δt small enough.

Thus, the right-winded finite difference scheme is stable under the CFL condition (2.26) for p odd and for Δt and Δx small enough.

Remark 2.27. We recognize that, if $\theta \geq \frac{1}{2}$, there is no CFL condition.

Proposition 2.19 is a straightforward consequence. □

Remark 2.28. For the left-winded finite difference scheme, the only difference in the amplification coefficient is $e^{i\pi\xi}$ instead of $e^{-i\pi\xi}$ (in both the numerator and denominator) :

$$A_-^{2p+1}(\xi) = \frac{\left(1 - \frac{(1-\theta)\Delta t}{(\Delta x)^{2p+1}} e^{i\pi\xi} (-2i\sin(\pi\xi))^{2p+1} \right)}{\left(1 + \frac{\theta\Delta t}{(\Delta x)^{2p+1}} e^{i\pi\xi} (-2i\sin(\pi\xi))^{2p+1} \right)}.$$

The parity needed for the stability changes because of that difference.

For the central finite difference scheme, $e^{-i\pi\zeta}$ is replaced with $\cos(\pi\zeta)$ in the numerator and the denominator of the amplification coefficient :

$$A_c^{2p+1}(\zeta) = \frac{1 - \frac{(1-\theta)\Delta t}{(\Delta x)^{2p+1}} \cos(\pi\zeta) (-2i\sin(\pi\zeta))^{2p+1}}{1 + \frac{\theta\Delta t}{(\Delta x)^{2p+1}} \cos(\pi\zeta) (-2i\sin(\pi\zeta))^{2p+1}}.$$

2.2.1.4 Error estimates

We recall the well-known discrete Grönwall lemma, whose proof can be found in [42], for example.

Lemma 2.29. *Let $(t^n)_{n \in \mathbb{N}}$ be an increasing sequence and $(v^n)_{n \in \mathbb{N}}$ be a nonnegative sequence such that we define $h^n := t^{n+1} - t^n > 0$, for $n \in \mathbb{N}$. Let $\Lambda > 0$ and $(w^n)_{n \in \mathbb{N}}$ be a sequence satisfying*

$$v^{n+1} \leq (1 + \Lambda h^n) v^n + |w^n|, \quad \forall n \in \mathbb{N}.$$

Then, the following inequality comes true

$$v^n \leq \exp(\Lambda(t^n - t^0)) v^0 + \sum_{k=0}^{n-1} \exp(\Lambda(t^n - t^{k+1})) |w^k|, \quad \forall n \in \mathbb{N}.$$

With Definition 2.3 of the convergence error, we are now able to state the main result of this section.

Theorem 2.30.

For an initial datum $u_0 \in H^{4p+2}(\mathbb{R})$ (and $u_0 \in H^{6p+3}(\mathbb{R})$ if $\theta = \frac{1}{2}$), there exists $C > 0$ a constant independent of Δt and Δx such that the error estimate of the right-winded finite difference scheme (2.5a) (if p is odd) or of the left-winded finite difference scheme (2.5b) (if p is even) satisfies

$$\|e\|_{\ell^\infty([0,N];\ell_\Delta^2(\mathbb{Z}))} \leq C \left(\Delta t \left| \frac{1}{2} - \theta \right| \left\| \partial_x^{4p+2} u_0 \right\|_{L^2(\mathbb{R})} + \Delta x \left\| \partial_x^{2p+2} u_0 \right\|_{L^2(\mathbb{R})} + \Delta t^2 \left\| \partial_x^{6p+3} u_0 \right\|_{L^2(\mathbb{R})} \right),$$

under the CFL conditions recalled in Table 2.1.

For the central finite difference scheme (2.5c), the convergence rate becomes

$$\|e\|_{\ell^\infty([0,N];\ell_\Delta^2(\mathbb{Z}))} \leq C \left(\Delta t \left| \frac{1}{2} - \theta \right| \left\| \partial_x^{4p+2} u_0 \right\|_{L^2(\mathbb{R})} + \Delta x^2 \left\| \partial_x^{2p+3} u_0 \right\|_{L^2(\mathbb{R})} + \Delta t^2 \left\| \partial_x^{6p+3} u_0 \right\|_{L^2(\mathbb{R})} \right),$$

under the CFL conditions of Table 2.1.

Remark 2.31. All those results are gathered in Table 2.6.

Proof of Theorem 2.30. We suppose p odd, so we work with the right-winded finite difference scheme. The case p even, with the left-winded scheme is similar. The definition of the convergence error implies

$$\widehat{e^{n+1}}(\zeta) = A_+^{2p+1}(\zeta) \widehat{e^n}(\zeta) - \frac{\Delta t}{1 + \frac{\theta\Delta t}{\Delta x^{2p+1}} e^{-i\pi\zeta} (-i)^{2p+1} (2\sin(\pi\zeta))^{2p+1}} \widehat{e^n}(\zeta).$$

One has $\left\| \frac{1}{1 + \frac{\theta\Delta t}{\Delta x^{2p+1}} e^{-i\pi\zeta} (-i)^{2p+1} (2\sin(\pi\zeta))^{2p+1}} \right\|_{L^\infty([0,1])} \leq 1$ and the stability condition gives

$\left\| A_+^{2p+1} \right\|_{L^\infty([0,1])} \leq 1 + C\Delta t$. Thus, we obtain the following estimate, by discrete Grönwall lemma (Lemma 2.29)

$$\|e^{n+1}\|_{\ell_\Delta^2} \leq (1 + C\Delta t) \|e^n\|_{\ell_\Delta^2} + \Delta t \|\epsilon^n\|_{\ell_\Delta^2} \leq e^{CT} \|e^0\|_{\ell_\Delta^2} + \Delta t e^{CT} \sum_{k=0}^n \|\epsilon^k\|_{\ell_\Delta^2}.$$

Section 2.2. High-order linear dispersive equations

The initial condition v_j^0 , Equation (2.4b), together with Definition 2.3 of u_Δ and the consistency error in Proposition 2.12 conclude the proof. $\square$

Remark 2.32. As expected, for the particular case $\theta = \frac{1}{2}$ (the so-called Crank-Nicolson case), the rate of convergence in time is better as illustrated in Table 2.6, provided $u_0 \in H^{6p+3}(\mathbb{R})$ (and not only in $H^{4p+2}(\mathbb{R})$).

2.2.2 Less smooth initial data

The previous order of accuracy is obtained for initial data u_0 at least in $H^{4p+2}(\mathbb{R})$ (or $H^{6p+3}(\mathbb{R})$ if $\theta = \frac{1}{2}$). In this section, our aim is to relax this hypothesis to obtain rates of convergence for non-smooth initial data, for example, $u_0 \in H^s(\mathbb{R})$ with $s > 0$.

2.2.2.1 A regularization of u_0

When u_0 is not smooth enough to verify $u_0 \in H^{4p+2}(\mathbb{R})$ (or $H^{6p+3}(\mathbb{R})$), we regularize it thanks to mollifiers $(\varphi^\delta)_{\delta>0}$. We recall here the construction introduced in general introduction (cf page 14).

Let χ be a C^∞ -function such that

- $0 \leq \chi \leq 1$,
- $\chi \equiv 1$ in $[-\frac{1}{2}, \frac{1}{2}]$ and $\text{supp}(\chi) \subset [-1, 1]$ (where supp is its support),
- $\chi(-\xi) = \chi(\xi)$, $\forall \xi \in [-1, 1]$.

Let φ be such as $\widehat{\varphi}(\xi) = \chi(\xi)$ (we recall that $\widehat{\varphi}$ stands for the Fourier transform of φ) and for all $\delta > 0$, we define φ^δ such that $\widehat{\varphi^\delta}(\xi) = \chi(\delta\xi)$, which implies $\varphi^\delta = \frac{1}{\delta} \varphi(\frac{\cdot}{\delta})$. Eventually,

- we shall still denote by u the exact solution of (2.2) starting from the initial datum u_0 .
- The numerical solution of (2.4a)-(2.4b) is still named $(v_j^n)_{(n,j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}}$.
- Let u^δ be the solution of (2.2) with $u_0^\delta := u_0 \star \varphi^\delta$ as initial datum, where $\star$ stands for the convolution product.
- We denote then $((v^\delta)_j^n)_{(n,j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}}$ the numerical solution obtained by applying the numerical scheme (2.4a) from the initial datum $(u_0^\delta)_\Delta$:

$$v_j^0 = (u_0^\delta)_\Delta = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u_0 \star \varphi^\delta(y) dy. \quad (2.27)$$

We first need to quantify the dependence of the Sobolev norms of u_0^δ with respect to δ .

Lemma 2.33. Assume $(m, s) \in \mathbb{R}^2$ with $m \geq s \geq 0$ and $u_0 \in H^s(\mathbb{R})$. Let u_0^δ be such as $u_0^\delta = u_0 \star \varphi^\delta$, then there exists a constant C independent of δ such as

$$\|u_0^\delta\|_{H^m(\mathbb{R})} \leq \frac{C}{\delta^{m-s}} \|u_0\|_{H^s(\mathbb{R})}. \quad (2.28)$$

Proof. According to (1.16), the $H^m(\mathbb{R})$ -norm of u_0^δ verifies

$$\begin{aligned} \|u_0 \star \varphi^\delta\|_{H^m(\mathbb{R})}^2 &= \int_{\mathbb{R}} (1 + |\xi|^2)^m |\chi(\delta\xi)|^2 |\widehat{u_0}(\xi)|^2 d\xi \\ &\leq \int_{\mathbb{R}} (1 + |\xi|^2)^s |\widehat{u_0}(\xi)|^2 (1 + |\xi|^2)^{m-s} |\chi(\delta\xi)|^2 d\xi. \end{aligned}$$

By hypothesis on χ and its support, there exists a constant $C > 0$ such as $(1 + |\xi|^2)^{m-s} \leq \frac{C}{\delta^{2(m-s)}}$ and moreover, one has $|\chi(\delta\xi)| \leq 1$, which concludes the proof. $\square$

We need then to know the rate of convergence of u_0^δ toward u_0 with respect to δ , which is summarized as follows.

Lemma 2.34. *Assume $u_0 \in H^s(\mathbb{R})$ with $0 \leq \ell \leq s$, then, there exists a constant C independent of δ such that*

$$\|u_0 - u_0^\delta\|_{H^\ell(\mathbb{R})} \leq C\delta^{s-\ell}\|u_0\|_{H^s(\mathbb{R})}.$$

Proof. By definition of the $H^\ell(\mathbb{R})$ -norm (1.16), we have, for $s \geq \ell$:

$$\begin{aligned} \|u_0 - u_0^\delta\|_{H^\ell(\mathbb{R})}^2 &= \int_{\mathbb{R}} (1 + |\xi|^2)^\ell |\widehat{u}_0(\xi)|^2 (1 - \chi(\delta\xi))^2 d\xi \\ &= \delta^{2(s-\ell)} \int_{\mathbb{R}} (1 + |\xi|^2)^\ell |\widehat{u}_0(\xi)|^2 \left(\frac{1 - \chi(\delta\xi)}{(\delta\xi)^{s-\ell}} \right)^2 \xi^{2(s-\ell)} d\xi. \end{aligned}$$

Hypothesis on χ implies that $\sup_{z \in \mathbb{R}} \left| \frac{1 - \chi(z)}{z^{s-\ell}} \right| \leq C_2$ with C_2 a constant. Hence, by using the inequality $(1 + |\xi|^2)^\ell |\xi|^{2(s-\ell)} \leq C(1 + |\xi|^2)^s$, with C a constant,

$$\begin{aligned} \|u_0 - u_0^\delta\|_{H^\ell(\mathbb{R})}^2 &\leq \delta^{2(s-\ell)} C C_2^2 \int_{\mathbb{R}} (1 + |\xi|^2)^s |\widehat{u}_0(\xi)|^2 d\xi \\ &\leq C C_2^2 \delta^{2(s-\ell)} \|u_0\|_{H^s(\mathbb{R})}^2. \end{aligned}$$

$\square$

Remark 2.35. We recognize the competition mentioned in general introduction between the regularity and the stiffness of the derivatives. Indeed, the derivation of Lemma 2.33 reveals $\frac{1}{\delta^{m-s}}$ whereas the regularization of Lemma 2.34 discloses $\delta^{s-\ell}$, for $0 \leq \ell \leq s \leq m$. More precise is the regularization, steeper is the derivation. A compromise has to be found to regularize enough to perform Taylor expansions but without needing to many derivatives.

2.2.2.2 Initial datum in $H^s(\mathbb{R})$ with $s \geq 2p + 2$

As explained previously (cf Remark 2.14), the regularity of u_0 is determined by the Taylor expansion in time. A first step is then to deal with the time term in error estimates. The following proposition provides that the time error prevails until $u_0 \in H^{2p+2}(\mathbb{R})$, for which the spatial error becomes predominant.

Proposition 2.36. *Assume $u_0 \in H^s(\mathbb{R})$ with $s \geq 2p + 2$, and let us fix $M = \min(s, 4p + 2)$ (and $M^{cn} = \min(s, 6p + 3)$ if $\theta = \frac{1}{2}$), then there exists a constant $C > 0$ independent of Δt and Δx , such as the error estimate for the right-winded (respectively left-winded) finite difference scheme, if p is odd (respectively even), yields*

$$\|e\|_{\ell^\infty([0, N]; \ell_\Delta^2(\mathbb{Z}))} \leq C \left(\Delta t^{\frac{M}{4p+2}} \left| \frac{1}{2} - \theta \right| \|\partial_x^M u_0\|_{L^2(\mathbb{R})} + \Delta x \|\partial_x^{2p+2} u_0\|_{L^2(\mathbb{R})} + \Delta t^{\frac{2M^{cn}}{6p+3}} \|\partial_x^{M^{cn}} u_0\|_{L^2(\mathbb{R})} \right),$$

under CLF conditions of Table 2.1.

For the central difference scheme, we suppose $s \geq 2p + 3$, and one has (for the same M and M^{cn})

$$\|e\|_{\ell^\infty([0, N]; \ell_\Delta^2(\mathbb{Z}))} \leq C \left(\Delta t^{\frac{M}{4p+2}} \left| \frac{1}{2} - \theta \right| \|\partial_x^M u_0\|_{L^2(\mathbb{R})} + \Delta x^2 \|\partial_x^{2p+3} u_0\|_{L^2(\mathbb{R})} + \Delta t^{\frac{2M^{cn}}{6p+3}} \|\partial_x^{M^{cn}} u_0\|_{L^2(\mathbb{R})} \right),$$

under CFL conditions of Table 2.1.

Section 2.2. High-order linear dispersive equations

Proof of Proposition 2.36. The triangular inequality applied to the convergence error gives $\|e^n\|_{\ell_\Delta^2} = \|u_\Delta - v\|_{\ell_\Delta^2} \leq \Xi_1^n + \Xi_2^n + \Xi_3^n$ with

$$\Xi_1^n = \left(\sum_{j \in \mathbb{Z}} \Delta x \left(\frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u(t^n, x) - u^\delta(t^n, x) dx \right)^2 \right)^{\frac{1}{2}}, \quad (2.29)$$

$$\Xi_2^n = \left(\sum_{j \in \mathbb{Z}} \Delta x \left(\frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u^\delta(t^n, x) dx - (v^\delta)_j^n \right)^2 \right)^{\frac{1}{2}}, \quad (2.30)$$

$$\Xi_3^n = \left(\sum_{j \in \mathbb{Z}} \Delta x \left((v^\delta)_j^n - v_j^n \right)^2 \right)^{\frac{1}{2}}. \quad (2.31)$$

Cauchy–Schwarz inequality together with the conservation of the L^2 -norm (cf Lemma 2.5) yield

$$\Xi_1^n \leq \left(\sum_{j \in \mathbb{Z}} \int_{x_j}^{x_{j+1}} |u(t^n, x) - u^\delta(t^n, x)|^2 dx \right)^{\frac{1}{2}} \leq \|u_0 - u_0^\delta\|_{L^2(\mathbb{R})}.$$

From Lemma 2.34 with $\ell = 0$ and $s = M$ (or $s = M^{cn}$ if $\theta = \frac{1}{2}$), there exists a constant C such that

$$\Xi_1^n \leq C \left(\delta^M \left| \frac{1}{2} - \theta \right| \|\partial_x^M u_0\|_{L^2(\mathbb{R})} + \delta^{M^{cn}} \mathbf{1}_{\{\theta = \frac{1}{2}\}} \|\partial_x^{M^{cn}} u_0\|_{L^2(\mathbb{R})} \right).$$

For the Ξ_2^n -term, we use the previous section (Section 2.2.1). Indeed, Ξ_2^n corresponds to the convergence error for a smooth initial datum u_0^δ . Hence, one has, if CFL conditions of Table 2.1 are true,

$$\begin{aligned} \Xi_2^n &\leq C \left(\Delta t \left| \frac{1}{2} - \theta \right| \|\partial_x^{4p+2} u_0^\delta\|_{L^2(\mathbb{R})} + \Delta x \|\partial_x^{2p+2} u_0^\delta\|_{L^2(\mathbb{R})} + \Delta t^2 \|\partial_x^{6p+3} u_0^\delta\|_{L^2(\mathbb{R})} \right) \\ &\leq C \left(\frac{\Delta t}{\delta^{4p+2-M}} \left| \frac{1}{2} - \theta \right| \|\partial_x^M u_0\|_{L^2(\mathbb{R})} + \Delta x \|\partial_x^{2p+2} u_0\|_{L^2(\mathbb{R})} + \frac{\Delta t^2}{\delta^{6p+3-M^{cn}}} \|\partial_x^{M^{cn}} u_0\|_{L^2(\mathbb{R})} \right), \end{aligned}$$

where the latest inequality comes from Lemma 2.33 with $(m, s) = (4p + 2, M)$, $(m, s) = (2p + 2, 2p + 2)$ and $(m, s) = (6p + 3, M^{cn})$.

Finally, the stability of the scheme gives the following estimate for Ξ_3^n :

$$\Xi_3^n = \|(v^\delta)^n - v^n\|_{\ell_\Delta^2} < \|u_0^\delta - u_0\|_{L^2(\mathbb{R})}.$$

Thus, the convergence error is upper bounded by

$$\begin{aligned} \|e^n\|_{\ell_\Delta^2} &\leq C \left(\delta^M \left| \frac{1}{2} - \theta \right| \|\partial_x^M u_0\|_{L^2(\mathbb{R})} + \delta^{M^{cn}} \mathbf{1}_{\{\theta = \frac{1}{2}\}} \|\partial_x^{M^{cn}} u_0\|_{L^2(\mathbb{R})} \right. \\ &\quad \left. + \frac{\Delta t}{\delta^{4p+2-M}} \left| \frac{1}{2} - \theta \right| \|\partial_x^M u_0\|_{L^2(\mathbb{R})} + \Delta x \|\partial_x^{2p+2} u_0\|_{L^2(\mathbb{R})} + \frac{\Delta t^2}{\delta^{6p+3-M^{cn}}} \|\partial_x^{M^{cn}} u_0\|_{L^2(\mathbb{R})} \right). \end{aligned}$$

Proposition 2.36 comes from the optimal choice for δ : $\delta = \Delta t^{\frac{1}{4p+2}}$ (or $\delta = \Delta t^{\frac{2}{6p+3}}$ if $\theta = \frac{1}{2}$). $\square$

Remark 2.37. The result for the central finite difference scheme is proved exactly in the same way, with the same s , m , ℓ and δ .

Remark 2.38. For $p = 0$, $\theta \neq \frac{1}{2}$ and the central finite difference scheme, we have to perform the same guidelines while exchanging the role of Δt and Δx . The choices of $\delta = \Delta x^{\frac{4}{3}}$ and $M = 2$ give :

$$\|e^n\|_{\ell^\infty([0, N]; \ell_\Delta^2(\mathbb{Z}))} \leq C \left(\Delta x^{\frac{4}{3}} \|\partial_x^2 u_0\|_{L^2(\mathbb{R})} + \Delta t \|\partial_x^2 u_0\|_{L^2(\mathbb{R})} \right).$$

2.2.2.3 Initial datum in $H^s(\mathbb{R})$ with $s \geq 0$

The main result of this chapter is summarized in the following theorem where the initial datum u_0 has any Sobolev regularity.

Theorem 2.39.

Assume that CFL conditions of Table 2.1 are true. Assume that $u_0 \in H^s(\mathbb{R})$ with $0 \leq s$, then, there exists a constant $C > 0$ independent of Δt and Δx , such that the right-winded (respectively left-winded) finite difference scheme if p is odd (respectively even) has the following error-estimate

$$\|e\|_{\ell^\infty(\llbracket 0, N \rrbracket; \ell_\Delta^2(\mathbb{Z}))} \leq C \left(\Delta t^{\frac{\min(s, 4p+2)}{4p+2}} \left| \frac{1}{2} - \theta \right| \|\partial_x^{\min(4p+2, s)} u_0\|_{L^2(\mathbb{R})} + \Delta x^{\frac{\min(s, 2p+2)}{2p+2}} \|\partial_x^{\min(2p+2, s)} u_0\|_{L^2(\mathbb{R})} + \Delta t^2 \frac{\min(s, 6p+3)}{6p+3} \|\partial_x^{\min(6p+3, s)} u_0\|_{L^2(\mathbb{R})} \right).$$

The previous inequality becomes for the central finite difference scheme

$$\|e\|_{\ell^\infty(\llbracket 0, N \rrbracket; \ell_\Delta^2(\mathbb{Z}))} \leq C \left(\Delta t^{\frac{\min(s, 4p+2)}{4p+2}} \left| \frac{1}{2} - \theta \right| \|\partial_x^{\min(4p+2, s)} u_0\|_{L^2(\mathbb{R})} + \Delta x^2 \frac{\min(s, 2p+3)}{2p+3} \|\partial_x^{\min(2p+3, s)} u_0\|_{L^2(\mathbb{R})} + \Delta t^2 \frac{\min(s, 6p+3)}{6p+3} \|\partial_x^{\min(6p+3, s)} u_0\|_{L^2(\mathbb{R})} \right).$$

Remark 2.40. The previous results are summarized in Table 2.6.

Proof of Theorem 2.39. Here again, we suppose that p is odd, we thus detail the proof for the right-winded finite difference scheme. We have already proved the case $s \geq 4p+2$ in Section 2.2.1 and the case $2p+2 \leq s \leq 4p+2$ in Subsection 2.2.2.2. Let us now focus on the case $0 \leq s \leq 2p+2$.

The proof of Theorem 2.39 follows the same guidelines as the proof of Proposition 2.36. Let $u_0 \in H^s(\mathbb{R})$, we regularize this initial datum thanks to mollifiers $(\varphi^\delta)_{\delta>0}$ whose properties are listed in Subsection 2.2.2.1. This involves introducing the same new unknowns u^δ , u_0^δ and $((v^\delta)_j^n)_{(j,n)}$.

CASE $\theta \neq \frac{1}{2}$: The convergence error $(e_j^n)_{(j,n)}$ is upper bounded by the same Ξ_1^n , Ξ_2^n and Ξ_3^n , defined in (2.29)-(2.31). Lemma 2.34 with $\ell = 0$ and $s = s$ leads to

$$\Xi_1^n + \Xi_3^n \leq C \delta^s \|\partial_x^s u_0\|_{L^2(\mathbb{R})}.$$

By definition $u_0^\delta \in H^k(\mathbb{R})$, $\forall k > 0$, therefore Proposition 2.36 applies with $k = 2p+2$ and $M = \min(k, 4p+2) = 2p+2$. It gives the following estimate for Ξ_2^n :

$$\Xi_2^n \leq C \left(\Delta t^{\frac{p+1}{2p+1}} \|\partial_x^{2p+2} u_0^\delta\|_{L^2(\mathbb{R})} + \Delta x \|\partial_x^{2p+2} u_0^\delta\|_{L^2(\mathbb{R})} \right) \leq C \left(\left(\Delta t^{\frac{p+1}{2p+1}} + \Delta x \right) \|\partial_x^{2p+2} u_0^\delta\|_{L^2(\mathbb{R})} \right).$$

We then apply Lemma 2.33 with $m = 2p+2$ and $s = s$. Finally, it yields

$$\|e^n\|_{\ell_\Delta^2} \leq C \left(\delta^s \|\partial_x^s u_0\|_{L^2(\mathbb{R})} + \frac{\left(\Delta t^{\frac{p+1}{2p+1}} + \Delta x \right)}{\delta^{2p+2-s}} \|\partial_x^s u_0\|_{L^2(\mathbb{R})} \right).$$

The conclusion comes from the optimal choice $\delta = \left(\Delta t^{\frac{p+1}{2p+1}} + \Delta x \right)^{\frac{1}{2p+2}}$.

Section 2.2. High-order linear dispersive equations

CASE $\theta = \frac{1}{2}$: The Crank-Nicolson case leads to

$$\|e^n\|_{\ell_\Delta^2} \leq C \left(\delta^s \|\partial_x^s u_0\|_{L^2(\mathbb{R})} + \frac{\left(\Delta t^{\frac{4p+4}{6p+3}} + \Delta x \right)}{\delta^{2p+2-s}} \|\partial_x^s u_0\|_{L^2(\mathbb{R})} \right).$$

The optimal choice is in this case $\delta = \left(\Delta t^{\frac{4p+4}{6p+3}} + \Delta x \right)^{\frac{1}{2p+2}}$. $\square$

Remark 2.41. The left-winded scheme, with p even, is exactly the same. The central finite difference scheme is proved with the same method except for the variable k , which is taken $k = 2p + 3$ for that scheme and $M = \min(k, 4p + 2) = 2p + 3$ (if $p \neq 0$). Thus, the upper bound for Ξ_2^n becomes, (if $\theta \neq \frac{1}{2}$)

$$\Xi_2^n \leq C \left(\Delta t^{\frac{2p+3}{4p+2}} \|\partial_x^{2p+3} u_0^\delta\|_{L^2(\mathbb{R})} + \Delta x^2 \|\partial_x^{2p+3} u_0^\delta\|_{L^2(\mathbb{R})} \right) \leq C \left(\left(\Delta t^{\frac{2p+3}{4p+2}} + \Delta x^2 \right) \|\partial_x^{2p+3} u_0^\delta\|_{L^2(\mathbb{R})} \right).$$

This implies (if $\theta \neq \frac{1}{2}$)

$$\|e^n\|_{\ell_\Delta^2} \leq C \left(\delta^s \|\partial_x^s u_0\|_{L^2(\mathbb{R})} + \frac{\left(\Delta t^{\frac{2p+3}{4p+2}} + \Delta x^2 \right)}{\delta^{2p+3-s}} \|\partial_x^s u_0\|_{L^2(\mathbb{R})} \right)$$

and (if $\theta = \frac{1}{2}$)

$$\|e^n\|_{\ell_\Delta^2} \leq C \left(\delta^s \|\partial_x^s u_0\|_{L^2(\mathbb{R})} + \frac{\left(\Delta t^{\frac{2p+3}{6p+3}} + \Delta x^2 \right)}{\delta^{2p+3-s}} \|\partial_x^s u_0\|_{L^2(\mathbb{R})} \right).$$

Remark 2.42. For $p = 0$, $\theta \neq \frac{1}{2}$ and the central scheme, the same holds to obtain

$$\|e^n\|_{\ell_\Delta^2} \leq \left(\delta^s + \frac{\Delta t + \Delta x^{\frac{4}{3}}}{\delta^{2-s}} \right) \|\partial_x^s u_0\|_{L^2(\mathbb{R})}$$

and we conclude with $\delta = \left(\Delta t + \Delta x^{\frac{4}{3}} \right)^{\frac{1}{2}}$

2.2.3 Numerical results

In order to illustrate numerically the previous results, we perform two sets of examples : on the one hand, we compute the numerical rate of convergence of various equations for a fix initial datum and on the other hand, the equation is fixed and we test different initial data.

In all examples, the computational domain is set to $[0, 50]$ subdivided into J cells with $J \in \{800, 1600, 3200, 6400, 12800, 25600, 51200, 102400\}$ and the numerical simulation is performed up to time $T = 5$. Not to have a too restricted Courant-Friedrichs-Lewy condition, we implement the implicit scheme ($\theta = 1$) and impose $\Delta t = \Delta x$.

Remark 2.43. To study convergence of the scheme, mesh size aims to be smaller and smaller : Δx^{2p+1} is so incredibly small. Therefore, dispersive CFL condition is a very restrictive condition, unworkable in most cases. However, it turns on only for $\theta < \frac{1}{2}$. That is why, all our numerical experiments will be carry out with $\theta \geq \frac{1}{2}$.

Since exact solutions are not always known, the convergence error is computed between the solution with J cells and a 'reference' solution with $2J$ cells in space.

Since the indicator function $\mathbb{1}_{[0, \frac{1}{2}]}$ belongs to $H^s([0, L])$ for all $s < \frac{1}{2}$, we build test functions in $H^s([0, L])$ with $s < \frac{1}{2} + k$ by integrating k -times the indicator function. Such functions will be denoted in $H^{\frac{1}{2}+k-}([0, L])$. For more details about this construction see Section 3.7.

The first test consists of fixing u_0 in $H^{\frac{3}{2}-}([0, L])$ or $H^{\frac{5}{2}-}([0, L])$ and compute the convergence rate for $p = 0$ (advection equation), $p = 1$ (Airy equation) and $p = 2$. The numerical results are gathered in Tables 2.2 and 2.3 and correctly match with the expected theoretical rates. We choose right-winded or left-winded schemes according to p .

Δx	$p = 0$ (Advection)		$p = 1$ (Airy)		$p = 2$ (Fifth-order derivative)	
	error in $\ell^\infty(0, N; \ell_\Delta^2(\mathbb{Z}))$	order	error in $\ell^\infty(0, N; \ell_\Delta^2(\mathbb{Z}))$	order	error in $\ell^\infty(0, N; \ell_\Delta^2(\mathbb{Z}))$	order
6.250.10 ⁻²					2.985.10 ⁻³	
3.125.10 ⁻²					2.757.10 ⁻³	0.115
1.563.10 ⁻²					2.441.10 ⁻³	0.175
7.813.10 ⁻³	1.194.10 ⁻⁴		1.348.10 ⁻³		2.176.10 ⁻³	0.166
3.906.10 ⁻³	7.381.10 ⁻⁵	0.694	1.125.10 ⁻³	0.261	1.961.10 ⁻³	0.149
1.953.10 ⁻³	4.471.10 ⁻⁵	0.723	9.670.10 ⁻⁴	0.219		
9.766.10 ⁻⁴	2.664.10 ⁻⁵	0.747	7.968.10 ⁻⁴	0.279		
4.883.10 ⁻⁴	1.585.10 ⁻⁵	0.749	6.572.10 ⁻⁴	0.278		
theoretical		0.750		0.250		0.150

 TABLE 2.2 – For a Sobolev regularity $H^{\frac{3}{2}-}$

Δx	$p = 0$ (Advection)		$p = 1$ (Airy)		$p = 2$ (Fifth-order derivative)	
	error in $\ell^\infty(0, N; \ell_\Delta^2(\mathbb{Z}))$	order	error in $\ell^\infty(0, N; \ell_\Delta^2(\mathbb{Z}))$	order	error in $\ell^\infty(0, N; \ell_\Delta^2(\mathbb{Z}))$	order
6.250.10 ⁻²					3.388.10 ⁻²	
3.125.10 ⁻²					3.639.10 ⁻²	0.093
1.563.10 ⁻²					3.032.10 ⁻²	0.263
7.813.10 ⁻³	2.586.10 ⁻³		1.011.10 ⁻²		2.528.10 ⁻²	0.262
3.906.10 ⁻³	1.347.10 ⁻³	0.940	7.507.10 ⁻³	0.430	2.138.10 ⁻²	0.242
1.953.10 ⁻³	6.873.10 ⁻⁴	0.971	6.267.10 ⁻³	0.261		
9.766.10 ⁻⁴	3.437.10 ⁻⁴	0.999	4.401.10 ⁻³	0.510		
4.883.10 ⁻⁴	1.719.10 ⁻⁴	1.000	3.074.10 ⁻³	0.520		
theoretical		1.000		0.417		0.250

 TABLE 2.3 – For a Sobolev regularity $H^{\frac{5}{2}-}$

For the second sample of examples, the equation is fixed ($p = 0$ for Subfigure 2.1(a) and $p = 1$ for Subfigure 2.1(b)) whereas the Sobolev regularity of the initial data is fluctuating. As shown in Figure 2.1, the theoretical rates are represented by the blue line and the numerical rates correspond to the red dots. The exponent of the Sobolev regularity of u_0 is shown in the x-axis. Again, the different rates match very well, which tends to indicate that the convergence orders we have proven are optimal.

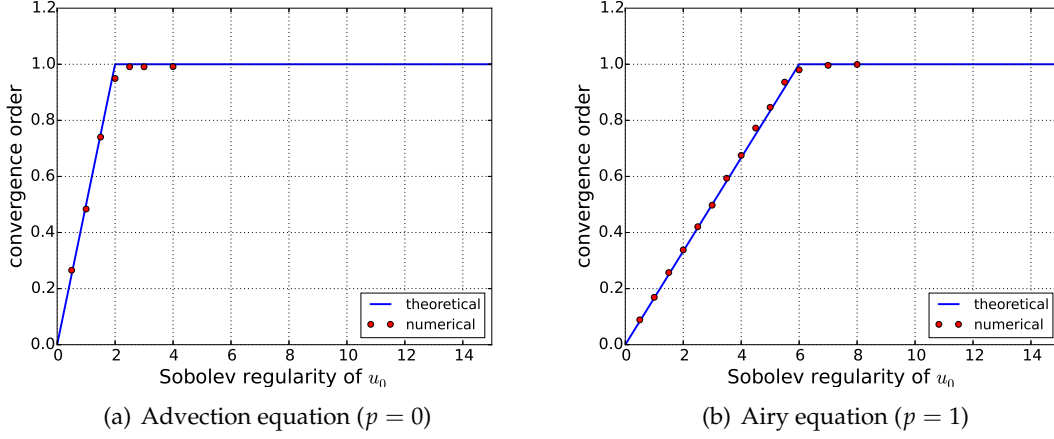


FIGURE 2.1 – Numerical versus theoretical orders

2.3 High-order polyharmonic heat equations

The high-order polyharmonic heat equation (2.3) is discretized according to the centered scheme (2.6a)-(2.6b) with (2.7). We follow exactly the same guidelines of the previous section to establish the convergence error : a first step consists on stating consistency estimates and we set the stability CFL-type condition in a second step. We conclude the proof by a discussion on the convergence rate according to the Sobolev regularity on the initial datum.

2.3.1 For an initial datum in $H^{4p}(\mathbb{R})$

2.3.1.1 Consistency estimate

For more simplicity in proofs of this section, we establish a link between the even case and the odd case in the following lemma.

Lemma 2.44. *Let $(a_j^n)_{(n,j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}}$ be a sequence, for all integer p , one has*

$$\Delta x \left(D_c^{2p} a \right)_j^n = \left(D_+^{2p-1} a \right)_j^n - \left(D_-^{2p-1} a \right)_j^n, \quad (n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}.$$

Proof. By definition of odd operators D_+^{2p-1} and D_-^{2p-1} (cf Relations (2.5a) and (2.5b)), it yields

$$\left(D_+^{2p-1} a \right)_j^n - \left(D_-^{2p-1} a \right)_j^n = \frac{1}{\Delta x^{2p-1}} \left[\sum_{k=0}^{2p-1} \binom{2p-1}{k} (-1)^k a_{p-k+j}^n - \sum_{k=0}^{2p-1} \binom{2p-1}{k} (-1)^k a_{p-k+j-1}^n \right].$$

The change of variable $k \rightarrow k+1$ in the second sum enables us to sum up the terms

$$\left(D_+^{2p-1} a \right)_j^n - \left(D_-^{2p-1} a \right)_j^n = \frac{1}{\Delta x^{2p-1}} \left[\sum_{k=1}^{2p-1} \left[\binom{2p-1}{k} + \binom{2p-1}{k-1} \right] (-1)^k a_{p-k+j}^n + a_{p+j}^n + a_{-p+j}^n \right].$$

We conclude due to the identity $\binom{2p-1}{k} + \binom{2p-1}{k-1} = \binom{2p}{k}$. $\square$

We are able now to set the consistency estimate for the polyharmonic heat equation.

Proposition 2.45. Assume $u_0 \in H^{4p}(\mathbb{R})$ (and $u_0 \in H^{6p}(\mathbb{R})$ if $\theta = \frac{1}{2}$), then there exists a constant $C > 0$ independent of Δt and Δx such that the consistency error verifies

$$\|\epsilon\|_{\ell^\infty(\llbracket 0, N \rrbracket; \ell^2_\Delta(\mathbb{Z}))} \leq C \left(\Delta t \left| \frac{1}{2} - \theta \right| \|\partial_x^{4p} u_0\|_{L^2(\mathbb{R})} + \Delta x^2 \|\partial_x^{2p+2} u_0\|_{L^2(\mathbb{R})} + \Delta t^2 \|\partial_x^{6p} u_0\|_{L^2(\mathbb{R})} \right).$$

Proof. It should be recall that the consistency error is defined in Definition 2.3. We first focus on the space part, by a Taylor expansion (in space) up to order $2p + 2$ and an exchange of the two sums

$$\begin{aligned} (D_c^{2p} u_\Delta)_j^n &= \sum_{k=0}^{2p} \frac{\binom{2p}{k} (-1)^k}{\Delta x^{2p}} \left(\frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u(t^n, y + (p-k)\Delta x) dy \right) \\ &= \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} \left(\sum_{\ell=0}^{2p} \frac{\partial_x^\ell u(t^n, y)}{\ell!} \frac{\left(\sum_{k=0}^{2p} \binom{2p}{k} (-1)^k (p-k)^\ell \Delta x^\ell \right)}{\Delta x^{2p}} \right. \\ &\quad \left. + \frac{\partial_x^{2p+1} u(t^n, y)}{(2p+1)!} \frac{\left(\sum_{k=0}^{2p} \binom{2p}{k} (-1)^k (p-k)^{2p+1} \Delta x^{2p+1} \right)}{\Delta x^{2p}} \right) dy + (R_c)_j^n \end{aligned}$$

with

$$(R_c)_j^n = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} \int_y^{y+(p-k)\Delta x} \frac{\partial_x^{2p+2} u(t^n, z)}{(2p+1)!} \frac{\sum_{k=0}^{2p} \binom{2p}{k} (-1)^k (y + (p-k)\Delta x - z)^{2p+1}}{\Delta x^{2p}} dz dy.$$

Lemma 2.9 simplifies the first sum :

$$(D_c^{2p} u_\Delta)_j^n = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} \partial_x^{2p} u(t^n, y) dy + \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} \frac{\partial_x^{2p+1} u(t^n, y)}{(2p+1)!} \frac{\left(\sum_{k=0}^{2p} \binom{2p}{k} (-1)^k (p-k)^{2p+1} \Delta x^{2p+1} \right)}{\Delta x^{2p}} dy + (R_c)_j^n.$$

Lemma 2.44 (with $a_k = k^{2p+1}$, $\Delta x = 1$ and $j = 0$) is used for the second sum

$$\sum_{k=0}^{2p} \binom{2p}{k} (-1)^k (p-k)^{2p+1} = \sum_{k=0}^{2p-1} \binom{2p-1}{k} (-1)^k (p-k)^{2p+1} - \sum_{k=0}^{2p-1} \binom{2p-1}{k} (-1)^k (p-k-1)^{2p+1}.$$

A change of variable $k \rightarrow 2p-1-k$ in the first sum simplifies the previous expression

$$\begin{aligned} \sum_{k=0}^{2p} \binom{2p}{k} (-1)^k (p-k)^{2p+1} &= (-1)^{2p+2} \sum_{k=0}^{2p-1} \binom{2p-1}{k} (-1)^k (p-k-1)^{2p+1} \\ &\quad - \sum_{k=0}^{2p-1} \binom{2p-1}{k} (-1)^k (p-k-1)^{2p+1} \\ &= 0. \end{aligned}$$

Let us now focus on the time term of the consistency error. A Taylor expansion (in time) gives us

$$\begin{aligned} \epsilon_j^n &= \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} \partial_t u(t^n, y) dy + \frac{(-1)^p}{\Delta x} \int_{x_j}^{x_{j+1}} \partial_x^{2p} u(t^n, y) dy + \frac{|\theta - \frac{1}{2}|}{\Delta x \Delta t} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \partial_t^2 u(s, y) (t^{n+1} - s) ds dy \\ &\quad + \mathbb{1}_{\{\theta = \frac{1}{2}\}} \frac{\Delta t}{2 \Delta x} \int_{x_j}^{x_{j+1}} \partial_t^2 u(t^n, y) dy + \mathbb{1}_{\{\theta = \frac{1}{2}\}} \frac{1}{\Delta x \Delta t} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \partial_t^3 u(s, y) \frac{(t^{n+1} - s)^2}{2} ds dy \\ &\quad + \frac{(-1)^p \theta \Delta t}{\Delta x} \int_{x_j}^{x_{j+1}} \partial_x^{2p} \partial_t u(t^n, y) dy + \frac{(-1)^p \theta}{\Delta x} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \partial_x^{2p} \partial_t^2 u(s, y) (t^{n+1} - s) ds dy \\ &\quad + \theta (R_c)_j^{n+1} + (1 - \theta) (R_c)_j^n. \end{aligned}$$

We recognize the polyharmonic heat equation in the two first terms. For the other, we square the expression and use Cauchy-Schwarz inequality. As for the odd case, there exists a constant C such that $\|(R_c^n)_j\|_{\ell^2_\Delta}^2 \leq C \Delta x^2 \|\partial_x^{2p+2} u(t^n, \cdot)\|_{L^2(\mathbb{R})}^2$. Proposition 2.45 is then proved. $\square$

2.3.1.2 Stability

Proposition 2.46. For small Δt and Δx , the scheme (2.6a) is stable under the CFL condition

$$\Delta t(1 - 2\theta) < \frac{\Delta x^{2p}}{2^{2p-1}}. \quad (2.32)$$

Once again, we adapt some lemmas of Section 2.2 to the case $\mathcal{P} = 2p$. For example, the equivalent of Lemma 2.22 is the following one.

Lemma 2.47. For all $\xi \in [0, 1]$ and $p \in \mathbb{N}$, it holds

$$\sum_{k=0}^{2p} \binom{2p}{k} (-1)^k e^{-2i\pi\xi(p-k)} = 2^{2p} (-1)^p \sin^{2p}(\pi\xi).$$

Proof of Lemma 2.47. We develop the sum to simplify

$$\begin{aligned} \sum_{k=0}^{2p} \binom{2p}{k} (-1)^k e^{-2i\pi\xi(p-k)} &= e^{2i\pi\xi p} \sum_{k=0}^{2p} \binom{2p}{k} (-1)^k (e^{-2i\pi\xi})^{2p-k} = e^{2i\pi\xi p} (e^{-2i\pi\xi} - 1)^{2p} \\ &= e^{2i\pi\xi p} [e^{-i\pi\xi} (e^{-i\pi\xi} - e^{i\pi\xi})]^{2p} = 2^{2p} (-1)^p \sin^{2p}(\pi\xi). \end{aligned}$$

□

Proof of Proposition 2.46. We use the Fourier transform as in the case $\mathcal{P} = 2p + 1$, to obtain

$$\widehat{v}^{n+1} \left[1 + (-1)^p \theta \frac{\Delta t}{\Delta x^{2p}} \sum_{k=0}^{2p} \binom{2p}{k} (-1)^k e^{-2i\pi\xi(p-k)} \right] = \widehat{v}^n \left[1 - (-1)^p (1 - \theta) \frac{\Delta t}{\Delta x^{2p}} \sum_{k=0}^{2p} \binom{2p}{k} (-1)^k e^{-2i\pi\xi(p-k)} \right].$$

Thanks to Lemma 2.47, the previous identity is simplified into

$$\widehat{v}^{n+1} \left[1 + \theta \frac{\Delta t}{\Delta x^{2p}} 2^{2p} \sin^{2p}(\pi\xi) \right] = \widehat{v}^n \left[1 - (1 - \theta) \frac{\Delta t}{\Delta x^{2p}} 2^{2p} \sin^{2p}(\pi\xi) \right].$$

The amplification coefficient A_c^{2p} is defined for all $\xi \in [0, 1]$ such as

$$A_c^{2p}(\xi) = \frac{1 - (1 - \theta) \frac{\Delta t}{\Delta x^{2p}} 2^{2p} \sin^{2p}(\pi\xi)}{1 + \theta \frac{\Delta t}{\Delta x^{2p}} 2^{2p} \sin^{2p}(\pi\xi)}.$$

As for the case $\mathcal{P} = 2p + 1$, we are looking for conditions ensuring $|A_c^{2p}(\xi)|^2 \leq (1 + C\Delta t)^2$ for any $\xi \in [0, 1]$. After developing the amplification coefficient, this conditions must ensure

$$\begin{aligned} 1 - 2^{2p+1} (1 - \theta) \frac{\Delta t}{\Delta x^{2p}} \sin^{2p}(\pi\xi) + (1 - 2\theta + \theta^2) \frac{\Delta t^2}{\Delta x^{4p}} 2^{4p} \sin^{4p}(\pi\xi) \\ \leq (1 + 2C\Delta t + C^2\Delta t^2) \left[1 + 2^{2p+1} \theta \frac{\Delta t}{\Delta x^{2p}} \sin^{2p}(\pi\xi) + \theta^2 \frac{\Delta t^2}{\Delta x^{4p}} 2^{4p} \sin^{4p}(\pi\xi) \right]. \end{aligned}$$

That means

$$\begin{aligned} \Delta t 2^{4p} \left[(1 - 2\theta) \sin^{4p}(\pi\xi) - \frac{C\theta}{2^{2p-2}} \Delta x^{2p} \sin^{2p}(\pi\xi) - 2C\theta^2 \Delta t \sin^{4p}(\pi\xi) - \frac{C^2}{2^{4p}} \Delta x^{4p} \right. \\ \left. - \frac{C^2 \theta \Delta x^{2p} \Delta t}{2^{2p-1}} \sin^{2p}(\pi\xi) - \theta^2 C^2 \Delta t^2 \sin^{4p}(\pi\xi) \right] \leq 2^{2p+1} \Delta x^{2p} \sin^{2p}(\pi\xi) + 2C\Delta x^{4p}. \end{aligned}$$

For Δt and Δx asymptotically small, this condition is equivalent to the following CFL-type condition

$$(1 - 2\theta)\Delta t < \frac{\Delta x^{2p}}{2^{2p-1}}.$$

□

2.3.1.3 Convergence error

The equivalent result of Theorem 2.30 is as follows.

Theorem 2.48.

For an initial datum $u_0 \in H^{4p}(\mathbb{R})$ (and $u_0 \in H^{6p}(\mathbb{R})$, if $\theta = \frac{1}{2}$), the error estimate of the finite difference scheme (2.6a) satisfies

$$\|e\|_{\ell^\infty([0,N];\ell_\Delta^2(\mathbb{Z}))} \leq C \left(\Delta t \left| \frac{1}{2} - \theta \right| \left\| \partial_x^{4p} u_0 \right\|_{L^2(\mathbb{R})} + \Delta x^2 \left\| \partial_x^{2p+2} u_0 \right\|_{L^2(\mathbb{R})} + \Delta t^2 \left\| \partial_x^{6p} u_0 \right\|_{L^2(\mathbb{R})} \right),$$

under the CFL condition (2.32).

Proof of Theorem 2.48. The same milestones as for the previous case (cf Section 2.2) apply. The convergence error verifies

$$\widehat{e^{n+1}}(\xi) = A_c^{2p}(\xi) \widehat{e^n}(\xi) - \frac{\Delta t}{1 + \frac{\theta \Delta t}{\Delta x^{2p}} (2\sin(\xi))^{2p}} \widehat{e^n}(\xi), \quad \forall \xi \in [0, 1].$$

The stability condition $\|A_c^{2p}(\xi)\|_{L^\infty([0,1])} \leq 1 + C\Delta t$, the upper bound $\left\| \frac{1}{1 + \frac{\theta \Delta t}{\Delta x^{2p}} (2\sin(\xi))^{2p}} \right\|_{L^\infty([0,1])} \leq 1$ together with the triangular inequality lead to

$$\|\widehat{e^{n+1}}\|_{L^2([0,1])} \leq \|\widehat{e^n}\|_{L^2([0,1])} (1 + C\Delta t) + \Delta t \|\widehat{e^n}\|_{L^2([0,1])}.$$

Thanks to Parseval identity (2.17) and Grönwall Lemma 2.29, the convergence error satisfies

$$\|e^n\|_{\ell_\Delta^2} \leq e^{CT} \|e^0\|_{\ell_\Delta^2} + T e^{CT} \sup_{k \in \llbracket 0, N \rrbracket} \|e^k\|_{\ell_\Delta^2}.$$

The conclusion of Theorem 2.48 is a straightforward consequence. $\square$

2.3.2 For a less smooth initial datum

We relax the strong regularity of u_0 and follow exactly the same key stones of previous section on $\mathcal{P}$ odd. As previously, we adopt a two-stage approach : first we weaken the regularity until $H^{2p+2}(\mathbb{R})$ in order to equalize the needed regularity in Δt -term and in Δx -term. We reduce only the regularity in Δt -term because it is the most demanding term. Then, we lower both regularity in Δt - and Δx -term until $L^2(\mathbb{R})$. Let us start with step one : a $H^{2p+2}(\mathbb{R})$ -reduction.

Proposition 2.49. Assume $u_0 \in H^s(\mathbb{R})$ with $s \geq 2p + 2$ and fix $M = \min(s, 4p)$ (and $M^{cn} = \min(s, 6p)$ for Crank-Nicolson case, $\theta = \frac{1}{2}$), then there exists a constant $C > 0$ independent of Δt and Δx , such that the following error estimate holds true, under CFL condition (2.32).

$$\|e\|_{\ell^\infty(0,N;\ell_\Delta^2(\mathbb{Z}))} \leq C \left(\Delta t^{\frac{M}{4p}} \left| \frac{1}{2} - \theta \right| \left\| \partial_x^M u_0 \right\|_{L^2(\mathbb{R})} + \Delta x^2 \left\| \partial_x^{2p+2} u_0 \right\|_{L^2(\mathbb{R})} + \Delta t^{\frac{M^{cn}}{3p}} \left\| \partial_x^{M^{cn}} u_0 \right\|_{L^2(\mathbb{R})} \right).$$

Proof. We regularize u_0 with the mollifiers introduced in Section 2.2.2.1 and keep the same notations : u_0^δ consists on the regularized initial datum, u^δ the exact solution from u_0^δ and $([v^\delta]_j^n)_{(n,j)}$ the numerical solution from $(u_0^\delta)_\Delta$.

The triangular inequality give $\|e^n\|_{\ell_\Delta^2} \leq \Xi_1^n + \Xi_2^n + \Xi_3^n$, with Ξ_i^n defined in (2.29)-(2.31). The upper-bound Ξ_1^n and Ξ_3^n are unchanged with respect to Section 2.2.2.2 :

$$\Xi_i^n \leq C \left(\delta^M \left| \frac{1}{2} - \theta \right| \left\| \partial_x^M u_0 \right\|_{L^2(\mathbb{R})} + \delta^{M^{cn}} \mathbb{1}_{\{\theta=\frac{1}{2}\}} \left\| \partial_x^{M^{cn}} u_0 \right\|_{L^2(\mathbb{R})} \right), \quad \text{for } i \in \{1, 3\}.$$

Section 2.3. High-order polyharmonic heat equations

Upper bound Ξ_2^n holds from Theorem 2.48, to become, under CFL condition (2.32),

$$\Xi_2^n \leq C \left(\frac{\Delta t}{\delta^{4p-M}} \left| \frac{1}{2} - \theta \right| \left\| \partial_x^M u_0 \right\|_{L^2(\mathbb{R})} + \Delta x^2 \left\| \partial_x^{2p+2} u_0 \right\|_{L^2(\mathbb{R})} + \frac{\Delta t^2}{\delta^{6p-M^{cn}}} \left\| \partial_x^{M^{cn}} u_0 \right\|_{L^2(\mathbb{R})} \right).$$

The optimal choice for δ is

$$\delta := \Delta t^{\frac{1}{4p}} \quad (\text{or } \delta := \Delta t^{\frac{1}{3p}} \text{ for Crank-Nicolson case}).$$

□

In the second step, we are now able to outline our main result for the $\mathcal{P}$ even case : the convergence rate for any regularity of u_0 .

Theorem 2.50.

Assume that $u_0 \in H^s(\mathbb{R})$ with $0 \leq s$, then there exists a constant $C > 0$ independent of Δt and Δx such that, under CFL condition (2.32)

$$\begin{aligned} \|e\|_{\ell^\infty(0,N;\ell_\Delta^2(\mathbb{Z}))} \leq C \left(\Delta t^{\frac{\min(s,4p)}{4p}} \left| \frac{1}{2} - \theta \right| \left\| \partial_x^{\min(s,4p)} u_0 \right\|_{L^2(\mathbb{R})} + \Delta x^{2\frac{\min(s,2p+2)}{2p+2}} \left\| \partial_x^{\min(s,2p+2)} u_0 \right\|_{L^2(\mathbb{R})} \right. \\ \left. + \Delta t^{2\frac{\min(s,6p)}{6p}} \left\| \partial_x^{\min(s,6p)} u_0 \right\|_{L^2(\mathbb{R})} \right). \end{aligned}$$

Proof. For $0 \leq s \leq 2p+2$, following the same guidelines as for the proof of Theorem 2.39, one has

$$\|e^n\|_{\ell_\Delta^2} \leq C \left(\delta^s \left\| \partial_x^s u_0 \right\|_{L^2(\mathbb{R})} + \frac{\left(\Delta t^{\frac{2p+2}{4p}} + \Delta x^2 \right)}{\delta^{2p+2-s}} \left\| \partial_x^s u_0 \right\|_{L^2(\mathbb{R})} \right), \quad \text{for } \theta \neq \frac{1}{2},$$

and

$$\|e^n\|_{\ell_\Delta^2} \leq C \left(\delta^s \left\| \partial_x^s u_0 \right\|_{L^2(\mathbb{R})} + \frac{\left(\Delta t^{\frac{2p+2}{3p}} + \Delta x^2 \right)}{\delta^{2p+2-s}} \left\| \partial_x^s u_0 \right\|_{L^2(\mathbb{R})} \right), \quad \text{for Crank-Nicolson case.}$$

We perform the following choices for δ

$$\delta = \left(\Delta t^{\frac{2p+2}{4p}} + \Delta x^2 \right)^{\frac{1}{2p+2}}, \quad \text{for } \theta \neq \frac{1}{2},$$

and

$$\delta = \left(\Delta t^{\frac{2p+2}{3p}} + \Delta x^2 \right)^{\frac{1}{2p+2}}, \quad \text{for Crank-Nicolson case.}$$

□

2.3.3 Numerical results

For the numerical simulations, we have fixed the final time $T = 5$ and the space domain $[0, L]$ with $L = 50$, discretized into J cells with

$$J \in \{800, 1600, 3200, 6400, 12800, 25600\}.$$

Not to have a too restrictive CFL condition we use the implicit scheme ($\theta = 1$) or the Crank-Nicolson scheme ($\theta = \frac{1}{2}$), and fix $\Delta t = \Delta x$. To find initial data in $H^s(\mathbb{R})$, we have used the same construction than in Section 2.2.3 (for more details, see Section 3.7 too).

Exactly as before, we perform two types of comparison. The first one consists of fixing the regularity of the initial datum and testing it on the heat equation (with $p = 1$) and the fourth-order equation ($p = 2$). We obtain Tables 2.4 and 2.5 for a Crank-Nicolson scheme $\theta = \frac{1}{2}$.

Δx	$p = 1$ (Heat equation)		$p = 2$ (Fourth-order)	
	error in $\ell^\infty(0, N; \ell_\Delta^2(\mathbb{Z}))$	order	error in $\ell^\infty(0, N; \ell_\Delta^2(\mathbb{Z}))$	order
$6.250 \cdot 10^{-2}$	$1.530 \cdot 10^{-3}$		$8.780 \cdot 10^{-3}$	
$3.125 \cdot 10^{-2}$	$8.656 \cdot 10^{-4}$	0.822	$6.706 \cdot 10^{-3}$	0.389
$1.563 \cdot 10^{-2}$	$4.984 \cdot 10^{-4}$	0.796	$5.153 \cdot 10^{-3}$	0.380
$7.813 \cdot 10^{-3}$	$2.909 \cdot 10^{-4}$	0.776	$3.969 \cdot 10^{-3}$	0.377
$3.906 \cdot 10^{-3}$	$1.713 \cdot 10^{-4}$	0.764	$3.059 \cdot 10^{-3}$	0.376
$1.953 \cdot 10^{-3}$	$1.013 \cdot 10^{-4}$	0.757	$2.359 \cdot 10^{-3}$	0.375
theoretical		0.500		0.250

TABLE 2.4 – For a Sobolev regularity $H^{\frac{3}{2}-}$. Case $\mathcal{P}$ even

Δx	$p = 1$ (Heat equation)		$p = 2$ (Fourth-order)	
	error in $\ell^\infty(0, N; \ell_\Delta^2(\mathbb{Z}))$	order	error in $\ell^\infty(0, N; \ell_\Delta^2(\mathbb{Z}))$	order
$6.250 \cdot 10^{-2}$	$2.977 \cdot 10^{-3}$		$4.727 \cdot 10^{-2}$	
$3.125 \cdot 10^{-2}$	$1.277 \cdot 10^{-3}$	1.221	$3.073 \cdot 10^{-2}$	0.622
$1.563 \cdot 10^{-2}$	$5.416 \cdot 10^{-4}$	1.237	$1.995 \cdot 10^{-2}$	0.624
$7.813 \cdot 10^{-3}$	$2.287 \cdot 10^{-4}$	1.244	$1.294 \cdot 10^{-2}$	0.625
$3.906 \cdot 10^{-3}$	$9.636 \cdot 10^{-5}$	1.247	$8.389 \cdot 10^{-3}$	0.625
$1.953 \cdot 10^{-3}$	$4.056 \cdot 10^{-5}$	1.249	$5.440 \cdot 10^{-3}$	0.625
theoretical		0.833		0.417

TABLE 2.5 – For a Sobolev regularity $H^{\frac{5}{2}-}$. Case $\mathcal{P}$ even

The second kind of comparison consists of fixing the equation (*i.e* choosing whether $p = 1$ or $p = 2$) and initializing the scheme with various initial data. This comparison is summarized in Figure 2.2 for $p = 1$ (the heat equation) and Figure 2.3 for $p = 2$ (the fourth-order equation). In both figures, the Crank-Nicolson scheme is illustrated to the left and the implicit scheme to the right.

We recall that we have chosen $\Delta t = \Delta x$. We anticipate a first order for implicit scheme and second order for Crank-Nicolson scheme, for smooth initial datum. Experimental results fit this

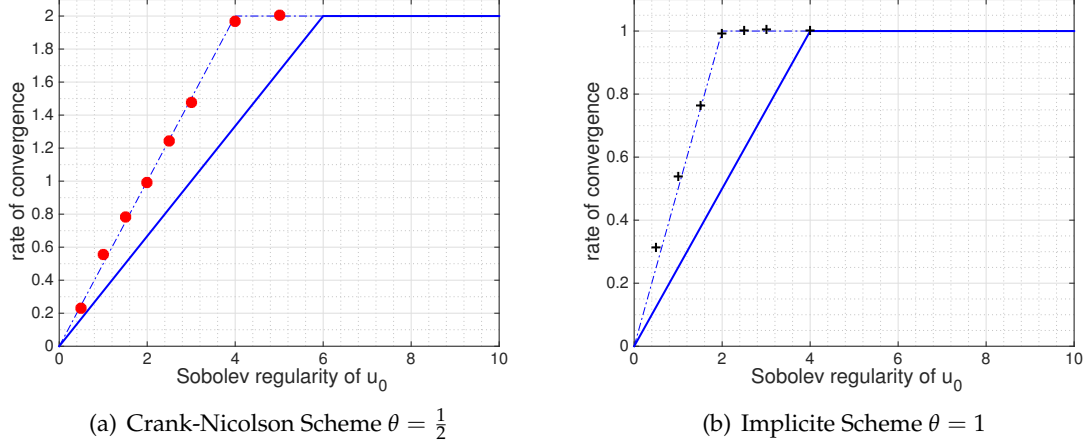


FIGURE 2.2 – Experimental versus theoretical orders : case $p = 1$ - Continuous line : theoretical orders, Dots : experimental orders, Dashed line : conjectured orders

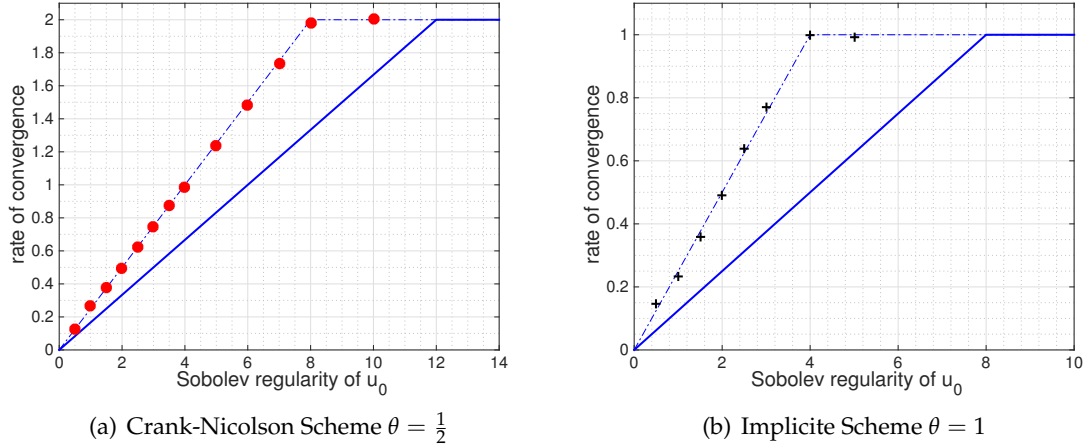


FIGURE 2.3 – Experimental versus theoretical orders : case $p = 2$ - Continuous line : theoretical orders, Dots : experimental orders, Dashed line : conjectured orders

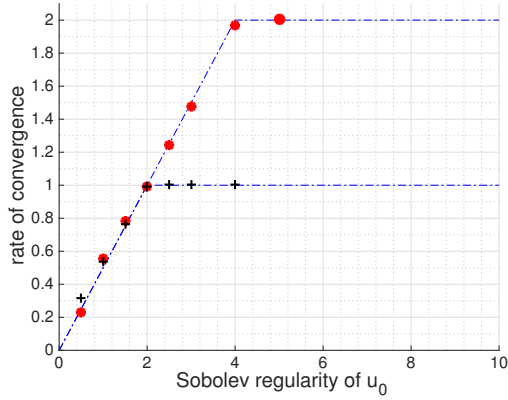
assumption. However, for non smooth initial datum, experimental orders seem better than theoretical ones.

An explanation of the differences between experimental and theoretical rates (in Tables 2.4-2.5 and Figures 2.2-2.3) could be as follows. Contrary to the high-order dispersive equations, the high-order polyharmonic heat equations have a regularizing effect, *cf* Proposition 2.1. Namely, the exact solution at time t would be significantly smoother than the initial datum u_0 . However, we do not have taken into account this natural regularization in our theoretical results : we keep thinking about only u_0 . This is why we obtain theoretically poorer results than the experimental simulations. From the experimental results, it appears that the optimal (conjectured) convergence order would be

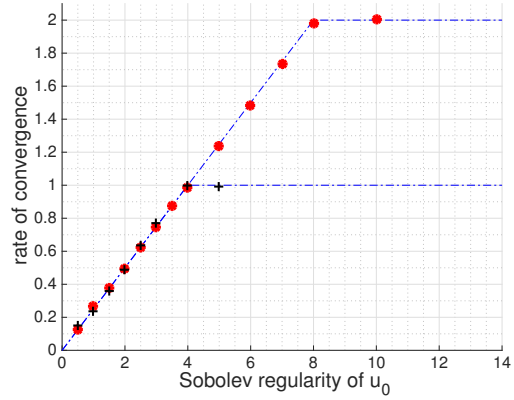
$$\left. \begin{array}{ll} \Delta x^{\frac{\min(s, 2p)}{2p}}, & \text{if } \theta = 1, \\ \Delta x^{\frac{\min(s, 4p)}{2p}}, & \text{if } \theta = \frac{1}{2}, \end{array} \right\} \text{ if } \Delta x = \Delta t.$$

We find interesting to notice that the experimental slope does not appear to depend on the

value of θ : the experimental slope is always $\frac{s}{2p}$ whether $\theta = \frac{1}{2}$ or $\theta = 1$, as illustrated in Figure 2.4.



(a) Case $p = 1$ (Heat equation)



(b) Case $p = 2$

FIGURE 2.4 – Experimental slopes - Red dots : Crank-Nicolson scheme ($\theta = \frac{1}{2}$) and Black crosses : Implicit scheme ($\theta = 1$)

For $\theta \neq \frac{1}{2}$	p even		p odd
	$p = 0$ (Advection)	$p \neq 0$	
Right-winded scheme	$\Delta t^{\frac{\min(s,2)}{2}} \left\ \partial_x^{\min(s,2)} u_0 \right\ _{L^2} + \Delta x^{\frac{\min(s,2)}{2}} \left\ \partial_x^{\min(s,2)} u_0 \right\ _{L^2}$		$\Delta t^{\frac{\min(s,4p+2)}{4p+2}} \left\ \partial_x^{\min(s,4p+2)} u_0 \right\ _{L^2} + \Delta x^{\frac{\min(s,2p+2)}{2p+2}} \left\ \partial_x^{\min(s,2p+2)} u_0 \right\ _{L^2}$
Left-winded scheme	$\Delta t^{\frac{\min(s,2)}{2}} \left\ \partial_x^{\min(s,2)} u_0 \right\ _{L^2} + \Delta x^{\frac{\min(s,2)}{2}} \left\ \partial_x^{\min(s,2)} u_0 \right\ _{L^2}$	$\Delta t^{\frac{\min(s,4p+2)}{4p+2}} \left\ \partial_x^{\min(s,4p+2)} u_0 \right\ _{L^2} + \Delta x^{\frac{\min(s,2p+2)}{2p+2}} \left\ \partial_x^{\min(s,2p+2)} u_0 \right\ _{L^2}$	
Central scheme	$\Delta t^{\frac{\min(s,2)}{2}} \left\ \partial_x^{\min(s,2)} u_0 \right\ _{L^2} + \Delta x^2 \frac{\min(s,3)}{3} \left\ \partial_x^{\min(s,3)} u_0 \right\ _{L^2}$	$\Delta t^{\frac{\min(s,4p+2)}{4p+2}} \left\ \partial_x^{\min(s,4p+2)} u_0 \right\ _{L^2} + \Delta x^2 \frac{\min(s,2p+3)}{2p+3} \left\ \partial_x^{\min(s,2p+3)} u_0 \right\ _{L^2}$	$\Delta t^{\frac{\min(s,4p+2)}{4p+2}} \left\ \partial_x^{\min(s,4p+2)} u_0 \right\ _{L^2} + \Delta x^2 \frac{\min(s,2p+3)}{2p+3} \left\ \partial_x^{\min(s,2p+3)} u_0 \right\ _{L^2}$
For $\theta = \frac{1}{2}$	(Crank-Nicolson case)		
Right-winded scheme	$\Delta t^2 \frac{\min(s,3)}{3} \left\ \partial_x^{\min(s,3)} u_0 \right\ _{L^2(\mathbb{R})} + \Delta x^{\frac{\min(s,2)}{2}} \left\ \partial_x^{\min(s,2)} u_0 \right\ _{L^2(\mathbb{R})}$		$\Delta t^2 \frac{\min(s,6p+3)}{6p+3} \left\ \partial_x^{\min(s,6p+3)} u_0 \right\ _{L^2(\mathbb{R})} + \Delta x^{\frac{\min(s,2p+2)}{2p+2}} \left\ \partial_x^{\min(s,2p+2)} u_0 \right\ _{L^2(\mathbb{R})}$
Left-winded scheme	$\Delta t^2 \frac{\min(s,3)}{3} \left\ \partial_x^{\min(s,3)} u_0 \right\ _{L^2(\mathbb{R})} + \Delta x^{\frac{\min(s,2)}{2}} \left\ \partial_x^{\min(s,2)} u_0 \right\ _{L^2(\mathbb{R})}$	$\Delta t^2 \frac{\min(s,6p+3)}{6p+3} \left\ \partial_x^{\min(s,6p+3)} u_0 \right\ _{L^2(\mathbb{R})} + \Delta x^{\frac{\min(s,2p+2)}{2p+2}} \left\ \partial_x^{\min(s,2p+2)} u_0 \right\ _{L^2(\mathbb{R})}$	
Central scheme	$\Delta t^2 \frac{\min(s,3)}{3} \left\ \partial_x^{\min(s,3)} u_0 \right\ _{L^2(\mathbb{R})} + \Delta x^2 \frac{\min(s,3)}{3} \left\ \partial_x^{\min(s,3)} u_0 \right\ _{L^2(\mathbb{R})}$	$\Delta t^2 \frac{\min(s,6p+3)}{6p+3} \left\ \partial_x^{\min(s,6p+3)} u_0 \right\ _{L^2(\mathbb{R})} + \Delta x^2 \frac{\min(s,2p+3)}{2p+3} \left\ \partial_x^{\min(s,2p+3)} u_0 \right\ _{L^2(\mathbb{R})}$	$\Delta t^2 \frac{\min(s,6p+3)}{6p+3} \left\ \partial_x^{\min(s,6p+3)} u_0 \right\ _{L^2(\mathbb{R})} + \Delta x^2 \frac{\min(s,2p+3)}{2p+3} \left\ \partial_x^{\min(s,2p+3)} u_0 \right\ _{L^2(\mathbb{R})}$

 TABLE 2.6 – Error estimates for $u_0 \in H^s(\mathbb{R})$, with $s \geq 0$

TROISIÈME PARTIE

Analyse numérique de l'équation KdV et du système *abcd*

CHAPITRE 3

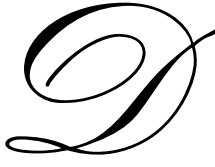
Error estimates for a finite difference scheme for the Korteweg-de Vries equation

This chapter deals with the numerical analysis of the Cauchy problem for the Korteweg-de Vries equation with a finite difference scheme. We consider the Rusanov scheme for the hyperbolic flux term and a 4-points θ -scheme for the dispersive term. We prove the convergence under a hyperbolic Courant-Friedrichs-Lewy condition when $\theta \geq \frac{1}{2}$ and under an "Airy" Courant-Friedrichs-Lewy condition when $\theta < \frac{1}{2}$. More precisely, we get the first order convergence rate for strong solutions in the Sobolev space $H^s(\mathbb{R})$, $s \geq 6$ and extend this result to the non-smooth case for initial data in $H^s(\mathbb{R})$, with $s \geq \frac{3}{4}$, to the price of a loss in the convergence order. The orders of convergence seem optimal with experimental simulations in some cases, at least when $s \geq 3$.

Summary

3.1	Introduction (en français)	67
3.2	Known results on the Cauchy problem for the KdV equation	72
3.3	Consistency error estimate	73
3.4	Stability estimate	75
3.4.1	Preliminary results	75
3.4.2	The discrete relative entropy inequality	77
3.5	Rate of convergence	84
3.5.1	Properties of dispersive operators	84
3.5.2	Simplification of Inequality (3.37)	86
3.5.3	Induction method	90
3.6	Convergence for less smooth initial data	95
3.6.1	Proof of convergence Theorem 3.10	96
3.7	Numerical results	99
3.7.1	Convergence rates	99
3.7.2	Smooth initial data	100
3.7.3	For a regularity between $H^3(\mathbb{R})$ and $H^{10}(\mathbb{R})$	101
3.7.4	For a regularity between $L^2(\mathbb{R})$ and $H^3(\mathbb{R})$	103
3.8	Afterword : go one step further in the Burgers and Airy structure of (KdV) . .	107
3.8.1	The Airy equation	107
3.8.2	The Burgers equation	107
3.8.2.1	The well-known hyperbolic CFL	108
3.8.2.2	Numerical entropy flux for Burgers equation	109

3.1 Introduction (en français)



ans ce chapitre, nous nous intéressons à l'analyse numérique pour le problème de Cauchy de l'équation de Korteweg-de Vries suivant

$$\begin{cases} \partial_t u(t, x) + \partial_x \left(\frac{u^2}{2} \right) (t, x) + \partial_x^3 u(t, x) = 0, & (t, x) \in [0, T] \times \mathbb{R} \quad (3.1a) \\ u|_{t=0}(x) = u_0(x), & x \in \mathbb{R}. \quad (3.1b) \end{cases}$$

Nous précisons dans les paragraphes suivants le schéma numérique que nous avons choisi d'utiliser.



Schéma numérique pour KdV. L'équation de Korteweg-de Vries possède un terme non linéaire $\partial_x \left(\frac{u^2}{2} \right)$ qui est le même que celui qui apparaît dans l'équation de Burgers non visqueuse, ainsi qu'un terme dispersif $\partial_x^3 u$ qui est celui de l'équation d'Airy, étudiée en partie II. Nous gardons les mêmes notations que pour les schémas numériques précédents : Δt sera le pas de temps et Δx le pas d'espace. Les points du maillage seront comme précédemment notés $t^n := n\Delta t$, pour $n \in \llbracket 0, N \rrbracket$ et $x_j := j\Delta x$, pour $j \in \mathbb{Z}$.

- Concernant la discrétisation du terme non linéaire, nous choisissons le schéma de Rusanov, initialement introduit dans [90], qui s'écrit pour l'équation de Burgers,

$$\begin{cases} \frac{v_j^{n+1} - v_j^n}{\Delta t} + D_c \left(\frac{v^2}{2} \right)_j^n = \frac{\tau \Delta x}{2} D_+ D_- (v)_j^n, & (n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}, \\ v_j^0 = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u_0(x) dx, & x \in \mathbb{Z}. \end{cases}$$

Remarque 3.1. Les opérateurs D_c , D_+ et D_- ont déjà été introduits au chapitre 1. Nous rappelons néanmoins que

$$D_c \left(\frac{v^2}{2} \right)_j^n = \frac{(v^2)_{j+1}^n - (v^2)_{j-1}^n}{4\Delta x} \quad \text{et} \quad D_+ D_- (v)_j^n = \frac{v_{j+1}^n - 2v_j^n + v_{j-1}^n}{\Delta x^2}.$$

La constante de Rusanov τ est choisie telle que $\|v\|_{\ell_n^\infty \ell_j^\infty} \leq \tau$, pour assurer la stabilité de ce schéma.

Remarque 3.2. Un des avantages du schéma de Rusanov est d'être moins diffusif que d'autres schémas (comme par exemple le schéma de Lax-Friedrichs) et d'avoir une diffusion linéaire (et explicite : $D_+ D_- (v)_j^n$ par rapport au schéma de Godunov par exemple), il est aussi facilement implémentable et converge (au plus à l'ordre un) vers la solution faible entropique de l'équation de Burgers si $\tau \Delta t \leq \Delta x$ et $\|v\|_{\ell_n^\infty \ell_j^\infty} \leq \tau$ [49].

- Concernant le terme dispersif d'Airy, nous utilisons les résultats de la partie II et choisissons un θ -schéma décentré à droite.

Cela aboutit au schéma suivant, pour l'équation de KdV, avec $\theta \in [0, 1]$,

$$\begin{cases} \frac{v_j^{n+1} - v_j^n}{\Delta t} + D_c \left(\frac{v^2}{2} \right)_j^n + \theta D_+ D_+ D_- (v)_j^{n+1} + (1 - \theta) D_+ D_+ D_- (v)_j^n = \frac{\tau \Delta x}{2} D_+ D_- (v)_j^n, & (n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}, \quad (3.2a) \\ v_j^0 = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u_0(y) dy, & j \in \mathbb{Z}. \quad (3.2b) \end{cases}$$

Section 3.1. Introduction (en français)

Remarque 3.3. En développant tous les opérateurs $D_\bullet$ selon la remarque 1.22 et les relations (1.12), nous obtenons le schéma

$$\begin{aligned} \frac{v_j^{n+1} - v_j^n}{\Delta t} + \frac{(v_{j+1}^n)^2 - (v_{j-1}^n)^2}{4\Delta x} + \theta \left(\frac{v_{j+2}^{n+1} - 3v_{j+1}^{n+1} + 3v_j^{n+1} - v_{j-1}^{n+1}}{\Delta x^3} \right) \\ + (1 - \theta) \left(\frac{v_{j+2}^n - 3v_{j+1}^n + 3v_j^n - v_{j-1}^n}{\Delta x^3} \right) = \frac{\tau \Delta x}{2} \left(\frac{v_{j+1}^n - 2v_j^n + v_{j-1}^n}{\Delta x^2} \right). \end{aligned}$$

Remarque 3.4. • Tous les résultats de cette partie sont valables avec un pas de temps variable Δt^n et un coefficient de Rusanov variable τ^n . Pour plus de simplicité, nous omettons la dépendance en n .

• Le schéma (3.2a)-(3.2b) est une généralisation de celui étudié par Holden, Koley et Risebro [55]. En effet, ils considèrent le schéma de Lax-Friedrichs pour le flux hyperbolique combiné avec un schéma implicite décentré à droite, pour la partie dispersive. Cela correspond à choisir $\tau \Delta t = \Delta x$ et $\theta = 1$ dans le schéma (3.2a)-(3.2b).

Remarque 3.5. En suivant la remarque 2.1 de [67], utiliser le schéma numérique (3.2a)-(3.2b) avec $\theta = 1$ revient d'une certaine manière à utiliser un opérateur de splitting mais la très forte condition de régularité de u_0 nécessaire pour les opérateurs de splitting (par exemple, u_0 dans $H^{r+5}(\mathbb{R})$, avec $r \geq 1$ quelconque, dans [56], pour assurer la convergence d'ordre deux dans $H^r(\mathbb{R})$) est ici remplacée par la condition CFL entre Δt et Δx . En effet, le schéma (3.2a)-(3.2b) peut être vu comme la résolution successive des deux schémas suivants

• résolution de l'équation de Burgers par le schéma de Rusanov

$$v_j^{n+\frac{1}{2}} = v_j^n - \Delta t D_c \left(\frac{v^2}{2} \right)_j + \frac{\tau \Delta x \Delta t}{2} D_+ D_- v_j^n,$$

• puis résolution de l'équation d'Airy par un schéma implicite du premier ordre

$$\frac{v_j^{n+1} - v_j^{n+\frac{1}{2}}}{\Delta t} + D_+ D_+ D_- v_j^{n+1} = 0.$$

Si on note $\mathcal{S}_{\Delta t}^{\text{Burgers}}$ et $\mathcal{S}_{\Delta t}^{\text{Airy}}$ ces deux opérateurs précédents, on trouve que le schéma (3.2a)-(3.2b) se réécrit

$$v_j^{n+1} = \left(\mathcal{S}_{\Delta t}^{\text{Airy}} \circ \mathcal{S}_{\Delta t}^{\text{Burgers}} \right) v_j^n, \quad \forall (n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}.$$

Nous rappelons la définition de l'erreur de convergence

$$e_j^n = v_\Delta(t^n, x_j) - u_\Delta(t^n, x_j), \quad (n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}, \quad (3.3)$$

où v_Δ et u_Δ sont les deux fonctions suivantes définies par morceaux :

pour $n \in \llbracket 0, N \rrbracket$ et $j \in \mathbb{Z}$

$$\begin{cases} u_\Delta(t, x) = (u_\Delta)_j^n, \\ v_\Delta(t, x) = v_j^n, \end{cases} \text{ si } (t, x) \in [t^n, \inf(t^{n+1}, T)) \times [x_j, x_{j+1}[. \quad (3.4)$$

Dans ce qui précède, $(v_j^n)_{(n,j)}$ est le vecteur de la solution numérique et $((u_\Delta)_j^n)_{(n,j)}$ le vecteur construit à partir de la solution exacte tel que

$$\begin{cases} (u_\Delta)_j^n = \frac{1}{\Delta x [\inf(t^{n+1}, T) - t^n]} \int_{t^n}^{\inf(t^{n+1}, T)} \int_{x_j}^{x_{j+1}} u(s, y) dy ds, & \text{si } (n, j) \in \llbracket 1, N \rrbracket \times \mathbb{Z}, \\ (u_\Delta)_j^0 = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u_0(y) dy, & \text{si } j \in \mathbb{Z}. \end{cases} \quad (3.5)$$

Remarque 3.6. Pour plus de simplicité, nous poserons $\Delta t = \inf(T, t^{n+1}) - t^n$ par la suite.

L'erreur de consistance est, quant à elle, définie par

$$\begin{aligned} \epsilon_j^n = & \frac{(u_\Delta)_j^{n+1} - (u_\Delta)_j^n}{\Delta t} + D_c \left(\frac{u_\Delta^2}{2} \right)_j^n + \theta D_+ D_+ D_- (u_\Delta)_j^{n+1} \\ & + (1 - \theta) D_+ D_+ D_- (u_\Delta)_j^n - \frac{\tau \Delta x}{2} D_+ D_- (u_\Delta)_j^n, \quad (n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}. \end{aligned} \quad (3.6)$$

⋈

Résultats principaux de ce chapitre. Le premier résultat concerne les données initiales très régulières.

Théorème 3.7. *Convergence pour une donnée initiale régulière*

Supposons $u_0 \in H^s(\mathbb{R})$, avec $s \geq 6$. Choisissons $\alpha_0 > 0$ et $\tau > 0$ tels que

$$\| (u_\Delta)^n \|_{\ell^\infty} + \alpha_0 = \sup_{j \in \mathbb{Z}} | (u_\Delta)_j^n | + \alpha_0 \leq \tau, \quad n \in \mathbb{N}. \quad (3.7)$$

Soit $\beta_0 \in]0, 1[$, il existe $\omega_0 > 0$ tel que

$$\Delta t \leq \omega_0, \quad \Delta x \leq \omega_0,$$

et

$$\begin{cases} 4(1 - 2\theta) \frac{\Delta t}{\Delta x^3} \leq 1 - \beta_0, \\ \left[\tau + \frac{1}{2} \right] \frac{\Delta t}{\Delta x} \leq 1 - \beta_0. \end{cases} \quad (3.8)$$

Alors, le schéma aux différences finies (3.2a)-(3.2b) avec la constante de Rusanov τ choisie précédemment, les pas de temps et d'espace Δt et Δx vérifiant la condition de Courant-Friedrichs-Lewy (condition CFL par la suite) (3.8), vérifie, pour tout $\eta \in]0, s - \frac{3}{2}]$,

$$\| e^n \|_{\ell^\infty(\llbracket 0, N \rrbracket, \ell_\Delta^2(\mathbb{Z}))} \leq \Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}} \left(1 + \|u_0\|_{H^{\frac{1}{2}+\eta}}^2 \right) \left(\frac{\|u_0\|_{H^6}}{\tau + \frac{1}{2}} + \|u_0\|_{H^4} + \|u_0\|_{H^{\frac{3}{2}+\eta}} \|u_0\|_{H^1} \right) \Delta x, \quad (3.9)$$

où e^n est défini par (3.3)-(3.4)-(3.5) et $\Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}}$ par

$$\Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}} = \exp \left(\frac{C}{2} (1 + \tau^2) \left(1 + \frac{\Delta t^2}{\Delta x^2} \right) \left(T + (T^{\frac{3}{4}} + T^{\frac{1}{2}}) \|u_0\|_{H^{\frac{3}{4}}} e^{\kappa_{\frac{3}{4}} T} \right) \right) C e^{\kappa T} \sqrt{T \left\{ 1 + \frac{\Delta t}{\Delta x} \right\}}, \quad (3.10)$$

la variable C est une constante, $\kappa_{\frac{3}{4}}$ et κ dépendent seulement de $\|u_0\|_{L^2(\mathbb{R})}$.

Remarque 3.8. L'hypothèse $\left[\tau + \frac{1}{2} \right] \frac{\Delta t}{\Delta x} \leq 1 - \beta_0$ semble être une hypothèse technique, qui peut sûrement être remplacée par la condition CFL hyperbolique classique $\tau \Delta t \leq \Delta x$. En effet, les résultats expérimentaux semblent en accord avec le théorème 3.7 tout en choisissant la CFL classique, voir la section 3.7.

Remarque 3.9. Par la suite, η sera choisi positif, aussi petit que possible. Ainsi, les normes $\|u_0\|_{H^{s+\eta}}(\mathbb{R})$ seront utilisées dans le sens de la norme $\|u_0\|_{H^{s+}}(\mathbb{R})$.

Le deuxième résultat traite des données initiales moins régulières $u_0 \in H^s(\mathbb{R})$, $s \geq 3/4$. Pour cela, il est nécessaire d'introduire une régularisation comme en partie II (et introduite dans l'introduction générale à la page 14) et de changer l'initialisation du schéma. Nous reprenons φ^δ la suite régularisante, comme précédemment

- nous appelons u la solution exacte de (KdV) partant de la donnée initiale u_0 .
- Soit u^δ la solution de (KdV) avec $u_0^\delta = u_0 \star \varphi^\delta$ comme donnée initiale, où $\star$ correspond au

Section 3.1. Introduction (en français)

produit de convolution.

• Nous notons ensuite $((v^\delta)_j^n)_{(n,j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}}$ la solution numérique obtenue en appliquant le schéma (3.2a) à la donnée initiale $(u_0^\delta)_\Delta$:

$$(v^\delta)_j^0 = (u_0^\delta)_\Delta = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u_0 \star \varphi^\delta(y) dy. \quad (3.11)$$

Théorème 3.10. Convergence pour une donnée initiale peu régulière

Supposons $u_0 \in H^s(\mathbb{R})$ avec $s \geq \frac{3}{4}$. Choisissons $\alpha_0 > 0$ et $\tau > 0$ tels que

$$\|(u_\Delta)^n\|_{\ell^\infty} + \alpha_0 \leq \tau, \quad n \in \mathbb{N}.$$

Soit $\beta_0 \in]0, 1[$, il existe $\omega_0 > 0$ tel que

$$\Delta t \leq \omega_0, \quad \Delta x \leq \omega_0,$$

et

$$\begin{cases} 4(1-2\theta) \frac{\Delta t}{\Delta x^3} \leq 1 - \beta_0, \\ \left[\tau + \frac{1}{2} \right] \frac{\Delta t}{\Delta x} \leq 1 - \beta_0. \end{cases} \quad (3.12)$$

Alors, le schéma aux différences finies (3.2a)-(3.11) avec la constante de Rusanov τ choisie précédemment et les pas de temps et d'espace Δt et Δx vérifiant la condition de type CFL (3.12), vérifie, pour tout η tel que $\eta \in]0, s - \frac{1}{2}]$,

$$\|e^n\|_{\ell^\infty(\llbracket 0, N \rrbracket, \ell_\Delta^2(\mathbb{Z}))} \leq \Gamma_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}} \left(1 + \|u_0\|_{H^{\frac{1}{2}+\eta}}^2 \right) \left(\frac{1}{\tau + \frac{1}{2}} + 1 + \|u_0\|_{H^{\min(1,s)}} \right) \|u_0\|_{H^s} \Delta x^\eta,$$

avec e^n définie par (3.3)-(3.4)-(3.5) avec $(v^\delta)_j^n$ à la place de v_j^n et avec

- $q = \frac{s}{12-2s}$, si $\frac{3}{4} \leq s \leq 3$,

- $q = \frac{\min(s,6)}{6}$, si $3 < s$.

La quantité $\Gamma_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}}$ est définie par $\Gamma_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}} = C \left[\Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}} + \exp \left(\frac{T^{\frac{3}{4}} C_{\frac{3}{4}} e^{\frac{\kappa_{\frac{3}{4}} T}{2}} \|u_0\|_{H^{\frac{3}{4}}}} \right) \right]$, où

$\Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}}$ est défini par (3.10), C et $C_{\frac{3}{4}}$ sont deux constantes et $\kappa_{\frac{3}{4}}$ ne dépend que de $\|u_0\|_{L^2(\mathbb{R})}$.

Remarque 3.11. • Si $u_0 \in H^s(\mathbb{R})$ avec $s \geq 6$, alors le théorème 3.10 implique un ordre de convergence égal à un et nous retrouvons le résultat du théorème 3.7 dans ce cas.

• Notez que le résultat est valide pour tout temps $T > 0$ puisque à ce niveau de régularité, nous avons des solutions globales de l'équation de KdV qui conservent leur régularité de Sobolev au cours du temps.



Méthode suivie. Pour démontrer le théorème 3.7, nous démontrons la consistance et la stabilité du schéma. La principale difficulté dans la preuve est de traiter en même temps le terme hyperbolique et le terme dispersif et la partie la plus délicate de la preuve consiste en l'étude de la stabilité du schéma. Une idée est alors de tirer profit du fait que les solutions exactes restent régulières sur tout $[0, T]$. L'idée principale est de transposer au niveau discret la propriété de stabilité fort-faible classique pour les lois de conservation hyperboliques. Cette méthode se base sur les estimations d'entropies relatives développées par Dafermos [39] ou DiPerna [44].

Définition 3.12. Considérons une équation hyperbolique écrite sous la forme conservative

$$\partial_t u + \partial_x(f(u)) = 0. \quad (3.13)$$

• Soit η une fonction convexe et g tel que $g' = \eta' f'$, alors (η, g) est appelé un couple entropie/flux d'entropie associée à l'EDP (3.13).

• Soit (η, g) un couple entropie-flux d'entropie avec en plus $\eta \in C^2(\mathbb{R})$ et $\eta''(u) \geq \alpha > 0$, pour tout $u \in \mathbb{R}$. Le couple d'entropie relative-flux d'entropie relative $(\eta(\cdot), g(\cdot))$ est défini par

$$\eta(u|v) := \eta(u) - \eta(v) - \eta'(v) \cdot (u - v) \quad (3.14a)$$

$$g(u;v) := g(u) - g(v) - \eta'(v) \cdot (f(u) - f(v)), \quad \forall (u, v) \in \mathbb{R}. \quad (3.14b)$$

La stricte convexité de η permet d'obtenir une estimation *a priori* de la solution dans $L^2(\mathbb{R})$, puisque pour tous u, v réels, on a $\eta(u|v) \geq \frac{\alpha}{2}|u - v|^2$. Estimer l'intégrale $\int_{\mathbb{R}} \eta(u|v) dx$ est donc un moyen d'estimer $u - v$ dans L^2 , et cela conduit à une inégalité de stabilité L^2_{loc} , à la base de l'unicité fort-faible de Dafermos [39] : considérons u la solution forte et w une solution faible d'un même problème de loi de conservation, on a pour tout $r > 0$:

$$\int_{|x| < r} |w(x, T) - u(x, T)|^2 dx \leq C(T, u) \int_{|x| < r+sT} |w_0(x) - u_0(x)|^2 dx, \quad (3.15)$$

avec $C(T, u) < +\infty$ si u est lipschitzienne. La constante $s > 0$ est choisie telle que $|g(w; u)| \leq s\eta(w|u)$. Cette inégalité (3.15) s'appelle l'unicité fort-faible : pour une même donnée initiale, toute solution faible coïncide avec la solution forte sur son domaine de définition (cf [39]).

Remarque 3.13. Notons que le facteur d'amplification $C(T, u)$, qui apparaît à l'équation (3.15), est indépendant de la solution faible w : il ne dépend que de la solution « la plus régulière » des deux et fait plus précisément intervenir la norme $\int_0^T \|\partial_x u(t, \cdot)\|_{L^\infty} dt$. Notons également que dans le cas particulier d'une entropie quadratique $\eta_0(u) = \frac{u^2}{2}$, l'entropie relative se simplifie considérablement $\eta_0(u|v) = \frac{(u-v)^2}{2}$, pour $\eta_0(u) := \frac{u^2}{2}$.

Pour de faibles régularités, le facteur d'amplification $\int_0^T \|\partial_x u^\delta(t, \cdot)\|_{L^\infty} dt$ est fini et peut être borné indépendamment du paramètre de régularisation δ du moment que la donnée initiale est dans H^s , $s \geq 3/4$ (grâce aux estimations de Strichartz qui assurent que $\partial_x u \in L^4([0, T], L^\infty(\mathbb{R}))$). Nous pouvons donc utiliser cette méthode, même pour de faibles régularités.

Nous suivrons cette heuristique mais non pas en comparant une solution forte u avec une solution faible w . Nous allons remplacer w par la solution numérique $(v_j^n)_{(n,j)}$ et essayer de trouver le pendant discret d'une inégalité d'entropie relative. Les étapes des preuves de convergence seront donc les suivantes :

Étapes clés : preuve de convergence pour (KdV)

- Étude de la consistance
 - Dérivation d'une inégalité d'entropie discrète entre la solution numérique $(v_j^n)_{(n,j)}$ et la solution exacte moyennée par morceaux u_Δ .
 - Obtention des conditions CFL pour assurer la stabilité ℓ_Δ^2
 - Inégalité finale par le lemme de Grönwall
-



Organisation du chapitre. Nous détaillons la preuve des deux résultats principaux énoncés ci-dessus. En section 3.2, nous énonçons les résultats de la théorie de Cauchy pour (KdV) que nous utiliserons par la suite. La section 3.3 est consacrée à l'analyse de l'erreur de consistance ϵ_j^n (les preuves des lemmes techniques sont répertoriées au niveau de l'annexe B). Le but de la section 3.4 est d'obtenir l'inégalité discrète d'entropie relative, équation cruciale par la suite pour établir la stabilité. Le taux de convergence pour une donnée initiale régulière est finalement déterminé en section 3.5 et pour une donnée initiale peu régulière en section 3.6. Finalement, certains résultats numériques illustrent le taux de convergence théorique et semblent montrer qu'il est optimal pour $s \geq 3$, en section 3.7. Une digression sur le flux numérique de l'équation de Burgers pour l'entropie quadratique conclut cette partie, en section 3.8.

Remarque 3.14. Par la suite, la lettre C représente une constante strictement positive qui peut être différente d'une ligne à l'autre et qui peut être choisie indépendamment de Δt , Δx , u , u_0 , T et δ . Nous noterons κ toute constante ne dépendant que de $\|u_0\|_{L^2(\mathbb{R})}$.



3.2 Known results on the Cauchy problem for the KdV equation

Let us recall the definition of Bourgain spaces. A tempered distribution $u(t, x)$ on $\mathbb{R}^2$ belongs to $X^{s,b}$ if its following norm is finite

$$\|u\|_{X^{s,b}} = \left(\int_{\mathbb{R}} \int_{\mathbb{R}} (1 + |\xi|)^{2s} (1 + |\tau - \xi^3|)^{2b} |\tilde{u}(\tau, \xi)|^2 d\xi d\tau \right)^{\frac{1}{2}},$$

where $\tilde{u}$ is the space and time Fourier transform of u . We shall also use a localized version of this space, $u \in X^{s,b}(I)$ where $I \subset \mathbb{R}$ is an interval if $\|u\|_{X^{s,b}(I)} < +\infty$ where

$$\|u\|_{X^{s,b}(I)} = \inf\{\|\bar{u}\|_{X^{s,b}}, \bar{u}|_I = u\}.$$

By using results from [64], [24], [65], see for example the book [80], we get the following theorem.

Theorem 3.15. *Consider $s \geq 0$, $1 > b > 1/2$. There exists a unique global solution u of (3.1a)-(3.1b), with $u_0 \in H^s(\mathbb{R})$, such that for every $T \geq 0$, $u \in \mathcal{C}([0, T], H^s(\mathbb{R})) \cap X^{s,b}([0, T])$. Moreover, there exists $\kappa_s > 0$, depending only on s and on the norm $\|u_0\|_{L^2}$ and $C_s > 0$, depending only on s , such that, for any $T \geq 0$,*

- $\sup_{t \in [0, T]} \|u(t)\|_{H^s(\mathbb{R})} \leq C_s \|u_0\|_{H^s(\mathbb{R})} e^{\kappa_s T},$
- if $s \geq \frac{3}{4}$, $\|\partial_x u\|_{L^i([0, T], L^\infty(\mathbb{R}))} \leq T^{\frac{4-i}{4i}} \|u_0\|_{H^{\frac{3}{4}}(\mathbb{R})} C_{\frac{3}{4}} e^{\kappa_{\frac{3}{4}} T}, \text{ for } i \in \{1, 2\}.$

The growth rate in the above estimates is not optimal.

Note that a local well-posedness result for $s > 3/4$ follows directly from [64]. In the present paper, we will be only interested in $s \geq 3/4$, nevertheless, to get the global well-posedness for $s \in [3/4, 1)$, we need to go through the L^2 local well-posedness result.

Proof. Let us just briefly explain how we can organize now classical arguments to get the result. We refer for example to [65], [80] for the details. The existence is proved by a fixed point argument on the following truncated problem :

$$v \mapsto F(v)$$

such that

$$F(v)(t) = \chi(t) e^{-t\partial_x^3} u_0 - \chi(t) \int_0^t e^{-(t-\tau)\partial_x^3} \partial_x \left(\chi\left(\frac{\tau}{\delta}\right) \frac{v^2}{2}(\tau) \right) d\tau,$$

where $\chi \in [0, 1]$ is a smooth compactly supported function which is equal to 1 on $[-1, 1]$ and supported in $[-2, 2]$. For $|t| \leq \delta \leq 1/2$, a fixed point of the above equation gives a solution of the original Cauchy problem, denoted by u .

Fix $C > 0$, that does not depend on u , such that

$$\|\chi(t) e^{-t\partial_x^3} u_0\|_{X^{0,b}} \leq C \|u_0\|_{L^2}.$$

We can first prove that F is a contraction on a suitable ball of $X^{0,b}$, provided $8C^2 \|u_0\|_{L^2} \delta^\beta \leq 1$ for some $\beta > 0$ (that is related to $1 > b > 1/2$) that does not depend on δ nor u . In particular, for the fixed point, we can ensure that

$$\|v\|_{X^{0,b}} \leq 2C \|u_0\|_{L^2}.$$

Next, by using again the Duhamel formula, we can obtain, for $s \geq 0$,

$$\|v\|_{X^{s,b}} \leq c_s \|u_0\|_{H^s} + c_s \delta^\beta \|v\|_{X^{0,b}} \|v\|_{X^{s,b}} \leq c_s \|u_0\|_{H^s} + 2c_s C \|u_0\|_{L^2} \delta^\beta \|v\|_{X^{s,b}},$$

where c_s depends only on s . In particular, by choosing δ , possibly smaller than previously, such that $2c_s C \|u_0\|_{L^2} \delta^\beta \leq 1/2$, we thus obtain that

$$\|v\|_{X^{s,b}} \leq 2c_s \|u_0\|_{H^s}.$$

Next, by using that the $X^{s,b}$ norm for $b > 1/2$ controls the $\mathcal{C}(\mathbb{R}, H^s)$ norm (see for example [99] Lemma 2.9 page 100), we obtain that

$$\|u\|_{\mathcal{C}([0,\delta], H^s(\mathbb{R}))} \leq \|v\|_{\mathcal{C}(\mathbb{R}, H^s(\mathbb{R}))} \leq B_s \|u_0\|_{H^s(\mathbb{R})},$$

where B_s depends only on s . Since the existence time δ depends only on the L^2 -norm of the initial datum and that the L^2 -norm is conserved for the KdV equation, we can iterate the above argument to get a global solution. Moreover, in a quantitative way, by choosing $n = \lfloor T/\delta \rfloor + 1$ and iterating n times, we obtain that

$$\|u\|_{\mathcal{C}([0,T], H^s)} + \|u\|_{X^{s,b}([0,T])} \leq B_s^n \|u_0\|_{H^s} \leq C_s \|u_0\|_{H^s} e^{\kappa_s T},$$

where κ_s depends only on s and $\|u_0\|_{L^2}$ while C_s depends only on s .

Finally, since the Strichartz estimate in the KdV context (see [64]) reads

$$\| |\partial_x|^{\frac{1}{4}} e^{-t\partial_x^3} u_0 \|_{L_t^4(\mathbb{R}, L_x^\infty)} \leq C \|u_0\|_{L^2},$$

by using the embedding properties of the Bourgain spaces (see again [99] Lemma 2.9 page 100), we obtain that

$$\|\partial_x u\|_{L_t^4([0,\delta], L_x^\infty)} \leq \|\partial_x v\|_{L_t^4(\mathbb{R}, L_x^\infty)} \leq \|v\|_{X^{\frac{3}{4},b}} \leq C \|u_0\|_{H^{\frac{3}{4}}}.$$

Again by iterating this estimate, we finally obtain that

$$\|\partial_x u\|_{L_t^4([0,T], L_x^\infty)} \leq C_{\frac{3}{4}} \|u_0\|_{H^{\frac{3}{4}}} e^{\kappa_{\frac{3}{4}} T}$$

and the desired estimate follows from the Hölder inequality. □

3.3 Consistency error estimate

This section is devoted to the computation of the consistency error defined by Equation (3.6). As a starting point, by using Theorem 3.15, we obtain the following estimates on the averaged solution u_Δ .

Lemma 3.16. *Let u be the exact solution of (3.1a)-(3.1b) from $u_0 \in H^s(\mathbb{R})$, $s > \frac{1}{2}$ and u_Δ be defined by (3.5). Then there exists $C > 0$, depending only on s , and $\kappa_s > 0$, depending only on s and on $\|u_0\|_{L^2}$, such that, for any $T \geq 0$ and any $n \in \llbracket 0, N \rrbracket$ with $N = \lfloor \frac{T}{\Delta t} \rfloor$,*

$$\begin{aligned} & \bullet \| (u_\Delta)^n \|_{\ell^\infty} \leq C e^{\kappa_s T} \|u_0\|_{H^s}, \\ & \bullet \text{ if } s \geq \frac{3}{4}, \quad \Delta t \|D_+ (u_\Delta)^n\|_{\ell^\infty}^i \leq \int_{t^n}^{t^{n+1}} \|\partial_x u(t, \cdot)\|_{L_x^\infty}^i dt \leq T^{\frac{4-i}{4i}} C e^{\kappa_{\frac{3}{4}} T} \|u_0\|_{H^{\frac{3}{4}}(\mathbb{R})}, \text{ for } i \in \{1, 2\}. \end{aligned} \tag{3.16}$$

Proof. The Sobolev embedding $H^s(\mathbb{R}) \hookrightarrow L^\infty(\mathbb{R})$, for $s > \frac{1}{2}$ yields the inequality

$$\| (u_\Delta)^n \|_{\ell^\infty} \leq \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \|u(t, \cdot)\|_{L^\infty(\mathbb{R})} dt \leq C \sup_{t \in [0, T]} \|u(t, \cdot)\|_{H^s(\mathbb{R})}.$$

Section 3.3. Consistency error estimate

Theorem 3.15 implies

$$\| (u_\Delta)^n \|_{\ell^\infty} \leq CC_s \|u_0\|_{H^s(\mathbb{R})} e^{\kappa_s T},$$

which proves the first estimate of Lemma 3.16.

To prove Equation (3.16) for $i = 1$, we use a Taylor expansion :

$$\begin{aligned} \Delta t \|D_+ (u_\Delta)^n\|_{\ell^\infty} &= \Delta t \left\| \frac{1}{\Delta t \Delta x^2} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} u(s, y + \Delta x) - u(s, y) dy ds \right\|_{\ell^\infty} \\ &= \Delta t \left\| \frac{1}{\Delta x^2 \Delta t} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_y^{y+\Delta x} \partial_x u(s, z) dz dy ds \right\|_{\ell^\infty} \\ &\leq \int_{t^n}^{t^{n+1}} \|\partial_x u(s, \cdot)\|_{L_x^\infty} ds. \end{aligned}$$

For $i = 2$, the same Taylor expansion gives, thanks to the Cauchy-Schwarz inequality,

$$\begin{aligned} \Delta t \|D_+ (u_\Delta)^n\|_{\ell^\infty}^2 &= \Delta t \left\| \frac{1}{\Delta x^2 \Delta t} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_y^{y+\Delta x} \partial_x u(s, z) dz dy ds \right\|_{\ell^\infty}^2 \\ &\leq \frac{\Delta t}{\Delta x^4 \Delta t^2} \Delta t \Delta x^2 \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_y^{y+\Delta x} \|\partial_x u(s, \cdot)\|_{L_x^\infty}^2 dz dy ds \\ &\leq \int_{t^n}^{t^{n+1}} \|\partial_x u(s, \cdot)\|_{L_x^\infty}^2 ds. \end{aligned}$$

Theorem 3.15 concludes the proof. $\square$

Remark 3.17. The Sobolev regularity of the initial datum is at least $H^{\frac{3}{4}}(\mathbb{R})$ in Theorem 3.10 because we need to control $\int_0^T \|\partial_x u(t, \cdot)\|_{L^\infty(\mathbb{R})}^i dt$, for $i \in \{1, 2\}$ in some of the proofs. This is explicitly needed in Lemma 3.16, Theorem 3.15 and in the definition of $\Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}}$ in (3.10).

As a consequence, we control the ℓ_Δ^2 -norm of the consistency error ϵ^n defined in (3.6) in terms of the initial datum thanks to the following proposition.

Proposition 3.18. Consider $u_0 \in H^s(\mathbb{R})$, with $s \geq 6$. There exist $C > 0$ and $\kappa > 0$, depending only on $\|u_0\|_{L^2}$, such that, for any $\eta \in (0, s - \frac{3}{2})$ and for any $T > 0$,

$$\|\epsilon^n\|_{\ell^\infty(\llbracket 0, N \rrbracket; \ell_\Delta^2(\mathbb{Z}))} \leq C e^{\kappa T} \left(1 + \|u_0\|_{H^{\frac{1}{2}+\eta}}^2\right) \left\{ \Delta t \|u_0\|_{H^6} + \Delta x \left[\|u_0\|_{H^4} + \|u_0\|_{H^{\frac{3}{2}+\eta}} \|u_0\|_{H^1} \right] \right\}. \quad (3.17)$$

Proof. The consistency error satisfies the following inequality (cf. Appendix A.1 for the proof) :

$$\begin{aligned} \|\epsilon^n\|_{\ell^\infty(\llbracket 0, N \rrbracket; \ell_\Delta^2)} &\leq B_1 \left\{ \Delta t \sup_{t \in [0, T]} \left[\left(1 + \|u\|_{L_x^\infty}^2\right) \|u\|_{H_x^6} \right] \right. \\ &\quad \left. + \Delta x \sup_{t \in [0, T]} \left[\left(1 + \|u\|_{L_x^\infty}\right) \|u\|_{H_x^4} + \|\partial_x u\|_{L_x^\infty} \|u\|_{H_x^1} \right] \right\}, \end{aligned}$$

with B_1 a constant that does not depend on u , u_0 , T , Δt and Δx . By using the Sobolev embedding $H^{\frac{1}{2}+\eta}(\mathbb{R}) \hookrightarrow L^\infty(\mathbb{R})$, with $\eta > 0$, we obtain

$$\begin{aligned} \|\epsilon^n\|_{\ell^\infty(\llbracket 0, N \rrbracket; \ell_\Delta^2)} &\leq B_1 \left\{ \Delta t \sup_{t \in [0, T]} \left[\left(1 + \|u\|_{H_x^{\frac{1}{2}+\eta}}^2\right) \|u\|_{H_x^6} \right] \right. \\ &\quad \left. + \Delta x \sup_{t \in [0, T]} \left[\left(1 + \|u\|_{H_x^{\frac{1}{2}+\eta}}\right) \|u\|_{H_x^4} + \|u\|_{H_x^{\frac{3}{2}+\eta}} \|u\|_{H_x^1} \right] \right\}. \end{aligned}$$

Theorem 3.15 enables to rewrite

$$\begin{aligned} \|\epsilon^n\|_{\ell^\infty([0,N],\ell_\Delta^2)} &\leq \Delta t B_1 C_6 C_{\frac{1}{2}+\eta}^2 e^{(2\kappa_{\frac{1}{2}+\eta}+\kappa_6)T} \left[\left(1 + \|u_0\|_{H^{\frac{1}{2}+\eta}}^2\right) \|u_0\|_{H^6} \right] \\ &\quad + \Delta x \bar{C} e^{\bar{\kappa}T} \left[\left(1 + \|u_0\|_{H^{\frac{1}{2}+\eta}}\right) \|u_0\|_{H^4} + \|u_0\|_{H^{\frac{3}{2}+\eta}} \|u_0\|_{H^1} \right], \end{aligned}$$

with $\bar{C} = \max \left(B_1 C_{\frac{1}{2}+\eta} C_4, B_1 C_{\frac{3}{2}+\eta} C_1, B_1 C_4 \right)$ and $\bar{\kappa} = \max \left(\kappa_{\frac{1}{2}+\eta} + \kappa_4, \kappa_{\frac{3}{2}+\eta} + \kappa_1, \kappa_4 \right)$.

Inequality (3.17) follows from the fact that there exists a constant $B_2 \geq 2$ such that

$$\left(1 + \|u_0\|_{H^{\frac{1}{2}+\eta}}\right) \leq B_2 \left(1 + \|u_0\|_{H^{\frac{1}{2}+\eta}}^2\right).$$

We fix $C = \max \left(B_1 C_6 C_{\frac{1}{2}+\eta}^2, B_2 \bar{C} \right)$ and $\kappa = \max \left(2\kappa_{\frac{1}{2}+\eta} + \kappa_6, \bar{\kappa} \right)$. □

3.4 Stability estimate

The stability property will be proved in stating a discrete weak-strong stability type inequality : Equation (3.37). This inequality gives an upper bound of the convergence error at time $n + 1$ with convergence error at time n . However, some terms C_i^{KdV} of (3.37) are in factor of derivatives on the convergence error $(e_j^n)_{(n,j)}$, which are not controlled at this first step : we will cancel those terms in Section 3.5.

3.4.1 Preliminary results

We here collect some discrete "Leibniz's rules" (Lemma 3.19), ℓ^2 -norm identities (Lemma 3.20) and discrete integrations by parts formulas (Lemma 3.22) which will be used in Subsection 3.4.2. Their proofs lie in Appendix B.1.

Lemma 3.19. *Let $(a_j)_{j \in \mathbb{Z}}$ and $(b_j)_{j \in \mathbb{Z}}$ be two sequences and let D_c , D_+ , D_- be the discrete operators defined in (1.12). One has, for any $j \in \mathbb{Z}$:*

$$\bullet \quad D_+ D_- (a)_j = D_- D_+ (a)_j, \quad (3.18)$$

$$\bullet \quad \begin{cases} D_+ (ab)_j = a_{j+1} D_+ (b)_j + b_j D_+ (a)_j, \\ D_- (ab)_j = a_{j-1} D_- (b)_j + b_j D_- (a)_j, \end{cases} \quad (3.19a)$$

$$\bullet \quad \begin{cases} D_+ (ab)_j = a_{j+1} D_+ (b)_j + b_j D_+ (a)_j, \\ D_- (ab)_j = a_{j-1} D_- (b)_j + b_j D_- (a)_j, \end{cases} \quad (3.19b)$$

$$\bullet \quad D_c(ab)_j = D_c(a)_j b_{j+1} + a_{j-1} D_c(b)_j, \quad (3.20)$$

$$\bullet \quad D_c(ab)_j = b_j D_c(a)_j + \frac{a_{j+1}}{2} D_+(b)_j + \frac{a_{j-1}}{2} D_-(b)_j, \quad (3.21)$$

$$\bullet \quad \begin{cases} a_j D_+ (a)_j = \frac{1}{2} D_+ (a^2)_j - \frac{\Delta x}{2} \left(D_+ (a)_j \right)^2, \\ a_j D_- (a)_j = \frac{1}{2} D_- (a^2)_j + \frac{\Delta x}{2} \left(D_- (a)_j \right)^2. \end{cases} \quad (3.22a)$$

$$\bullet \quad \begin{cases} a_j D_+ (a)_j = \frac{1}{2} D_+ (a^2)_j - \frac{\Delta x}{2} \left(D_+ (a)_j \right)^2, \\ a_j D_- (a)_j = \frac{1}{2} D_- (a^2)_j + \frac{\Delta x}{2} \left(D_- (a)_j \right)^2. \end{cases} \quad (3.22b)$$

Lemma 3.20. *For $(a_j)_{j \in \mathbb{Z}}$ a sequence in $\ell_\Delta^2(\mathbb{Z})$, one has*

$$\bullet \quad \sum_{j \in \mathbb{Z}} \left[D_+ (a)_j \right]^2 = \sum_{j \in \mathbb{Z}} \left[D_- (a)_j \right]^2, \quad (3.23)$$

$$\bullet \quad \sum_{j \in \mathbb{Z}} \left[D_c \left(\frac{a^2}{2} \right)_j \right]^2 = \sum_{j \in \mathbb{Z}} \left[D_c(a)_j \right]^2 \left(\frac{a_{j+1} + a_{j-1}}{2} \right)^2, \quad (3.24)$$

$$\bullet \quad \sum_{j \in \mathbb{Z}} \left[D_+ D_- (a)_j \right]^2 = \frac{4}{\Delta x^2} \sum_{j \in \mathbb{Z}} \left[D_+(a)_j \right]^2 - \frac{4}{\Delta x^2} \sum_{j \in \mathbb{Z}} \left[D_c(a)_j \right]^2. \quad (3.25)$$

Applying (3.25) to $D_+(a)_j$ rather than a_j enables to state

Corollary 3.21. *Let $(a_j)_{j \in \mathbb{Z}}$ be a sequence in $\ell_\Delta^2(\mathbb{Z})$. One has*

$$\sum_{j \in \mathbb{Z}} \left[D_+ D_+ D_- (a)_j \right]^2 = \frac{4}{\Delta x^2} \sum_{j \in \mathbb{Z}} \left[D_+ D_- (a)_j \right]^2 - \frac{4}{\Delta x^2} \sum_{j \in \mathbb{Z}} \left[D_+ D_c(a)_j \right]^2. \quad (3.26)$$

Lemma 3.22. *Let $(a_j)_{j \in \mathbb{Z}}$ and $(b_j)_{j \in \mathbb{Z}}$ be two sequences in $\ell_\Delta^2(\mathbb{Z})$. One has*

$$\bullet \quad \sum_{j \in \mathbb{Z}} D_+(a)_j b_j = - \sum_{j \in \mathbb{Z}} a_j D_-(b)_j, \quad (3.27)$$

$$\bullet \quad \sum_{j \in \mathbb{Z}} D_c(a)_j b_j = - \sum_{j \in \mathbb{Z}} a_j D_c(b)_j, \quad (3.28)$$

$$\bullet \quad \sum_{j \in \mathbb{Z}} a_j D_+(a)_j = - \sum_{j \in \mathbb{Z}} \frac{\Delta x}{2} \left[D_+(a)_j \right]^2, \quad (3.29)$$

$$\bullet \quad \sum_{j \in \mathbb{Z}} D_+(a)_j a_j a_{j+1} = - \sum_{j \in \mathbb{Z}} \frac{\Delta x^2}{3} \left[D_+(a)_j \right]^3, \quad (3.30)$$

$$\bullet \quad \sum_{j \in \mathbb{Z}} D_c(a)_j a_{j-1} a_{j+1} = - \sum_{j \in \mathbb{Z}} \frac{4\Delta x^2}{3} \left[D_c(a)_j \right]^3, \quad (3.31)$$

$$\bullet \quad \sum_{j \in \mathbb{Z}} a_j D_c(ab)_j = \sum_{j \in \mathbb{Z}} D_+(b)_j \frac{a_j a_{j+1}}{2}, \quad (3.32)$$

$$\bullet \quad \sum_{j \in \mathbb{Z}} D_+ D_- (a)_j D_c(ab)_j = - \sum_{j \in \mathbb{Z}} \frac{1}{\Delta x^2} D_+(b)_j a_j a_{j+1} + \sum_{j \in \mathbb{Z}} \frac{1}{\Delta x^2} D_c(b)_j a_{j-1} a_{j+1}. \quad (3.33)$$

With (3.30) and (3.31), taking $(b)_{j \in \mathbb{Z}} = \left(\frac{a_j}{2} \right)_{j \in \mathbb{Z}}$ in (3.32) and (3.33) gives the following corollary.

Corollary 3.23. *Let $(a_j)_{j \in \mathbb{Z}}$ be a sequence in $\ell_\Delta^2(\mathbb{Z})$. One has*

$$\bullet \quad \sum_{j \in \mathbb{Z}} a_j D_c \left(\frac{a^2}{2} \right)_j = - \sum_{j \in \mathbb{Z}} \frac{\Delta x^2}{12} \left[D_+(a)_j \right]^3, \quad (3.34)$$

$$\bullet \quad \sum_{j \in \mathbb{Z}} D_c \left(\frac{a^2}{2} \right)_j D_+ D_- (a)_j = \sum_{j \in \mathbb{Z}} \frac{1}{6} \left[D_+(a)_j \right]^3 - \frac{2}{3} \left[D_c(a)_j \right]^3. \quad (3.35)$$

We focus on the derivation of the ℓ_Δ^2 -stability inequality (3.37), which corresponds to a discrete equivalent of a weak-strong estimate. Combining (3.2a), (3.3) and (3.6), we obtain

$$\begin{aligned} e_j^{n+1} + \theta \Delta t D_+ D_+ D_- (e)_j^{n+1} &= e_j^n - (1 - \theta) \Delta t D_+ D_+ D_- (e)_j^n - \Delta t D_c \left(\frac{e^2}{2} \right)_j \\ &\quad - \Delta t D_c (u_\Delta e)_j^n + \frac{\tau \Delta x \Delta t}{2} D_+ D_- (e)_j^n - \Delta t \epsilon_j^n, \quad (n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}. \end{aligned} \quad (3.36)$$

Remark 3.24. There is an additional term $D_c (u_\Delta e)_j^n$ with respect to the scheme (3.2a) because of the nonlinearity.

Proposition 3.25 (ℓ_Δ^2 -stability inequality). Let $(e_j^n)_{(j,n)}$ be the convergence error defined by (3.3) with respect to Scheme (3.2a)-(3.2b). For every $\theta \in [0, 1]$, $\Delta t > 0$ and $\Delta x > 0$, for every $(n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}$ and $\gamma \in [0, \frac{1}{2})$ and $\sigma \in \{0, 1\}$, one has

$$\begin{aligned} \left\| e^{n+1} + \theta \Delta t D_+ D_+ D_- (e)^{n+1} \right\|_{\ell_\Delta^2}^2 &\leq \left\| e^n + \theta \Delta t D_+ D_+ D_- (e)_j^n \right\|_{\ell_\Delta^2}^2 + \Delta t C_a^{KdV} \|e^n\|_{\ell_\Delta^2}^2 \\ &\quad + \Delta t \|e^n - (1 - \theta) \Delta t D_+ D_+ D_- (e)^n\|_{\ell_\Delta^2}^2 + \Delta t C_b^{KdV} \|D_+ (e)^n\|_{\ell_\Delta^2}^2 + \Delta t^2 C_c^{KdV} \|D_c (e)^n\|_{\ell_\Delta^2}^2 \\ &\quad + \Delta t C_d^{KdV} \|D_+ D_- (e)^n\|_{\ell_\Delta^2}^2 + \Delta t C_e^{KdV} \|D_+ D_c (e)^n\|_{\ell_\Delta^2}^2 + \Delta t C_f^{KdV} \|D_+ D_+ D_- (e)^n\|_{\ell_\Delta^2}^2 \\ &\quad + \Delta t \|\epsilon^n\|_{\ell_\Delta^2}^2 \left\{ 1 + 4 \frac{\Delta t}{\Delta x} + \Delta t \right\}, \end{aligned} \quad (3.37)$$

where the coefficients C_i^{KdV} , for $i \in \{a, b, c, d, e, f\}$, are defined in Equations (3.38a)-(3.38f).

$$\begin{aligned} C_a^{KdV} &:= \|u_\Delta^n\|_{\ell^\infty}^2 + \|D_+ (u_\Delta)^n\|_{\ell^\infty} \left(2 - \theta + \frac{\Delta t}{\Delta x} \left[2\tau + \frac{2}{3} \|e^n\|_{\ell^\infty} + \frac{3}{2} \|(u_\Delta)^n\|_{\ell^\infty} \right] \right) \\ &\quad + \frac{\Delta t^2}{\Delta x^2} \|D_+ (u_\Delta)^n\|_{\ell^\infty}^2 + \frac{\Delta t}{\Delta x} (\|u_\Delta^n\|_{\ell^\infty}^2 + 2\tau^2), \end{aligned} \quad (3.38a)$$

$$C_b^{KdV} := \left(\frac{\Delta x}{6} D_+ (e)_j^n - \tau \right) (\Delta x - \tau \Delta t) + (1 - \theta) \Delta t \|D_+ (u_\Delta)^n\|_{\ell^\infty}^{2-\sigma}, \quad (3.38b)$$

$$C_c^{KdV} := \|e^n\|_{\ell^\infty}^2 [1 + \Delta x] + \|(u_\Delta)^n\|_{\ell^\infty}^2 - \tau^2 + 2\|e^n\|_{\ell^\infty} \|(u_\Delta)^n\|_{\ell^\infty} + \frac{2\tau}{3} \|e^n\|_{\ell^\infty}, \quad (3.38c)$$

$$C_d^{KdV} := \Delta t (1 - \theta) \left[\|D_+ (u_\Delta)^n\|_{\ell^\infty}^\sigma + \frac{\Delta x}{2} \|D_- (u_\Delta)^n\|_{\ell^\infty} \right] \quad (3.38d)$$

$$C_e^{KdV} := 2(1 - \theta) \Delta t \left\{ \|(u_\Delta)^n\|_{\ell^\infty} + \|e^n\|_{\ell^\infty} + \left[\frac{\Delta x^{\frac{1}{2}-\gamma} + \|e^n\|_{\ell^\infty} + 9\|e^n\|_{\ell^\infty}^2 \Delta x^{\gamma-\frac{1}{2}}}{2} \right] \right\} - \Delta x, \quad (3.38e)$$

$$\begin{aligned} C_f^{KdV} &:= \Delta t \left\{ (1 - 2\theta) + \frac{(1 - \theta) \Delta x^2}{2} \left[\tau + \frac{\Delta x^{\frac{1}{2}-\gamma} + \|e^n\|_{\ell^\infty} + 9\|e^n\|_{\ell^\infty}^2 \Delta x^{\gamma-\frac{1}{2}}}{2} \right] \right. \\ &\quad \left. + \Delta t (1 - \theta) \|D_+ (u_\Delta)^n\|_{\ell^\infty} \right\} - \frac{\Delta x^3}{4}. \end{aligned} \quad (3.38f)$$

Remark 3.26. The purpose of Proposition 3.42, here below, will be to control the right hand side terms C_i^{KdV} with $i \in \{b, c, d, e, f\}$ only in terms of u_Δ and not v . This is why, this inequality can be viewed as a weak-strong inequality.

Proof of Proposition 3.25. We square Equality (3.36)

$$\sum_{j \in \mathbb{Z}} \Delta x \left(e_j^{n+1} + \theta \Delta t D_+ D_+ D_- (e)_j^{n+1} \right)^2 = \|\text{RHS}_a^{KdV}\|_{\ell_\Delta^2}^2 + \|\text{RHS}_b^{KdV}\|_{\ell_\Delta^2}^2 + \|\text{RHS}_c^{KdV}\|_{\ell_\Delta^2}^2 + \Delta t^2 \|\epsilon^n\|_{\ell_\Delta^2}^2.$$

with

$$\begin{aligned}
 ||\text{RHS}_a^{KdV}||_{\ell_\Delta^2}^2 &:= \sum_{j \in \mathbb{Z}} \Delta x \left(e_j^n \right)^2 + \sum_{j \in \mathbb{Z}} \Delta x (1 - \theta)^2 \Delta t^2 \left[D_+ D_+ D_- (e)_j^n \right]^2 + \sum_{j \in \mathbb{Z}} \Delta x \Delta t^2 \left[D_c \left(\frac{e^2}{2} \right)_j^n \right]^2 \\
 &+ \sum_{j \in \mathbb{Z}} \Delta x \Delta t^2 \left[D_c (u_\Delta e)_j^n \right]^2 + \sum_{j \in \mathbb{Z}} \Delta x \frac{\tau^2 \Delta t^2 \Delta x^2}{4} \left[D_+ D_- (e)_j^n \right]^2, \\
 ||\text{RHS}_b^{KdV}||_{\ell_\Delta^2}^2 &:= - \sum_{j \in \mathbb{Z}} 2(1 - \theta) \Delta x \Delta t e_j^n D_+ D_+ D_- (e)_j^n - \sum_{j \in \mathbb{Z}} 2 \Delta x \Delta t e_j^n D_c \left(\frac{e^2}{2} \right)_j^n - \sum_{j \in \mathbb{Z}} 2 \Delta x \Delta t e_j^n D_c (u_\Delta e)_j^n \\
 &+ \sum_{j \in \mathbb{Z}} \tau \Delta x^2 \Delta t e_j^n D_+ D_- (e)_j^n + \sum_{j \in \mathbb{Z}} 2(1 - \theta) \Delta x \Delta t^2 D_+ D_+ D_- (e)_j^n D_c (u_\Delta e)_j^n \\
 &+ \sum_{j \in \mathbb{Z}} 2(1 - \theta) \Delta x \Delta t^2 D_+ D_+ D_- (e)_j^n D_c \left(\frac{e^2}{2} \right)_j^n - \sum_{j \in \mathbb{Z}} \tau \Delta x^2 \Delta t^2 (1 - \theta) D_+ D_+ D_- (e)_j^n D_+ D_- (e)_j^n \\
 &+ \sum_{j \in \mathbb{Z}} 2 \Delta x \Delta t^2 D_c \left(\frac{e^2}{2} \right)_j^n D_c (u_\Delta e)_j^n - \sum_{j \in \mathbb{Z}} \tau \Delta x^2 \Delta t^2 D_c \left(\frac{e^2}{2} \right)_j^n D_+ D_- (e)_j^n \\
 &- \sum_{j \in \mathbb{Z}} \tau \Delta x^2 \Delta t^2 D_c (u_\Delta e)_j^n D_+ D_- (e)_j^n, \tag{3.39}
 \end{aligned}$$

and

$$\begin{aligned}
 ||\text{RHS}_c^{KdV}||_{\ell_\Delta^2}^2 &:= - \sum_{j \in \mathbb{Z}} 2 \Delta x \Delta t e_j^n \left(e_j^n - (1 - \theta) \Delta t D_+ D_+ D_- (e)_j^n \right) + \sum_{j \in \mathbb{Z}} 2 \Delta x \Delta t^2 D_c \left(\frac{e^2}{2} \right)_j^n \epsilon_j^n \\
 &+ \sum_{j \in \mathbb{Z}} 2 \Delta x \Delta t^2 D_c (u_\Delta e)_j^n \epsilon_j^n - \sum_{j \in \mathbb{Z}} \tau \Delta x^2 \Delta t^2 D_+ D_- (e)_j^n \epsilon_j^n.
 \end{aligned}$$

Right Hand Side $||\text{RHS}_a^{KdV}||_{\ell_\Delta^2}^2$. We here will bound $||\text{RHS}_a^{KdV}||_{\ell_\Delta^2}^2$.

- To this aim, we use the discrete integrations by parts formulas of Subsection 3.4.1, to see that, thanks to Identity (3.24),

$$\sum_{j \in \mathbb{Z}} \Delta x \Delta t^2 \left[D_c \left(\frac{e^2}{2} \right)_j^n \right]^2 = \sum_{j \in \mathbb{Z}} \Delta x \Delta t^2 \left[D_c (e)_j^n \right]^2 \left(\frac{e_{j+1}^n + e_{j-1}^n}{2} \right)^2.$$

- To bound $\sum_{j \in \mathbb{Z}} \Delta x \Delta t^2 \left[D_c (u_\Delta e)_j^n \right]^2$, we use the following result, proved in Appendix B.1.

Lemma 3.27. Let $(a_j)_{j \in \mathbb{Z}}$ and $(b_j)_{j \in \mathbb{Z}}$ be two sequences in $\ell_\Delta^2(\mathbb{Z})$, then they satisfy the following inequality

$$\begin{aligned}
 \sum_{j \in \mathbb{Z}} \left[D_c (ab)_j \right]^2 &\leq \sum_{j \in \mathbb{Z}} \left\{ (b_j)^2 + \frac{\Delta t}{2} \left[(D_+ b_j)^2 + (D_- b_j)^2 \right] \right\} (D_c a_j)^2 \\
 &+ \sum_{j \in \mathbb{Z}} \frac{1}{2} \left\{ \frac{(b_{j-1})^2 + (b_{j+1})^2}{\Delta t} + \frac{3}{4} (D_+ b_j)^2 + \frac{3}{4} (D_- b_j)^2 \right\} (a_j)^2. \tag{3.40}
 \end{aligned}$$

Relation (3.40) gives

$$\begin{aligned}
 \sum_{j \in \mathbb{Z}} \Delta x \Delta t^2 \left[D_c (u_\Delta e)_j^n \right]^2 &\leq \sum_{j \in \mathbb{Z}} \Delta x \Delta t^2 \left\{ \left([u_\Delta]_j^n \right)^2 + \frac{\Delta t}{2} \left(D_+ (u_\Delta)_j^n \right)^2 + \frac{\Delta t}{2} \left(D_- (u_\Delta)_j^n \right)^2 \right\} \left(D_c e_j^n \right)^2 \\
 &+ \sum_{j \in \mathbb{Z}} \frac{\Delta t \Delta x}{2} \left\{ \left([u_\Delta]_{j-1}^n \right)^2 + \left([u_\Delta]_{j+1}^n \right)^2 + \frac{3 \Delta t}{4} \left(D_+ (u_\Delta)_j^n \right)^2 + \frac{3 \Delta t}{4} \left(D_- (u_\Delta)_j^n \right)^2 \right\} \left(e_j^n \right)^2.
 \end{aligned}$$

We turn our attention to the term $\sum_{j \in \mathbb{Z}} \frac{\Delta x \Delta t^3}{2} \left\{ \left(D_+(u_\Delta)_j^n \right)^2 + \left(D_-(u_\Delta)_j^n \right)^2 \right\} \left(D_c e_j^n \right)^2$. Definition of $D_c e_j^n$ gives

$$\begin{aligned} \sum_{j \in \mathbb{Z}} \frac{\Delta x \Delta t^3}{2} \left\{ \left(D_+(u_\Delta)_j^n \right)^2 + \left(D_-(u_\Delta)_j^n \right)^2 \right\} \left(D_c e_j^n \right)^2 &\leq \sum_{j \in \mathbb{Z}} \Delta x \Delta t^3 \|D_+(u_\Delta)^n\|_{\ell^\infty}^2 \left(\frac{|e_j^n|}{\Delta x} \right)^2 \\ &\leq \sum_{j \in \mathbb{Z}} \Delta x \frac{\Delta t^2}{\Delta x^2} \Delta t \|D_+(u_\Delta)^n\|_{\ell^\infty}^2 |e_j^n|^2. \end{aligned}$$

- Thanks to Relation (3.25), one has

$$\sum_{j \in \mathbb{Z}} \frac{\tau^2 \Delta t^2 \Delta x^3}{4} \left[D_+ D_-(e)_j^n \right]^2 = \sum_{j \in \mathbb{Z}} \tau^2 \Delta t^2 \Delta x \left[D_+(e)_j^n \right]^2 - \sum_{j \in \mathbb{Z}} \tau^2 \Delta t^2 \Delta x \left[D_c(e)_j^n \right]^2.$$

All this yields

$$\begin{aligned} \|\text{RHS}_a^{KdV}\|_{\ell_\Delta^2}^2 &\leq \sum_{j \in \mathbb{Z}} \Delta x \Delta t^2 \left[D_+ D_+ D_-(e)_j^n \right]^2 (\theta^2 + (1 - 2\theta)) + \sum_{j \in \mathbb{Z}} \tau^2 \Delta t^2 \Delta x \left[D_+(e)_j^n \right]^2 \\ &+ \sum_{j \in \mathbb{Z}} \Delta x \Delta t^2 \left[D_c(e)_j^n \right]^2 \left\{ \left(\frac{e_{j+1}^n + e_{j-1}^n}{2} \right)^2 + \left[(u_\Delta)_j^n \right]^2 - \tau^2 \right\} + \sum_{j \in \mathbb{Z}} \Delta x (e_j^n)^2 \left\{ 1 + \frac{\Delta t}{2} \left[([u_\Delta]_{j-1}^n)^2 \right. \right. \\ &\left. \left. + ([u_\Delta]_{j+1}^n)^2 + \frac{3\Delta t}{4} (D_+(u_\Delta)_j^n)^2 + \frac{3\Delta t}{4} (D_-(u_\Delta)_j^n)^2 + 2 \frac{\Delta t^2}{\Delta x^2} \|D_+(u_\Delta)^n\|_{\ell^\infty}^2 \right] \right\}. \end{aligned} \quad (3.41)$$

Right Hand Side $\|\text{RHS}_b^{KdV}\|_{\ell_\Delta^2}^2$. We next focus on $\|\text{RHS}_b^{KdV}\|_{\ell_\Delta^2}^2$ and on its different ten terms.

- By Relation (3.27) together with (3.29), we obtain

$$\begin{aligned} \sum_{j \in \mathbb{Z}} -2(1 - \theta) \Delta t \Delta x e_j^n D_+ D_+ D_-(e)_j^n &= \sum_{j \in \mathbb{Z}} 2\theta \Delta t \Delta x e_j^n D_+ D_+ D_-(e)_j^n + \sum_{j \in \mathbb{Z}} 2\Delta t \Delta x D_-(e)_j^n D_+ D_-(e)_j^n \\ &= \sum_{j \in \mathbb{Z}} 2\theta \Delta t \Delta x e_j^n D_+ D_+ D_-(e)_j^n - \sum_{j \in \mathbb{Z}} \Delta t \Delta x^2 \left[D_+ D_-(e)_j^n \right]^2. \end{aligned}$$

Identity (3.26) enables to write

$$\begin{aligned} \sum_{j \in \mathbb{Z}} -2(1 - \theta) \Delta t \Delta x e_j^n D_+ D_+ D_-(e)_j^n &= \sum_{j \in \mathbb{Z}} 2\theta \Delta t \Delta x e_j^n D_+ D_+ D_-(e)_j^n \\ &- \sum_{j \in \mathbb{Z}} \Delta t \frac{\Delta x^4}{4} \left[D_+ D_+ D_-(e)_j^n \right]^2 - \sum_{j \in \mathbb{Z}} \Delta t \Delta x^2 \left[D_+ D_c(e)_j^n \right]^2. \end{aligned}$$

- Thanks to Identity (3.34), one has

$$- \sum_{j \in \mathbb{Z}} 2\Delta x \Delta t e_j^n D_c \left(\frac{e^2}{2} \right)_j^n = \sum_{j \in \mathbb{Z}} \frac{\Delta x^3 \Delta t}{6} \left[D_+(e)_j^n \right]^3.$$

- Identity (3.32) gives

$$- \sum_{j \in \mathbb{Z}} 2\Delta x \Delta t e_j^n D_c (u_\Delta e)_j^n = \sum_{j \in \mathbb{Z}} -\Delta x \Delta t D_+(u_\Delta)_j^n e_j^n e_{j+1}^n \leq \sum_{j \in \mathbb{Z}} \Delta x \Delta t (e_j^n)^2 \|D_+(u_\Delta)^n\|_{\ell^\infty}.$$

- Moreover, Relations (3.18) and (3.27) imply

$$\sum_{j \in \mathbb{Z}} \tau \Delta x^2 \Delta t e_j^n D_+ D_-(e)_j^n = \sum_{j \in \mathbb{Z}} -\tau \Delta x^2 \Delta t \left[D_+(e)_j^n \right]^2.$$

- To bound $\sum_{j \in \mathbb{Z}} 2(1 - \theta) \Delta t^2 \Delta x D_+ D_+ D_-(e)_j^n D_c (u_\Delta e)_j^n$ we use Lemma 3.28, whose proof lies in Appendix B.1.

Section 3.4. Stability estimate

Lemma 3.28. Let $(a_j)_{j \in \mathbb{Z}}$ and $(b_j)_{j \in \mathbb{Z}}$ be two sequences in $\ell_\Delta^2(\mathbb{Z})$ and $\sigma \in \{0, 1\}$. One has

$$\begin{aligned} \sum_{j \in \mathbb{Z}} D_+ D_+ D_- (a)_j D_c (ab)_j &\leq \sum_{j \in \mathbb{Z}} \frac{\Delta t}{4} \left\{ |D_+ (b)_j| + |D_- (b)_j| \right\} \left(D_+ D_+ D_- (a)_j \right)^2 \\ &\quad + \sum_{j \in \mathbb{Z}} \frac{1}{2} \left\{ \|D_+ (b)\|_{\ell^\infty}^\sigma - \frac{\Delta x}{2} D_- (b)_j \right\} \left(D_+ D_- (a)_j \right)^2 + \sum_{j \in \mathbb{Z}} \frac{1}{2} \|D_+ (b)\|_{\ell^\infty}^{2-\sigma} \left(D_+ (a)_j \right)^2 \\ &\quad - \sum_{j \in \mathbb{Z}} b_j \left(D_+ D_c (a)_j \right)^2 + \sum_{j \in \mathbb{Z}} \frac{1}{4\Delta t} \left\{ |D_- (b)_j| + |D_+ (b)_j| \right\} (a_j)^2. \end{aligned} \quad (3.42)$$

Thanks to this lemma applied with $a_j = e_j^n$ and $b_j = (u_\Delta)_j^n$, one has

$$\begin{aligned} \sum_{j \in \mathbb{Z}} 2(1-\theta)\Delta t^2 \Delta x D_+ D_+ D_- (e)_j^n D_c (u_\Delta e)_j^n \\ \leq \sum_{j \in \mathbb{Z}} \frac{\Delta t^3 \Delta x}{2} (1-\theta) \left\{ |D_+ (u_\Delta)_j^n| + |D_- (u_\Delta)_j^n| \right\} \left(D_+ D_+ D_- (e)_j^n \right)^2 \\ + \sum_{j \in \mathbb{Z}} \frac{\Delta t \Delta x}{2} (1-\theta) \left\{ |D_- (u_\Delta)_j^n| + |D_+ (u_\Delta)_j^n| \right\} \left(e_j^n \right)^2 \\ + \sum_{j \in \mathbb{Z}} (1-\theta)\Delta t^2 \Delta x \left\{ \|D_+ (u_\Delta)_j^n\|_{\ell^\infty}^\sigma - \frac{\Delta x}{2} D_- (u_\Delta)_j^n \right\} \left(D_+ D_- (e)_j^n \right)^2 \\ + \sum_{j \in \mathbb{Z}} (1-\theta)\Delta x \Delta t^2 \|D_+ (u_\Delta)_j^n\|_{\ell^\infty}^{2-\sigma} \left(D_+ (e)_j^n \right)^2 - \sum_{j \in \mathbb{Z}} 2(1-\theta)\Delta x \Delta t^2 (u_\Delta)_j^n \left(D_+ D_c (e)_j^n \right)^2, \end{aligned}$$

for $\sigma \in \{0, 1\}$.

• To bound $\sum_{j \in \mathbb{Z}} 2(1-\theta)\Delta t^2 \Delta x D_+ D_+ D_- (e)_j^n D_c \left(\frac{e^2}{2} \right)_j^n$, we use Lemma 3.29 (proved in Appendix B.1).

Lemma 3.29. Let $(a_j)_{j \in \mathbb{Z}}$ be a sequence in $\ell_\Delta^2(\mathbb{Z})$ and $\gamma \in [0, \frac{1}{2})$, one has

$$\begin{aligned} \sum_{j \in \mathbb{Z}} D_+ D_+ D_- (a)_j D_c \left(\frac{a^2}{2} \right)_j &\leq \sum_{j \in \mathbb{Z}} \frac{\Delta x^{\frac{1}{2}-\gamma} + \|a\|_{\ell^\infty} + 9\|a\|_{\ell^\infty}^2 \Delta x^{\gamma-\frac{1}{2}}}{2} \left(D_+ D_- (a)_j \right)^2 \\ &\quad + \sum_{j \in \mathbb{Z}} \|a\|_{\ell^\infty} \left(D_+ D_c (a)_j \right)^2. \end{aligned}$$

Applying this inequality to $a_j = e_j^n$, one gets

$$\begin{aligned} \sum_{j \in \mathbb{Z}} 2(1-\theta)\Delta t^2 \Delta x D_+ D_+ D_- (e)_j^n D_c \left(\frac{e^2}{2} \right)_j^n \\ \leq \sum_{j \in \mathbb{Z}} \Delta t^2 \Delta x (1-\theta) \left(\Delta x^{\frac{1}{2}-\gamma} + \|e^n\|_{\ell^\infty} + 9\|e^n\|_{\ell^\infty}^2 \Delta x^{\gamma-\frac{1}{2}} \right) \left(D_+ D_- (e)_j^n \right)^2 \\ + \sum_{j \in \mathbb{Z}} 2(1-\theta)\Delta t^2 \Delta x \|e^n\|_{\ell^\infty} \left(D_+ D_c (e)_j^n \right)^2. \end{aligned}$$

Once again, Relation (3.26) transforms

$$\sum_{j \in \mathbb{Z}} \Delta t^2 \Delta x (1-\theta) \left(\Delta x^{\frac{1}{2}-\gamma} + \|e^n\|_{\ell^\infty} + 9\|e^n\|_{\ell^\infty}^2 \Delta x^{\gamma-\frac{1}{2}} \right) \left(D_+ D_- (e)_j^n \right)^2$$

$$\begin{aligned}
& \sum_{j \in \mathbb{Z}} 2(1-\theta) \Delta t^2 \Delta x D_+ D_+ D_- (e)_j^n D_c \left(\frac{e^2}{2} \right)_j \\
& \leq \sum_{j \in \mathbb{Z}} \Delta t^2 \Delta x (1-\theta) \left[\Delta x^{\frac{1}{2}-\gamma} + \|e^n\|_{\ell^\infty} + 9\|e^n\|_{\ell^\infty}^2 \Delta x^{\gamma-\frac{1}{2}} \right] \left(D_+ D_c (e)_j^n \right)^2 \\
& \quad + \sum_{j \in \mathbb{Z}} 2(1-\theta) \Delta t^2 \Delta x \|e^n\|_{\ell^\infty} \left(D_+ D_c (e)_j^n \right)^2 \\
& \quad + \sum_{j \in \mathbb{Z}} (1-\theta) \frac{\Delta t^2 \Delta x^3}{4} \left[\Delta x^{\frac{1}{2}-\gamma} + \|e^n\|_{\ell^\infty} + 9\|e^n\|_{\ell^\infty}^2 \Delta x^{\gamma-\frac{1}{2}} \right] \left(D_+ D_+ D_- (e)_j^n \right)^2.
\end{aligned}$$

Remark 3.30. Thereafter, a_j will be replaced by the unknown e_j^n whereas b_j will be replaced by the exacte solution $[u_\Delta]_j^n$. We could not use Lemma 3.28 with $b_j = \frac{a_j}{2}$ instead of Lemma 3.29 because $D_+(b)_j$ in Lemma 3.28 will be replaced by $D_+(\frac{a}{2})_j = D_+(\frac{e}{2})_j^n$ which is always unknown.

- Relation (3.29) gives

$$- \sum_{j \in \mathbb{Z}} \tau(1-\theta) \Delta t^2 \Delta x^2 D_+ D_+ D_- (e)_j^n D_+ D_- (e)_j^n = \sum_{j \in \mathbb{Z}} (1-\theta) \tau \frac{\Delta x^3 \Delta t^2}{2} \left[D_+ D_+ D_- (e)_j^n \right]^2.$$

- To deal with $\sum_{j \in \mathbb{Z}} 2\Delta x \Delta t^2 D_c \left(\frac{e^2}{2} \right)_j^n D_c(u_\Delta e)_j^n$, we use Lemma 3.31 (see Appendix B.1 for its proof).

Lemma 3.31. Let $(a_j)_{j \in \mathbb{Z}}$ and $(b_j)_{j \in \mathbb{Z}}$ be two sequences in $\ell_\Delta^2(\mathbb{Z})$, then one has

$$\begin{aligned}
\sum_{j \in \mathbb{Z}} D_c(ab)_j D_c \left(\frac{a^2}{2} \right)_j &= \sum_{j \in \mathbb{Z}} \left[D_c(a)_j \right]^2 \frac{a_{j+1}b_{j+1} + a_{j-1}b_{j-1}}{2} \\
&\quad - \sum_{j \in \mathbb{Z}} D_c(b)_j \left[D_c(a)_j \right]^3 \frac{4\Delta x^2}{3} - \sum_{j \in \mathbb{Z}} \frac{1}{3} D_c D_c(b)_j a_j^3. \quad (3.43)
\end{aligned}$$

Identity (3.43) with $a_j = e_j^n$ and $b_j = (u_\Delta)_j^n$ gives

$$\begin{aligned}
\sum_{j \in \mathbb{Z}} 2\Delta x \Delta t^2 D_c \left(\frac{e^2}{2} \right)_j^n D_c(u_\Delta e)_j^n &= \sum_{j \in \mathbb{Z}} \Delta x \Delta t^2 \left[D_c(e)_j^n \right]^2 \left[(u_\Delta)_{j+1}^n e_{j+1}^n + (u_\Delta)_{j-1}^n e_{j-1}^n \right] \\
&\quad - \sum_{j \in \mathbb{Z}} \frac{8\Delta x^3 \Delta t^2}{3} D_c(u_\Delta)_j^n \left[D_c(e)_j^n \right]^3 - \sum_{j \in \mathbb{Z}} \frac{2\Delta x \Delta t^2}{3} D_c D_c(u_\Delta)_j^n (e_j^n)^3.
\end{aligned}$$

- Relation (3.35) yields

$$- \sum_{j \in \mathbb{Z}} \tau \Delta x^2 \Delta t^2 D_c \left(\frac{e^2}{2} \right)_j^n D_+ D_- (e)_j^n = \sum_{j \in \mathbb{Z}} -\frac{\tau \Delta x^2 \Delta t^2}{6} \left[D_+(e)_j^n \right]^3 + \sum_{j \in \mathbb{Z}} \frac{2\tau \Delta x^2 \Delta t^2}{3} \left[D_c(e)_j^n \right]^3.$$

- Relation (3.33) implies

$$- \sum_{j \in \mathbb{Z}} \tau \Delta x^2 \Delta t^2 D_c(u_\Delta e)_j^n D_+ D_- (e)_j^n = \sum_{j \in \mathbb{Z}} \tau \Delta t^2 D_+(u_\Delta)_j^n e_j^n e_{j+1}^n - \sum_{j \in \mathbb{Z}} \tau \Delta t^2 D_c(u_\Delta)_j^n e_{j-1}^n e_{j+1}^n.$$

Section 3.4. Stability estimate

Thus, thanks to the Cauchy-Schwarz inequality,

$$- \sum_{j \in \mathbb{Z}} \tau \Delta x^2 \Delta t^2 D_c(u_\Delta e)_j^n D_+ D_-(e)_j^n \leq \sum_{j \in \mathbb{Z}} \tau \Delta t^2 \|D_+(u_\Delta)^n\|_{\ell^\infty} (e_j^n)^2 + \sum_{j \in \mathbb{Z}} \tau \Delta t^2 \|D_c(u_\Delta)^n\|_{\ell^\infty} (e_j^n)^2.$$

Gathering all these relations yields the following inequality, for $\sigma \in \{0, 1\}$,

$$\begin{aligned} \|\text{RHS}_b^{KdV}\|_{\ell_\Delta^2}^2 &\leq \sum_{j \in \mathbb{Z}} 2\theta \Delta t \Delta x e_j^n D_+ D_+ D_-(e)_j^n + \sum_{j \in \mathbb{Z}} \Delta x \Delta t \left\{ \|D_+ u_\Delta^n\|_{\ell^\infty} - \frac{2\Delta t}{3} D_c D_c(u_\Delta)_j^n e_j^n \right. \\ &\quad \left. + \frac{\tau \Delta t}{\Delta x} \|D_+ u_\Delta^n\|_{\ell^\infty} + \frac{\tau \Delta t}{\Delta x} \|D_c u_\Delta^n\|_{\ell^\infty} + \frac{(1-\theta)}{2} [|D_+ u_\Delta^n| + |D_- u_\Delta^n|] \right\} (e_j^n)^2 \\ &+ \sum_{j \in \mathbb{Z}} (1-\theta) \Delta t^2 \Delta x \left[\|D_+(u_\Delta)^n\|_{\ell^\infty}^\sigma + \frac{\Delta x}{2} \|D_-(u_\Delta)^n\|_{\ell^\infty} \right] (D_+ D_-(e)_j^n)^2 \\ &+ \sum_{j \in \mathbb{Z}} \Delta x \Delta t \left\{ \frac{\Delta x^2}{6} D_+(e)_j^n - \tau \Delta x - \frac{\tau \Delta t \Delta x}{6} D_+(e)_j^n + (1-\theta) \Delta t \|D_+(u_\Delta)^n\|_{\ell^\infty}^{2-\sigma} \right\} [D_+(e)_j^n]^2 \\ &+ \sum_{j \in \mathbb{Z}} \Delta x \Delta t^2 \left\{ (u_\Delta)_{j+1}^n e_{j+1}^n + (u_\Delta)_{j-1}^n e_{j-1}^n - \frac{8\Delta x^2}{3} D_c(u_\Delta)_j^n D_c(e)_j^n + \frac{2\tau \Delta x}{3} D_c(e)_j^n \right\} [D_c(e)_j^n]^2 \\ &+ \sum_{j \in \mathbb{Z}} \Delta t \Delta x \left\{ -\Delta x - 2(1-\theta) \Delta t (u_\Delta)_j^n + 2(1-\theta) \Delta t \|e^n\|_{\ell^\infty} \right. \\ &\quad \left. + \Delta t (1-\theta) \left[\Delta x^{\frac{1}{2}-\gamma} + \|e^n\|_{\ell^\infty} + 9\|e^n\|_{\ell^\infty}^2 \Delta x^{\gamma-\frac{1}{2}} \right] \right\} (D_+ D_c(e)_j^n)^2 \\ &+ \sum_{j \in \mathbb{Z}} \Delta t \Delta x \left\{ -\frac{\Delta x^3}{4} + \tau \frac{(1-\theta) \Delta x^2 \Delta t}{2} + \frac{\Delta t^2 (1-\theta)}{2} [|D_+(u_\Delta)_j^n| + |D_-(u_\Delta)_j^n|] \right. \\ &\quad \left. + (1-\theta) \frac{\Delta t \Delta x^2}{4} \left(\Delta x^{\frac{1}{2}-\gamma} + \|e^n\|_{\ell^\infty} + 9\|e^n\|_{\ell^\infty}^2 \Delta x^{\gamma-\frac{1}{2}} \right) \right\} [D_+ D_+ D_-(e)_j^n]^2. \end{aligned} \quad (3.44)$$

Right Hand Side $\|\text{RHS}_c^{KdV}\|_{\ell_\Delta^2}^2$. Let us now focus on $\|\text{RHS}_c^{KdV}\|_{\ell_\Delta^2}^2$ and its four different terms.

- From Young's inequality,

$$\begin{aligned} - \sum_{j \in \mathbb{Z}} 2\Delta x \Delta t (e_j^n - (1-\theta) \Delta t D_+ D_+ D_-(e)_j^n) \epsilon_j^n &\leq \sum_{j \in \mathbb{Z}} \Delta x \Delta t (e_j^n - (1-\theta) \Delta t D_+ D_+ D_-(e)_j^n)^2 \\ &\quad + \sum_{j \in \mathbb{Z}} \Delta t \Delta x (\epsilon_j^n)^2. \end{aligned} \quad (3.45)$$

- Once again, we apply Young's inequality to obtain

$$\sum_{j \in \mathbb{Z}} 2\Delta x \Delta t^2 D_c \left(\frac{e^2}{2} \right)_j \epsilon_j^n \leq \sum_{j \in \mathbb{Z}} \Delta t^2 [\epsilon_j^n]^2 + \sum_{j \in \mathbb{Z}} \Delta t^2 \Delta x^2 \left[D_c \left(\frac{e^2}{2} \right)_j \right]^2.$$

Then, Identity (3.24) gives

$$\sum_{j \in \mathbb{Z}} 2\Delta x \Delta t^2 D_c \left(\frac{e^2}{2} \right)_j \epsilon_j^n \leq \sum_{j \in \mathbb{Z}} \Delta t^2 [\epsilon_j^n]^2 + \sum_{j \in \mathbb{Z}} \Delta t^2 \Delta x^2 [D_c(e)_j^n]^2 \left(\frac{e_{j+1}^n + e_{j-1}^n}{2} \right)^2,$$

- One also has

$$\begin{aligned} \sum_{j \in \mathbb{Z}} 2\Delta x \Delta t^2 D_c(u_\Delta e)_j^n \epsilon_j^n &= \sum_{j \in \mathbb{Z}} \Delta t^2 \left((u_\Delta)_{j+1}^n e_{j+1}^n - (u_\Delta)_{j-1}^n e_{j-1}^n \right) \epsilon_j^n \\ &\leq \sum_{j \in \mathbb{Z}} \Delta t^2 (e_j^n)^2 \|u_\Delta^n\|_{\ell^\infty}^2 + \sum_{j \in \mathbb{Z}} \Delta t^2 (\epsilon_j^n)^2. \end{aligned}$$

• Finally, we see that, thanks to Young's inequality

$$\begin{aligned} \sum_{j \in \mathbb{Z}} -\tau \Delta x^2 \Delta t^2 D_+ D_- (e)_j^n \epsilon_j^n &= \sum_{j \in \mathbb{Z}} -\tau \Delta t^2 \left(e_{j+1}^n - 2e_j^n + e_{j-1}^n \right) \epsilon_j^n \\ &\leq \sum_{j \in \mathbb{Z}} 2\tau^2 \Delta t^2 \left(e_j^n \right)^2 + \sum_{j \in \mathbb{Z}} 2\Delta t^2 \left(\epsilon_j^n \right)^2. \end{aligned}$$

Final inequality

To conclude, the convergence error satisfies the following inequality

$$\begin{aligned} \sum_{j \in \mathbb{Z}} \Delta x \left(e_j^{n+1} + \theta \Delta t D_+ D_+ D_- (e)_j^{n+1} \right)^2 &\leq \sum_{j \in \mathbb{Z}} \Delta x \left(e_j^n + \theta \Delta t D_+ D_+ D_- (e)_j^n \right)^2 + \Delta t \|e^n\|_{\ell_\Delta^2}^2 A_a \\ &+ \sum_{j \in \mathbb{Z}} \Delta t \Delta x \left(e_j^n - (1 - \theta) \Delta t D_+ D_+ D_- (e)_j^n \right)^2 + \sum_{j \in \mathbb{Z}} \Delta x \Delta t A_b \left[D_+ (e)_j^n \right]^2 + \sum_{j \in \mathbb{Z}} \Delta x \Delta t^2 A_c \left[D_c (e)_j^n \right]^2 \\ &+ \sum_{j \in \mathbb{Z}} \Delta x \Delta t A_d \left(D_+ D_- (e)_j^n \right)^2 + \sum_{j \in \mathbb{Z}} \Delta x \Delta t A_e \left(D_+ D_c (e)_j^n \right)^2 + \sum_{j \in \mathbb{Z}} \Delta t \Delta x A_f \left[D_+ D_+ D_- (e)_j^n \right]^2 \\ &+ \Delta t \|e^n\|_{\ell_\Delta^2}^2 \left\{ 1 + 4 \frac{\Delta t}{\Delta x} + \Delta t \right\}, \end{aligned}$$

with

$$\begin{aligned} A_a &:= \frac{\left([u_\Delta]_{j-1}^n \right)^2}{2} + \frac{\left([u_\Delta]_{j+1}^n \right)^2}{2} + \frac{\Delta t}{2} \left[\frac{3}{4} \left(D_- (u_\Delta)_j^n \right)^2 + \frac{3}{4} \left(D_+ (u_\Delta)_j^n \right)^2 \right] + \frac{\Delta t^2}{\Delta x^2} \|D_+ (u_\Delta)^n\|_{\ell^\infty}^2 \\ &\quad + \frac{(1 - \theta)}{2} \left[|D_- (u_\Delta)_j^n| + |D_+ (u_\Delta)_j^n| \right] + \|D_+ (u_\Delta)^n\|_{\ell^\infty} \left(1 + \frac{\tau \Delta t}{\Delta x} \right) \\ &\quad + \frac{\tau \Delta t}{\Delta x} \|D_c (u_\Delta)^n\|_{\ell^\infty} - \frac{2\Delta t}{3} D_c D_c (u_\Delta)_j^n e_j^n + \frac{\Delta t}{\Delta x} \left(\| (u_\Delta)^n \|_{\ell^\infty}^2 + 2\tau^2 \right), \\ A_b &:= \tau^2 \Delta t + \frac{\Delta x^2}{6} D_+ (e)_j^n - \tau \Delta x - \frac{\tau \Delta x \Delta t}{6} D_+ (e)_j^n + (1 - \theta) \Delta t \|D_+ (u_\Delta)^n\|_{\ell^\infty}^{2-\sigma}, \\ A_c &:= \left(\frac{e_{j+1}^n + e_{j-1}^n}{2} \right)^2 [1 + \Delta x] + \left([u_\Delta]_j^n \right)^2 - \tau^2 + (u_\Delta)_{j+1}^n e_{j+1}^n + (u_\Delta)_{j-1}^n e_{j-1}^n \\ &\quad - \frac{8\Delta x^2}{3} D_c (u_\Delta)_j^n D_c (e)_j^n + \frac{2\tau \Delta x}{3} D_c (e)_j^n, \\ A_d &:= (1 - \theta) \Delta t \left[\|D_+ (u_\Delta)^n\|_{\ell^\infty}^\sigma + \frac{\Delta x}{2} \|D_- (u_\Delta)^n\|_{\ell^\infty} \right], \\ A_e &:= -2(1 - \theta) \Delta t (u_\Delta)_j^n + 2(1 - \theta) \Delta t \|e^n\|_{\ell^\infty} - \Delta x + \Delta t (1 - \theta) \left[\Delta x^{\frac{1}{2}-\gamma} + \|e^n\|_{\ell^\infty} + 9 \|e^n\|_{\ell^\infty}^2 \Delta x^{\gamma-\frac{1}{2}} \right], \end{aligned}$$

and

$$\begin{aligned} A_f &:= \Delta t \left[(1 - 2\theta) + \frac{\tau(1 - \theta) \Delta x^2}{2} + \frac{\Delta t(1 - \theta)}{2} \left[|D_+ (u_\Delta)_j^n| + |D_- (u_\Delta)_j^n| \right] \right. \\ &\quad \left. + (1 - \theta) \frac{\Delta x^2}{4} \left(\Delta x^{\frac{1}{2}-\gamma} + \|e^n\|_{\ell^\infty} + 9 \|e^n\|_{\ell^\infty}^2 \Delta x^{\gamma-\frac{1}{2}} \right) \right] - \frac{\Delta x^3}{4}. \end{aligned}$$

• Since $\|D_c D_c (u_\Delta)^n\|_{\ell^\infty} \leq \frac{1}{\Delta x} \|D_c (u_\Delta)^n\|_{\ell^\infty}$, $\|D_c (u_\Delta)^n\|_{\ell^\infty} \leq \|D_+ (u_\Delta)^n\|_{\ell^\infty}$ and $\Delta t \|D_+ (u_\Delta)^n\|_{\ell^\infty} \leq \frac{2\Delta t}{\Delta x} \|u_\Delta^n\|_{\ell^\infty}$, then

$$A_a \leq C_a^{KdV},$$

Section 3.5. Rate of convergence

where C_a^{KdV} is defined by (3.38a).

- For A_b , we recognize the definition (3.38b) of C_b^{KdV} .
- For the term A_c , we have

$$A_c \leq \|e^n\|_{\ell^\infty}^2 [1 + \Delta x] + \|(u_\Delta)^n\|_{\ell^\infty}^2 - \tau^2 + \frac{1}{3}e_{j+1}^n (u_\Delta)_{j+1}^n + \frac{1}{3}e_{j-1}^n (u_\Delta)_{j-1}^n + \frac{2}{3}(u_\Delta)_{j+1}^n e_{j-1}^n + \frac{2}{3}(u_\Delta)_{j-1}^n e_{j+1}^n + \frac{2\tau}{3}\|e^n\|_{\ell^\infty}.$$

Thus, one has $A_c \leq C_c^{KdV}$ (3.38c).

- Furthermore, from (3.38d) and (3.38e)

$$A_d = C_d^{KdV}$$

and

$$A_e \leq C_e^{KdV}.$$

- At last, we see that $A_f \leq C_f^{KdV}$ defined by (3.38f). Proposition 3.25 is then proved. $\square$

3.5 Rate of convergence

In the left hand side of the ℓ_Δ^2 -stability inequality (3.37), e_j^{n+1} appears in the operator $I + \theta\Delta t D_+ D_+ D_-$. The study of this dispersive operator is the aim of Subsection 3.5.1.

In the right hand side of (3.37), $D_+(e)_j^n, D_+ D_-(e)_j^n$ appear in factor of some terms C_i^{KdV} . Since we have no control on these derivatives of the convergence error, we reorganize terms C_i^{KdV} in Subsection 3.5.2 to obtain non-positive terms : the E_i and F_i terms of Corollaries 3.38 and 3.40. We notice that the reorganization is not exactly the same for $\theta \geq \frac{1}{2}$ and $\theta < \frac{1}{2}$, therefore, we state two corollaries.

In Subsection 3.5.3, Proposition 3.42 gives correct CFL hypothesis to cancel extra terms E_i and F_i and concludes the convergence proof, by an induction method.

3.5.1 Properties of dispersive operators

Definition 3.32. Let $(a_j)_{j \in \mathbb{Z}}$ a sequence in $\ell_\Delta^2(\mathbb{Z})$, we define the two following dispersive operators :

$$(\mathcal{A}a)_j := a_j + \theta\Delta t D_+ D_+ D_-(a)_j, \quad (3.46)$$

$$(\mathcal{B}a)_j := a_j - (1 - \theta)\Delta t D_+ D_+ D_-(a)_j. \quad (3.47)$$

Proposition 3.33. For every $\Delta t > 0$ and $\Delta x > 0$, the operator $\mathcal{A}$ is

- continuous (with a norm depending of $\frac{\Delta t}{\Delta x^3}$) from ℓ_Δ^2 to ℓ_Δ^2 ,
- invertible.

Moreover, one has the following inequality, for any sequence $(a_j)_{j \in \mathbb{Z}} \in \ell_\Delta^2(\mathbb{Z})$

$$\|a\|_{\ell_\Delta^2}^2 \leq \|\mathcal{A}a\|_{\ell_\Delta^2}^2 \leq \left\{ 1 + \frac{16\theta\Delta t}{\Delta x^3} \left[1 + \frac{4\theta\Delta t}{\Delta x^3} \right] \right\} \|a\|_{\ell_\Delta^2}^2. \quad (3.48)$$

Remark 3.34. Inequality (3.48) implies that the inverse of $\mathcal{A}$ is continuous too from ℓ_Δ^2 to ℓ_Δ^2 with a norm independent of $\frac{\Delta t}{\Delta x^3}$.

Chapitre 3. Error estimates for a finite difference scheme for the Korteweg-de Vries equation

Proof. Given $a \in \ell_\Delta^2(\mathbb{Z})$, we may define the function $\widehat{a} \in L^2([0, 1])$ and the shift operator $\mathcal{S}^\ell$ such as in the beginning of Subsection 2.2.1.3. The function $\widehat{\mathcal{A}a}$ corresponds to the denominator of the amplification factor $A_+^{2p+1}(\xi)$ computed in (2.18), with $p = 1$. Therefore, we obtain

$$\widehat{\mathcal{A}a}(\xi) = \widehat{a} \left\{ 1 + 8i\theta \frac{\Delta t}{\Delta x^3} e^{-i\pi\xi} \sin^3(\pi\xi) \right\}.$$

The operator $\mathcal{A}$ is so inversible and its inverse is defined by $\widehat{\mathcal{A}^{-1}a}(\xi) = \frac{1}{1 + 8i\theta \frac{\Delta t}{\Delta x^3} e^{-i\pi\xi} \sin^3(\pi\xi)} \widehat{a}(\xi)$. Moreover, this operator (and its inverse) is continuous since

$$\|\mathcal{A}a\|_{\ell_\Delta^2}^2 = \Delta x \int_0^1 \left| 1 + 8i\theta \frac{\Delta t}{\Delta x^3} e^{-i\pi\xi} \sin^3(\pi\xi) \right|^2 |\widehat{a}(\xi)|^2 d\xi.$$

Yet, the module $\left| 1 + 8i\theta \frac{\Delta t}{\Delta x^3} e^{-i\pi\xi} \sin^3(\pi\xi) \right|^2$ is equal to

$$\begin{aligned} \left| 1 + 8i\theta \frac{\Delta t}{\Delta x^3} e^{-i\pi\xi} \sin^3(\pi\xi) \right|^2 &= \left(1 + 8\theta \frac{\Delta t}{\Delta x^3} \sin^4(\pi\xi) \right)^2 + \left(8\theta \frac{\Delta t}{\Delta x^3} \cos(\pi\xi) \sin^3(\pi\xi) \right)^2 \\ &= 1 + 16\theta \frac{\Delta t}{\Delta x^3} \sin^4(\pi\xi) \left(1 + 4\theta \frac{\Delta t}{\Delta x^3} \sin^2(\pi\xi) \right) \\ &\in [1, 1 + 16\theta \frac{\Delta t}{\Delta x^3} \left(1 + 4\theta \frac{\Delta t}{\Delta x^3} \right)]. \end{aligned}$$

Thus, the operator $\mathcal{A}$ verifies

$$\Delta x \int_0^1 |\widehat{a}(\xi)|^2 d\xi \leq \|\mathcal{A}a\|_{\ell_\Delta^2}^2 \leq \left\{ 1 + 16\theta \frac{\Delta t}{\Delta x^3} \left(1 + 4\theta \frac{\Delta t}{\Delta x^3} \right) \right\} \Delta x \int_0^1 |\widehat{a}(\xi)|^2 d\xi.$$

We conclude by using Parseval identity (2.17). $\square$

Remark 3.35. The norm of the inverse operator $\mathcal{A}^{-1}$ is upper bounded by 1 (independent of $\frac{\Delta t}{\Delta x^3}$). This independence is crucial to be able to impose a hyperbolic Courant-Friedrichs-Lewy condition, and not a dispersive one.

The operator $\mathcal{A}$ enables us to control not only the ℓ_Δ^2 -norm (as proved in Proposition 3.33) but also an h_Δ^2 -discrete norm and h_Δ^3 -discrete norm as explained in the following proposition.

Proposition 3.36. Let $\mathcal{A}$ be the operator defined by (3.46), then for any sequence $(a_j)_{j \in \mathbb{Z}}$ in $\ell_\Delta^2(\mathbb{Z})$, one has

$$\|\mathcal{A}a\|_{\ell_\Delta^2}^2 = \|a\|_{\ell_\Delta^2}^2 + \theta \Delta t \Delta x \|D_+ D_-(a)\|_{\ell_\Delta^2}^2 + \theta^2 \Delta t^2 \|D_+ D_+ D_-(a)\|_{\ell_\Delta^2}^2.$$

Proof. We develop the ℓ_Δ^2 -norm of $(\mathcal{A}a_j)_{j \in \mathbb{Z}}$:

$$\sum_{j \in \mathbb{Z}} \Delta x (a_j + \theta \Delta t D_+ D_+ D_-(a))_j^2 = \sum_{j \in \mathbb{Z}} \Delta x (a_j)^2 + \sum_{j \in \mathbb{Z}} 2\theta \Delta x \Delta t a_j D_+ D_+ D_-(a)_j + \sum_{j \in \mathbb{Z}} \theta^2 \Delta x \Delta t^2 (D_+ D_+ D_-(a))_j^2.$$

Let us focus on the cross term. Discrete integration by parts (3.27) together with (3.29) (with $D_-(a)_j$ instead of a_j) give

$$\sum_{j \in \mathbb{Z}} 2\theta \Delta t \Delta x a_j D_+ D_+ D_-(a)_j = - \sum_{j \in \mathbb{Z}} 2\theta \Delta t \Delta x D_-(a)_j D_+ D_-(a)_j = \sum_{j \in \mathbb{Z}} \theta \Delta t \Delta x^2 (D_+ D_-(a))_j^2,$$

which concludes the proof. $\square$

Furthermore, the following proposition enables us to compare $\|\mathcal{B}a\|_{\ell_\Delta^2}$ to $\|\mathcal{A}a\|_{\ell_\Delta^2}$ modulo a classical dispersive CFL condition.

Section 3.5. Rate of convergence

Proposition 3.37. For $\theta \in [0, 1]$, assume the following CFL condition be true $\Delta t(1 - 2\theta) \leq \frac{\Delta x^3}{4}$, then, for any sequence $(a_j)_{j \in \mathbb{Z}}$ in $\ell_\Delta^2(\mathbb{Z})$, it holds

$$\sum_{j \in \mathbb{Z}} \Delta x \left(a_j - (1 - \theta) \Delta t D_+ D_+ D_- (a)_j \right)^2 \leq \|\mathcal{A}a\|_{\ell_\Delta^2}^2. \quad (3.49)$$

Proof. We develop the expression

$$\begin{aligned} \sum_{j \in \mathbb{Z}} \Delta x \left(a_j - (1 - \theta) \Delta t D_+ D_+ D_- (a)_j \right)^2 &= \sum_{j \in \mathbb{Z}} \Delta x \left(a_j + \theta \Delta t D_+ D_+ D_- (a)_j \right)^2 - \sum_{j \in \mathbb{Z}} 2\Delta x \Delta t a_j D_+ D_+ D_- (a)_j \\ &\quad + \sum_{j \in \mathbb{Z}} \Delta x \Delta t^2 (1 - 2\theta) \left(D_+ D_+ D_- (a)_j \right)^2. \end{aligned}$$

By applying Relations (3.27) and (3.29) (with $D_- (a)_j$ instead of a_j), the previous equation becomes

$$\begin{aligned} \sum_{j \in \mathbb{Z}} \Delta x \left(a_j - (1 - \theta) \Delta t D_+ D_+ D_- (a)_j \right)^2 &= \sum_{j \in \mathbb{Z}} \Delta x \left(a_j + \theta \Delta t D_+ D_+ D_- (a)_j \right)^2 - \sum_{j \in \mathbb{Z}} \Delta x^2 \Delta t \left(D_+ D_- (a)_j \right)^2 \\ &\quad + \sum_{j \in \mathbb{Z}} \Delta x \Delta t^2 (1 - 2\theta) \left(D_+ D_+ D_- (a)_j \right)^2. \end{aligned}$$

CASE $\theta \geq \frac{1}{2}$: Proposition 3.37 is proved.

CASE $\theta < \frac{1}{2}$: We transform the third derivative thanks to Identity (3.26), to obtain

$$\begin{aligned} \sum_{j \in \mathbb{Z}} \Delta x \left(a_j - (1 - \theta) \Delta t D_+ D_+ D_- (a)_j \right)^2 &= \sum_{j \in \mathbb{Z}} \Delta x \left(a_j + \theta \Delta t D_+ D_+ D_- (a)_j \right)^2 \\ &\quad - \sum_{j \in \mathbb{Z}} \Delta x^2 \Delta t \left(D_+ D_- (a)_j \right)^2 + \sum_{j \in \mathbb{Z}} \frac{4\Delta t^2(1 - 2\theta)}{\Delta x} \left(D_+ D_- (a)_j \right)^2 - \sum_{j \in \mathbb{Z}} \frac{4\Delta t^2(1 - 2\theta)}{\Delta x} \left(D_+ D_c (a)_j \right)^2. \end{aligned}$$

Since $\Delta t(1 - 2\theta) \leq \frac{\Delta x^3}{4}$, the term $\frac{4\Delta t^2(1 - 2\theta)}{\Delta x}$ is upper bounded by $\Delta t \Delta x^2$, which transforms the previous equation into

$$\begin{aligned} \sum_{j \in \mathbb{Z}} \Delta x \left(a_j - (1 - \theta) \Delta t D_+ D_+ D_- (a)_j \right)^2 &= \sum_{j \in \mathbb{Z}} \Delta x \left(a_j + \theta \Delta t D_+ D_+ D_- (a)_j \right)^2 \\ &\quad - \sum_{j \in \mathbb{Z}} \Delta x^2 \Delta t \left(D_+ D_- (a)_j \right)^2 + \sum_{j \in \mathbb{Z}} \Delta t \Delta x^2 \left(D_+ D_- (a)_j \right)^2 - \sum_{j \in \mathbb{Z}} \frac{4\Delta t^2(1 - 2\theta)}{\Delta x} \left(D_+ D_c (a)_j \right)^2. \end{aligned}$$

The conclusion of the proposition is a straightforward consequence, since $1 - 2\theta > 0$. $\square$

3.5.2 Simplification of Inequality (3.37)

The previous study of the dispersive operator $\mathcal{A}$ enables us to reorganize terms in ℓ_Δ^2 -stability inequality (3.37) in a way simpler to study : signs of new terms are easier to identify. The reorganization is not exactly the same for $\theta \geq \frac{1}{2}$ and $\theta < \frac{1}{2}$, as seen in the two following corollaries of Proposition 3.25.

Corollary 3.38 (Corollary of Proposition 3.25). Consider Scheme (3.2a)-(3.2b). Let $(e_j^n)_{(j,n)}$ be the convergence error defined by (3.3). Then, for every $n \in \llbracket 0, N \rrbracket$, $\gamma \in [0, \frac{1}{2})$ and $\theta \geq \frac{1}{2}$, one has

$$\begin{aligned} \|\mathcal{A}e^{n+1}\|_{\ell_\Delta^2}^2 &\leq \|\mathcal{A}e^n\|_{\ell_\Delta^2}^2 \left[1 + \Delta t G_a^{KdV} \right] + \Delta t E_b \|D_+ (e)^n\|_{\ell_\Delta^2}^2 + \Delta t^2 E_c \|D_c (e)^n\|_{\ell_\Delta^2}^2 \\ &\quad + \Delta t E_e \|D_+ D_c (e)^n\|_{\ell_\Delta^2}^2 + \Delta t E_f \|D_+ D_+ D_- (e)^n\|_{\ell_\Delta^2}^2 + \Delta t \|\epsilon^n\|_{\ell_\Delta^2}^2 \left\{ 1 + 4 \frac{\Delta t}{\Delta x} + \Delta t \right\}, \end{aligned} \quad (3.50)$$

with

$$G_a^{KdV} := \|u_\Delta^n\|_{\ell^\infty}^2 \left(1 + \frac{\Delta t}{\Delta x}\right) + \|D_+(u_\Delta)^n\|_{\ell^\infty} \left(7 + \frac{\Delta t}{\Delta x} \left[2\tau + \frac{2}{3}\|e^n\|_{\ell^\infty} + \frac{3}{2}\|(u_\Delta)^n\|_{\ell^\infty}\right]\right) \\ + \|D_+(u_\Delta)^n\|_{\ell^\infty}^2 \left[\sqrt{2}\frac{\sqrt{\Delta t}}{\sqrt{\Delta x}} + \frac{\Delta t^2}{\Delta x^2}\right] + 1 + 2\tau^2 \frac{\Delta t}{\Delta x}, \quad (3.51a)$$

$$E_b := \left(\frac{\Delta x}{6} D_+(e)_j^n - \tau\right) (\Delta x - \tau \Delta t), \quad (3.51b)$$

$$E_c := \|(u_\Delta)^n\|_{\ell^\infty}^2 + \left\{ \|e^n\|_{\ell^\infty}^2 [1 + \Delta x] + 2\|e^n\|_{\ell^\infty} \|(u_\Delta)^n\|_{\ell^\infty} + \frac{2\tau}{3}\|e^n\|_{\ell^\infty} \right\} - \tau^2, \quad (3.51c)$$

$$E_e := 2(1 - \theta)\Delta t \left\{ \|(u_\Delta)^n\|_{\ell^\infty} + \|e^n\|_{\ell^\infty} + \frac{1}{2} + \left[\frac{\Delta x^{\frac{1}{2}-\gamma} + \|e^n\|_{\ell^\infty} + 9\|e^n\|_{\ell^\infty}^2 \Delta x^{\gamma-\frac{1}{2}}}{2} \right] \right\} - \Delta x, \quad (3.51d)$$

$$E_f := \Delta t \left\{ (1 - 2\theta) + \frac{(1 - \theta)\Delta x^2}{2} \left[\tau + \frac{1}{2} + \frac{\Delta x^{\frac{1}{2}-\gamma} + \|e^n\|_{\ell^\infty} + 9\|e^n\|_{\ell^\infty}^2 \Delta x^{\gamma-\frac{1}{2}}}{2} \right] \right\} - \frac{\Delta x^3}{4}. \quad (3.51e)$$

Remark 3.39. Corollary 3.38 is, in fact, true for all $\theta \neq 0$ provided that $\Delta t(1 - 2\theta) \leq \frac{\Delta x^3}{4}$ for $\theta < \frac{1}{2}$, but we essentially use it for $\theta \geq \frac{1}{2}$.

Proof. We choose $\sigma = 0$ in Inequality (3.37).

- First, we upper bound $\|e^n - (1 - \theta)\Delta t D_+ D_+ D_-(e)^n\|_{\ell_\Delta^2}^2$ in (3.37) by $\|e^n + \theta\Delta t D_+ D_+ D_-(e)^n\|_{\ell_\Delta^2}^2$ thanks to Proposition 3.37.
- We tranform C_b^{KdV} in (3.38b) into

$$C_b^{KdV} = E_b + (1 - \theta)\Delta t \|D_+(u_\Delta)^n\|_{\ell^\infty}^2,$$

with

$$E_b = \left(\frac{\Delta x}{6} D_+(e)_j^n - \tau\right) (\Delta x - \tau \Delta t). \quad (3.52)$$

The C_b^{KdV} -term in (3.37) thus is

$$\sum_{j \in \mathbb{Z}} \Delta x \Delta t C_b^{KdV} \left(D_+ e_j^n\right)^2 = \sum_{j \in \mathbb{Z}} \Delta x \Delta t E_b \left(D_+ e_j^n\right)^2 + (1 - \theta)\Delta t^2 \|D_+ u_\Delta^n\|_{\ell^\infty}^2 \|D_+ e^n\|_{\ell_\Delta^2}^2. \quad (3.53)$$

For any sequence $(a_j)_{j \in \mathbb{Z}}$ in $\ell_\Delta^2(\mathbb{Z})$, the following Gagliardo-Nirenberg inequality

$$\|D_+(a)\|_{\ell_\Delta^2}^2 \leq \|a\|_{\ell_\Delta^2} \|D_+ D_-(a)\|_{\ell_\Delta^2}$$

is valid even with the ℓ_Δ^2 -norm. We will use it on $\|D_+(e)^n\|_{\ell_\Delta^2}^2$ in (3.53), to obtain

$$(1 - \theta)\Delta t^2 \|D_+(u_\Delta)^n\|_{\ell^\infty}^2 \|D_+ e^n\|_{\ell_\Delta^2}^2 \leq (1 - \theta)\Delta t^2 \|D_+(u_\Delta)^n\|_{\ell^\infty}^2 \frac{\|e^n\|_{\ell_\Delta^2}^2 \sqrt{\theta \Delta t \Delta x} \|D_+ D_-(e)^n\|_{\ell_\Delta^2}^2}{\sqrt{\theta \Delta t \Delta x}}.$$

Proposition 3.36 enables to make $\|Ae^n\|_{\ell_\Delta^2}^2$ appear and

$$(1 - \theta)\Delta t^2 \|D_+(u_\Delta)^n\|_{\ell^\infty}^2 \|D_+ e^n\|_{\ell_\Delta^2}^2 \leq (1 - \theta)\Delta t^2 \|D_+(u_\Delta)^n\|_{\ell^\infty}^2 \frac{\|e^n + \theta\Delta t D_+ D_+ D_-(e)^n\|_{\ell_\Delta^2}^2}{\sqrt{\theta \Delta t \Delta x}} \\ \leq \frac{(1 - \theta)}{\sqrt{\theta}} \frac{\sqrt{\Delta t}}{\sqrt{\Delta x}} \Delta t \|D_+(u_\Delta)^n\|_{\ell^\infty}^2 \|e^n + \theta\Delta t D_+ D_+ D_-(e)^n\|_{\ell_\Delta^2}^2.$$

Section 3.5. Rate of convergence

- As a third step, we transform the C_d^{KdV} -term of (3.37) (recall that $\sigma = 0$) :

$$\begin{aligned} \sum_{j \in \mathbb{Z}} \Delta x \Delta t C_d^{KdV} \left(D_+ D_- (e)_j^n \right)^2 &= \sum_{j \in \mathbb{Z}} (1 - \theta) \Delta t^2 \Delta x \left(D_+ D_- (e)_j^n \right)^2 \\ &\quad + \frac{(1 - \theta)}{2\theta} \Delta t \|D_+(u_\Delta)^n\|_{\ell^\infty} \theta \Delta t \Delta x \|D_+ D_- (e)^n\|_{\ell_\Delta^2}^2. \end{aligned}$$

Relation (3.26) allows to rewrite the term $\sum_{j \in \mathbb{Z}} (1 - \theta) \Delta t^2 \Delta x \left(D_+ D_- (e)_j^n \right)^2$:

$$\sum_{j \in \mathbb{Z}} (1 - \theta) \Delta t^2 \Delta x \left(D_+ D_- (e)_j^n \right)^2 = \sum_{j \in \mathbb{Z}} (1 - \theta) \Delta t^2 \Delta x \left(D_+ D_c (e)_j^n \right)^2 + \sum_{j \in \mathbb{Z}} (1 - \theta) \frac{\Delta t^2 \Delta x^3}{4} \left(D_+ D_+ D_- (e)_j^n \right)^2.$$

Proposition 3.36 gives

$$\frac{(1 - \theta)}{2\theta} \Delta t \|D_+(u_\Delta)^n\|_{\ell^\infty} \theta \Delta t \Delta x \|D_+ D_- (e)^n\|_{\ell_\Delta^2}^2 \leq \frac{(1 - \theta)}{2\theta} \Delta t \|D_+(u_\Delta)^n\|_{\ell^\infty} \|e^n + \theta \Delta t D_+ D_+ D_- (e)^n\|_{\ell_\Delta^2}^2.$$

- Eventually, we focus on the C_f^{KdV} -term in (3.37). We decompose C_f^{KdV} into

$$C_f^{KdV} = C_g + \Delta t^2 (1 - \theta) \|D_+(u_\Delta)^n\|_{\ell^\infty}$$

with

$$C_g = \Delta t \left\{ (1 - 2\theta) + \frac{(1 - \theta) \Delta x^2}{2} \left[\tau + \frac{\Delta x^{\frac{1}{2} - \gamma} + \|e^n\|_{\ell^\infty} + 9 \|e^n\|_{\ell^\infty}^2 \Delta x^{\gamma - \frac{1}{2}}}{2} \right] \right\} - \frac{\Delta x^3}{4} \quad (3.54)$$

which leads to the following inequality (thanks to Proposition 3.36) :

$$\begin{aligned} \Delta t C_f^{KdV} \|D_+ D_+ D_- (e)^n\|_{\ell_\Delta^2}^2 &= \Delta t C_g \|D_+ D_+ D_- (e)^n\|_{\ell_\Delta^2}^2 + \frac{(1 - \theta)}{\theta^2} \Delta t \|D_+(u_\Delta)^n\|_{\ell^\infty} \|\theta \Delta t D_+ D_+ D_- (e)^n\|_{\ell_\Delta^2}^2 \\ &\leq \Delta t C_g \|D_+ D_+ D_- (e)^n\|_{\ell_\Delta^2}^2 + \frac{(1 - \theta)}{\theta^2} \Delta t \|D_+(u_\Delta)^n\|_{\ell^\infty} \|e^n + \theta \Delta t D_+ D_+ D_- (e)^n\|_{\ell_\Delta^2}^2. \end{aligned}$$

Thanks to all the previous relations, we rewrite Inequality (3.37) as

$$\begin{aligned} \|\mathcal{A}e^{n+1}\|_{\ell_\Delta^2}^2 &\leq \|\mathcal{A}e^n\|_{\ell_\Delta^2}^2 \left[1 + \Delta t E_a^{KdV} \right] + \Delta t E_b \|D_+ (e)^n\|_{\ell_\Delta^2}^2 + \Delta t^2 C_c^{KdV} \|D_c (e)^n\|_{\ell_\Delta^2}^2 \\ &\quad + \Delta t \left[C_e^{KdV} + (1 - \theta) \Delta t \right] \|D_+ D_c (e)^n\|_{\ell_\Delta^2}^2 + \Delta t \left[C_g + (1 - \theta) \frac{\Delta t \Delta x^2}{4} \right] \|D_+ D_+ D_- (e)^n\|_{\ell_\Delta^2}^2 \\ &\quad + \Delta t \|e^n\|_{\ell_\Delta^2}^2 \left\{ 1 + 4 \frac{\Delta t}{\Delta x} + \Delta t \right\}, \end{aligned}$$

with

$$\begin{aligned} E_a^{KdV} &= \|u_\Delta^n\|_{\ell^\infty}^2 \left(1 + \frac{\Delta t}{\Delta x} \right) + \|D_+(u_\Delta)^n\|_{\ell^\infty} \left(2 - \theta + \frac{1 - \theta}{2\theta} + \frac{1 - \theta}{\theta^2} + \frac{\Delta t}{\Delta x} \left[2\tau + \frac{2}{3} \|e^n\|_{\ell^\infty} + \frac{3}{2} \|(u_\Delta)^n\|_{\ell^\infty} \right] \right) \\ &\quad + \|D_+(u_\Delta)^n\|_{\ell^\infty}^2 \left[\frac{(1 - \theta)}{\sqrt{\theta}} \frac{\sqrt{\Delta t}}{\sqrt{\Delta x}} + \frac{\Delta t^2}{\Delta x^2} \right] + 1 + 2\tau^2 \frac{\Delta t}{\Delta x}. \end{aligned}$$

For $\theta \in [\frac{1}{2}, 1]$, one has $E_a^{KdV} \leq G_a^{KdV}$ with G_a^{KdV} defined in (3.51a). Finally, we define $E_c := C_c^{KdV}$ and $E_e := C_e^{KdV} + (1 - \theta) \Delta t$ and $E_f := C_g + (1 - \theta) \frac{\Delta t \Delta x^2}{4}$.

□

Corollary 3.40 (Corollary of Proposition 3.25). Consider Scheme (3.2a)-(3.2b). Let $(e_j^n)_{(j,n)}$ be the convergence error defined by (3.3). Then, for every $n \in \llbracket 0, N \rrbracket$, $\gamma \in [0, \frac{1}{2}]$ and $\theta < \frac{1}{2}$, one has, if $\Delta t(1 - 2\theta) \leq \frac{\Delta x^3}{4}$,

$$\begin{aligned} \|\mathcal{A}e^{n+1}\|_{\ell_\Delta^2}^2 &\leq \|\mathcal{A}e^n\|_{\ell_\Delta^2}^2 \left[1 + G_a^{KdV} \Delta t \right] + \Delta t F_b \|D_+(e)^n\|_{\ell_\Delta^2}^2 + \Delta t^2 F_c \|D_c(e)^n\|_{\ell_\Delta^2}^2 + \Delta t F_d \|D_+ D_-(e)^n\|_{\ell_\Delta^2}^2 \\ &\quad + \Delta t F_e \|D_+ D_c(e)^n\|_{\ell_\Delta^2}^2 + \Delta t \|\epsilon^n\|_{\ell_\Delta^2}^2 \left\{ 1 + 4 \frac{\Delta t}{\Delta x} + \Delta t \right\}, \end{aligned}$$

with

$$\begin{aligned} G_a^{KdV} &:= \|u_\Delta^n\|_{\ell^\infty}^2 \left(1 + \frac{\Delta t}{\Delta x} \right) + \|D_+(u_\Delta)^n\|_{\ell^\infty} \left(7 + \frac{\Delta t}{\Delta x} \left[2\tau + \frac{2}{3} \|e^n\|_{\ell^\infty} + \frac{3}{2} \|(u_\Delta)^n\|_{\ell^\infty} \right] \right) \\ &\quad + \|D_+(u_\Delta)^n\|_{\ell^\infty}^2 \left[\sqrt{2} \frac{\sqrt{\Delta t}}{\sqrt{\Delta x}} + \frac{\Delta t^2}{\Delta x^2} \right] + 1 + 2\tau^2 \frac{\Delta t}{\Delta x}, \end{aligned} \quad (3.55a)$$

$$F_b := \left(\frac{\Delta x}{6} D_+(e)_j^n - \tau \right) (\Delta x - \tau \Delta t) + (1 - \theta) \Delta t \|D_+(u_\Delta)^n\|_{\ell^\infty}, \quad (3.55b)$$

$$F_c := \|(u_\Delta)^n\|_{\ell^\infty}^2 + \left\{ \|e^n\|_{\ell^\infty}^2 [1 + \Delta x] + 2 \|e^n\|_{\ell^\infty} \|(u_\Delta)^n\|_{\ell^\infty} + \frac{2\tau}{3} \|e^n\|_{\ell^\infty} \right\} - \tau^2, \quad (3.55c)$$

$$F_d := \frac{4}{\Delta x^2} \left\{ \Delta t \left[(1 - 2\theta) + \frac{(1 - \theta) \Delta x^2}{2} \left[\tau + \frac{\Delta x^{\frac{1}{2}-\gamma} + \|e^n\|_{\ell^\infty} + 9 \|e^n\|_{\ell^\infty}^2 \Delta x^{\gamma-\frac{1}{2}}}{2} \right] \right] \right\} \quad (3.55d)$$

$$\begin{aligned} &+ \Delta t(1 - \theta) \|D_+(u_\Delta)^n\|_{\ell^\infty} + \frac{(1 - \theta) \Delta x^2}{4} \left\{ \|D_+(u_\Delta)^n\|_{\ell^\infty} + \frac{\Delta x}{2} \|D_-(u_\Delta)^n\|_{\ell^\infty} \right\} \left[- \frac{\Delta x^3}{4} \right\}, \\ F_e &:= 2(1 - \theta) \Delta t \left\{ \|(u_\Delta)^n\|_{\ell^\infty} + \|e^n\|_{\ell^\infty} + \left[\frac{\Delta x^{\frac{1}{2}-\gamma} + \|e^n\|_{\ell^\infty} + 9 \|e^n\|_{\ell^\infty}^2 \Delta x^{\gamma-\frac{1}{2}}}{2} \right] \right\} \\ &- \frac{4\Delta t}{\Delta x^2} \left\{ (1 - 2\theta) + \frac{(1 - \theta) \Delta x^2}{2} \left[\tau + \frac{\Delta x^{\frac{1}{2}-\gamma} + \|e^n\|_{\ell^\infty} + 9 \|e^n\|_{\ell^\infty}^2 \Delta x^{\gamma-\frac{1}{2}}}{2} \right] \right\} \\ &\quad + \Delta t(1 - \theta) \|D_+(u_\Delta)^n\|_{\ell^\infty}. \end{aligned} \quad (3.55e)$$

Remark 3.41. The variables G_a^{KdV} are identical in both previous corollaries. It is noticed that Corollary 3.40 is valid for all θ but thereafter, it will be mainly used for $\theta < \frac{1}{2}$.

Proof. We choose $\sigma = 1$ in Inequality (3.37).

- From Relation (3.26), we transform the C_f^{KdV} -term in Inequality (3.37) into

$$\sum_{j \in \mathbb{Z}} \Delta t \Delta x C_f^{KdV} \left(D_+ D_+ D_-(e)_j^n \right)^2 = \sum_{j \in \mathbb{Z}} \Delta t \Delta x C_f^{KdV} \left[\frac{4}{\Delta x^2} \left(D_+ D_-(e)_j^n \right)^2 - \frac{4}{\Delta x^2} \left(D_+ D_c(e)_j^n \right)^2 \right].$$

- We upper bound $\|e^n - (1 - \theta) \Delta t D_+ D_+ D_-(e)^n\|_{\ell_\Delta^2}^2$ by $\|e^n + \theta \Delta t D_+ D_+ D_-(e)^n\|_{\ell_\Delta^2}^2$ thanks to Proposition 3.37, to obtain, instead of Inequality (3.37),

$$\begin{aligned} \|e^{n+1} + \theta \Delta t D_+ D_+ D_-(e)^n\|_{\ell_\Delta^2}^2 &\leq \|e^n + \Delta t \theta D_+ D_+ D_-(e)^n\|_{\ell_\Delta^2}^2 \left[1 + C_a^{KdV} \Delta t + \Delta t \right] + \Delta t C_b^{KdV} \|D_+(e)^n\|_{\ell_\Delta^2}^2 \\ &\quad + \Delta t^2 C_c^{KdV} \|D_c(e)^n\|_{\ell_\Delta^2}^2 + \Delta t \left\{ C_d^{KdV} + \frac{4C_f^{KdV}}{\Delta x^2} \right\} \|D_+ D_-(e)^n\|_{\ell_\Delta^2}^2 \\ &\quad + \Delta t \left\{ C_e^{KdV} - \frac{4C_f^{KdV}}{\Delta x^2} \right\} \|D_+ D_c(e)^n\|_{\ell_\Delta^2}^2 + \Delta t \|\epsilon^n\|_{\ell_\Delta^2}^2 \left\{ 1 + 4 \frac{\Delta t}{\Delta x} + \Delta t \right\}. \end{aligned}$$

Section 3.5. Rate of convergence

We note $F_a := C_a^{KdV} + 1$ and verify $F_a \leq G_a^{KdV}$. Finally, we fix $F_b := C_b^{KdV}$ with $\sigma = 1$, $F_c := C_c^{KdV}$, $F_d := C_d^{KdV} + \frac{4C_f^{KdV}}{\Delta x^2}$ with $\sigma = 1$ and $F_e := C_e^{KdV} - \frac{4C_f^{KdV}}{\Delta x^2}$.

□

In the following, we will have to show that E_i and F_i are non-positive to loop the estimates.

3.5.3 Induction method

Before proving our main theorem 3.7, we state a fruitful proposition, which concludes the stability stage and gives the CFL conditions to obtain the non-positivity of extra terms E_i and F_i .

Proposition 3.42. *Let $\alpha_0 > 0$ and $\beta_0 \in (0, 1)$ such that*

$$\|(u_\Delta)^n\|_{\ell^\infty} + \alpha_0 \leq \tau, \quad n \in \llbracket 0, N \rrbracket, \quad (3.56)$$

there exists $\omega_0 > 0$ such that

$$\Delta t \leq \omega_0, \quad \Delta x \leq \omega_0,$$

and such that the following CFL conditions be true

$$\begin{cases} 4(1 - 2\theta) \frac{\Delta t}{\Delta x^3} \leq 1 - \beta_0, \\ \left[\tau + \frac{1}{2} \right] \frac{\Delta t}{\Delta x} \leq 1 - \beta_0. \end{cases} \quad (3.57a)$$

$$\begin{cases} \left[\tau + \frac{1}{2} \right] \frac{\Delta t}{\Delta x} \leq 1 - \beta_0. \end{cases} \quad (3.57b)$$

If there exists $\gamma \in [0, \frac{1}{2})$ such that, for every $n \in \llbracket 0, N \rrbracket$, $\|e^n\|_{\ell^\infty} \leq \Delta x^{\frac{1}{2}-\gamma}$, then, for all $\theta \in [0, 1]$, we have

$$\|\mathcal{A}e^{n+1}\|_{\ell_\Delta^2}^2 \leq \|\mathcal{A}e^n\|_{\ell_\Delta^2}^2 \left\{ 1 + \Delta t G_a^{KdV} \right\} + \Delta t \|e^n\|_{\ell_\Delta^2}^2 \left\{ 1 + 4 \frac{\Delta t}{\Delta x} + \Delta t \right\}, \quad (3.58)$$

with G_a^{KdV} defined by (3.51a).

Remark 3.43. *The upper bound ω_0 must be chosen for the following inequalities to be verified*

$$\omega_0^{\frac{1}{2}-\gamma} \leq 3\tau,$$

- for $\theta \geq \frac{1}{2}$,

$$\bullet \quad \omega_0^{\frac{1}{4}-\frac{\gamma}{2}} \sqrt{\left[\omega_0^{\frac{1}{2}-\gamma} + \omega_0^{\frac{3}{2}-\gamma} \right] + 2\|(u_\Delta)^n\|_{\ell^\infty} + \frac{2\tau}{3}} \leq \alpha_0, \quad (3.59)$$

$$\bullet \quad \frac{13(1 - \beta_0)}{2\tau + 1} \omega_0^{\frac{1}{2}-\gamma} \leq \beta_0, \quad (3.60)$$

$$\bullet \quad (1 - 2\theta) + \frac{(1 - \theta)\omega_0^2}{2} \left[\tau + \frac{1}{2} + \frac{11}{2} \omega_0^{\frac{1}{2}-\gamma} \right] \leq 0, \quad \text{if } \theta > \frac{1}{2}, \quad (3.61)$$

$$\frac{11(1 - \beta_0)}{2\tau + 1} \omega_0^{\frac{1}{2}-\gamma} \leq \beta_0, \quad \text{if } \theta = \frac{1}{2}, \quad (3.62)$$

- for $\theta < \frac{1}{2}$,

$$\bullet \quad \omega_0^{\frac{1}{4}-\gamma} \sqrt{\left[\omega_0^{\frac{1}{2}-\gamma} + \omega_0^{\frac{3}{2}-\gamma} \right]} + 2\|(u_\Delta)^n\|_{\ell^\infty} + \frac{2\tau}{3} \leq \alpha_0, \quad (3.63)$$

$$\bullet \quad 12\omega_0^{\frac{1}{2}-\gamma} \leq \alpha_0, \quad (3.64)$$

$$\bullet \quad \frac{(1-\theta)(1-\beta_0)}{2(1-2\theta)\tau} \|(u_\Delta)^n\|_{\ell^\infty} \omega_0 + \frac{(1-\beta_0)}{3\tau + \frac{3}{2}} \omega_0^{\frac{1}{2}-\gamma} + \frac{\omega_0^{\frac{1}{2}-\gamma}}{3\tau} \leq \beta_0, \quad (3.65)$$

$$\bullet \quad \frac{(1-\theta)(1-\beta_0)}{2(1-2\theta)} \omega_0^2 \left[\tau + \frac{11}{2} \omega_0^{\frac{1}{2}-\gamma} \right] + (1-\theta) \|(u_\Delta)^n\|_{\ell^\infty} \frac{(1-\beta_0)}{(1-2\theta)} \left[\frac{(1-\beta_0)}{2(1-2\theta)} \omega_0^2 + \frac{\omega_0(2+\omega_0)}{4} \right] \leq \beta_0. \quad (3.66)$$

However, the previous required conditions are certainly not optimal.

Proof. We dissociate two cases according to the value of θ .

CASE $\theta \geq \frac{1}{2}$:

We show the non-positivity of coefficients E_i in Corollary 3.38, for $i \in \{b, c, e, f\}$.

- **Sign of E_b** : We get by developing $D_+(e)_j^n$

$$\frac{\Delta x}{6} D_+(e)_j^n \leq \frac{\|e^n\|_{\ell^\infty}}{3}.$$

The hypothesis on $\|e^n\|_{\ell^\infty}$ of Proposition 3.42 and the choice of a small $\Delta x \leq \omega_0$ (cf. Remark 3.43) give

$$\frac{\|e^n\|_{\ell^\infty}}{3} \leq \frac{\Delta x^{\frac{1}{2}-\gamma}}{3} \leq \frac{\omega_0^{\frac{1}{2}-\gamma}}{3} \leq \tau,$$

and due to the CFL condition (3.57b), one has

$$\Delta x - \tau \Delta t \geq 0.$$

Thus, $E_b \leq 0$.

- **Sign of E_c** : For the term E_c , thanks to the hypothesis $\|e^n\|_{\ell^\infty} \leq \Delta x^{\frac{1}{2}-\gamma}$, we obtain

$$E_c \leq \|(u_\Delta)^n\|_{\ell^\infty}^2 + \left\{ \left[\Delta x^{1-2\gamma} + \Delta x^{2-2\gamma} \right] + 2\Delta x^{\frac{1}{2}-\gamma} \|(u_\Delta)^n\|_{\ell^\infty} + \frac{2\tau \Delta x^{\frac{1}{2}-\gamma}}{3} \right\} - \tau^2.$$

As $\tau \geq \alpha_0 + \|(u_\Delta)^n\|_{\ell^\infty}$, if $\Delta x \leq \omega_0$ is sufficiently small (cf Remark 3.43, Inequality (3.59)), one has

$$E_c \leq \|(u_\Delta)^n\|_{\ell^\infty}^2 + \left\{ \left[\omega_0^{1-2\gamma} + \omega_0^{2-2\gamma} \right] + 2\omega_0^{\frac{1}{2}-\gamma} \|(u_\Delta)^n\|_{\ell^\infty} + \frac{2\tau \omega_0^{\frac{1}{2}-\gamma}}{3} \right\} - \tau^2 \leq 0.$$

- **Sign of E_e** : since we suppose $\|e^n\|_{\ell^\infty} \leq \Delta x^{\frac{1}{2}-\gamma}$, the term E_e satisfies

$$E_e \leq 2(1-\theta)\Delta t \left\{ \|(u_\Delta)^n\|_{\ell^\infty} + \frac{1}{2} + \frac{13}{2} \Delta x^{\frac{1}{2}-\gamma} \right\} - \Delta x.$$

Section 3.5. Rate of convergence

As $\theta \geq \frac{1}{2}$, $2(1 - \theta) \leq 1$, and, thanks to the choice of τ (3.56), one has

$$E_e \leq \Delta t \left\{ \tau + \frac{1}{2} + \frac{13}{2} \Delta x^{\frac{1}{2}-\gamma} \right\} - \Delta x = \Delta x \left\{ \frac{\Delta t}{\Delta x} \left[\tau + \frac{1}{2} \right] - 1 + \frac{13}{2} \frac{\Delta t}{\Delta x} \Delta x^{\frac{1}{2}-\gamma} \right\}.$$

If $\Delta x \leq \omega_0$ is small enough and using hyperbolic CFL (3.57b), one has

$$\frac{13}{2} \frac{\Delta t}{\Delta x} \Delta x^{\frac{1}{2}-\gamma} \leq \frac{13}{2} \frac{(1 - \beta_0)}{\tau + \frac{1}{2}} \Delta x^{\frac{1}{2}-\gamma} \leq \frac{13(1 - \beta_0)}{2\tau + 1} \omega_0^{\frac{1}{2}-\gamma}$$

which is less than β_0 thanks to Inequality (3.60). Thus one has

$$E_e \leq 0.$$

- **Sign of E_f** : the dispersive CFL-type condition (3.57a) together with hypothesis $\|e^n\|_{\ell^\infty} \leq \Delta x^{\frac{1}{2}-\gamma}$ give

$$E_f \leq \Delta t \left\{ (1 - 2\theta) + \frac{(1 - \theta)\Delta x^2}{2} \left[\tau + \frac{1}{2} + \frac{11}{2} \Delta x^{\frac{1}{2}-\gamma} \right] \right\} - \frac{\Delta x^3}{4},$$

which is non-positive if $\Delta x \leq \omega_0$ is sufficiently small. Indeed :

- if $\theta > \frac{1}{2}$, one has chosen ω_0 such that

$$(1 - 2\theta) + \frac{(1 - \theta)}{2} \Delta x^2 \left[\tau + \frac{1}{2} + \frac{11}{2} \Delta x^{\frac{1}{2}-\gamma} \right] \leq (1 - 2\theta) + \frac{(1 - \theta)}{2} \omega_0^2 \left[\tau + \frac{1}{2} + \frac{11}{2} \omega_0^{\frac{1}{2}-\gamma} \right] \leq 0,$$

thanks to Inequality (3.61),

- if $\theta = \frac{1}{2}$,

$$E_f \leq \frac{\Delta t \Delta x^2}{4} \left[\tau + \frac{1}{2} + \frac{11}{2} \Delta x^{\frac{1}{2}-\gamma} \right] - \frac{\Delta x^3}{4} = \frac{\Delta x^3}{4} \left\{ \frac{\Delta t}{\Delta x} \left[\tau + \frac{1}{2} \right] - 1 + \frac{11\Delta t}{2\Delta x} \Delta x^{\frac{1}{2}-\gamma} \right\},$$

and Condition (3.57b) together with $\Delta x \leq \omega_0$ for ω_0 verifying Inequality (3.62) enable us to conclude about the non-positivity of E_f .

Thus Estimate (3.58) is proved for $\theta \geq \frac{1}{2}$.

CASE $\theta < \frac{1}{2}$:

In the same way, from Corollary 3.40, we show the non-positivity of F_i , for $i \in \{b, c, d, e\}$.

- **Sign of F_b** : one has, by definition of F_b and by hypothesis $\|e^n\|_{\ell^\infty} \leq \Delta x^{\frac{1}{2}-\gamma}$

$$\begin{aligned} F_b &\leq \left(\frac{\Delta x}{6} D_+(e)_j^n - \tau \right) (\Delta x - \tau \Delta t) + 2(1 - \theta) \frac{\Delta t}{\Delta x} \|(u_\Delta)^n\|_{\ell^\infty} \\ &\leq \frac{\Delta x \|e^n\|_{\ell^\infty}}{3} + \frac{\tau \Delta t \|e^n\|_{\ell^\infty}}{3} - \tau \Delta x + \tau^2 \Delta t + 2(1 - \theta) \frac{\Delta t}{\Delta x} \|(u_\Delta)^n\|_{\ell^\infty} \\ &\leq \tau \left[\tau \Delta t \left(1 + \frac{\Delta x^{\frac{1}{2}-\gamma}}{3\tau} \right) - \Delta x \left(1 - \frac{\Delta x^{\frac{1}{2}-\gamma}}{3\tau} - 2(1 - \theta) \frac{\Delta t}{\Delta x^2 \tau} \|(u_\Delta)^n\|_{\ell^\infty} \right) \right] \\ &\leq \tau \Delta x \left[\tau \frac{\Delta t}{\Delta x} + \frac{\Delta t}{\Delta x} \frac{\Delta x^{\frac{1}{2}-\gamma}}{3} - 1 + \frac{\Delta x^{\frac{1}{2}-\gamma}}{3\tau} + 2(1 - \theta) \frac{\Delta t}{\Delta x^2 \tau} \|(u_\Delta)^n\|_{\ell^\infty} \right]. \end{aligned}$$

The hyperbolic CFL condition (3.57b) and the dispersive one (3.57a) (we recall that $1 - 2\theta > 0$ in that case) imply

$$F_b \leq \tau \Delta x \left[1 - \beta_0 + \frac{(1 - \beta_0)\Delta x^{\frac{1}{2}-\gamma}}{3\tau + \frac{3}{2}} - 1 + \frac{\Delta x^{\frac{1}{2}-\gamma}}{3\tau} + (1 - \theta) \frac{\Delta x(1 - \beta_0)}{2\tau(1 - 2\theta)} \|(u_\Delta)^n\|_{\ell^\infty} \right].$$

The choice of ω_0 small enough to satisfy Inequalities (3.65) (cf Remark 3.43) implies $F_b \leq 0$.

- **Sign of F_c** : since $F_c = E_c$, we follow exactly the same proof as for $\theta \geq \frac{1}{2}$ to show $F_c \leq 0$.
- **Sign of F_d** : thanks to Relation (3.55d), one has

$$F_d = \frac{4}{\Delta x^2} \left\{ \Delta t \left[(1-2\theta) + \frac{(1-\theta)\Delta x^2}{2} \left[\tau + \frac{\Delta x^{\frac{1}{2}-\gamma} + \|e^n\|_{\ell^\infty} + 9\|e^n\|_{\ell^\infty}^2 \Delta x^{\gamma-\frac{1}{2}}}{2} \right] \right. \right. \\ \left. \left. + \Delta t(1-\theta)\|D_+(u_\Delta)^n\|_{\ell^\infty} + \frac{(1-\theta)\Delta x^2}{4} \left\{ \|D_+(u_\Delta)^n\|_{\ell^\infty} + \frac{\Delta x}{2}\|D_-(u_\Delta)^n\|_{\ell^\infty} \right\} \right] - \frac{\Delta x^3}{4} \right\}.$$

Since $\|e^n\|_{\ell^\infty} \leq \Delta x^{\frac{1}{2}-\gamma}$, it becomes, thanks to dispersive CFL (3.57a),

$$F_d = \Delta x \left\{ \frac{4\Delta t}{\Delta x^3}(1-2\theta) + \frac{2\Delta t}{\Delta x}(1-\theta) \left[\tau + \frac{11\Delta x^{\frac{1}{2}-\gamma}}{2} \right] \right. \\ \left. + 8\frac{\Delta t^2}{\Delta x^4}(1-\theta)\|u_\Delta^n\|_{\ell^\infty} + 2(1-\theta)\frac{\Delta t}{\Delta x^2}\|u_\Delta^n\|_{\ell^\infty} + (1-\theta)\frac{\Delta t}{\Delta x}\|u_\Delta^n\|_{\ell^\infty} - 1 \right\} \\ \leq \Delta x \left\{ \frac{4\Delta t}{\Delta x^3}(1-2\theta) + \frac{\Delta x^2(1-\beta_0)}{2(1-2\theta)}(1-\theta) \left[\tau + \frac{11\Delta x^{\frac{1}{2}-\gamma}}{2} \right] + \frac{(1-\beta_0)^2\Delta x^2}{2(1-2\theta)^2}(1-\theta)\|u_\Delta^n\|_{\ell^\infty} \right. \\ \left. + (1-\theta)\frac{(1-\beta_0)\Delta x}{2(1-2\theta)}\|u_\Delta^n\|_{\ell^\infty} + (1-\theta)\frac{\Delta x^2(1-\beta_0)}{4(1-2\theta)}\|u_\Delta^n\|_{\ell^\infty} - 1 \right\} \\ = \Delta x \left\{ \frac{4\Delta t}{\Delta x^3}(1-2\theta) + \frac{\Delta x^2(1-\beta_0)}{2(1-2\theta)}(1-\theta) \left[\tau + \frac{11\Delta x^{\frac{1}{2}-\gamma}}{2} \right] \right. \\ \left. + (1-\theta)\|u_\Delta^n\|_{\ell^\infty} \frac{(1-\beta_0)}{(1-2\theta)} \left[\frac{(1-\beta_0)}{2(1-2\theta)}\Delta x^2 + \frac{\Delta x(2+\Delta x)}{4} \right] - 1 \right\}$$

For $\Delta x \leq \omega_0$ small enough such that (3.66) be verified and thanks to the CFL condition (3.57a), one has

$$F_d \leq 0.$$

- **Sign of F_e** : we develop F_e to obtain

$$F_e \leq 2(1-\theta)\Delta t \left\{ \|(u_\Delta)^n\|_{\ell^\infty} + \frac{13}{2}\Delta x^{\frac{1}{2}-\gamma} \right\} - \frac{4\Delta t}{\Delta x^2}(1-2\theta) - 2(1-\theta)\Delta t \left[\tau - \frac{11\Delta x^{\frac{1}{2}-\gamma}}{2} \right] \\ - \frac{8\Delta t^2}{\Delta x^3}(1-\theta)\|(u_\Delta)^n\|_{\ell^\infty} \\ \leq 2(1-\theta)\Delta t \left\{ \|(u_\Delta)^n\|_{\ell^\infty} + 12\Delta x^{\frac{1}{2}-\gamma} - \tau \right\} - \frac{4\Delta t}{\Delta x^2} \left[(1-2\theta) + \frac{2\Delta t}{\Delta x}(1-\theta)\|(u_\Delta)^n\|_{\ell^\infty} \right].$$

Since $\theta < \frac{1}{2}$, one has $1-2\theta > 0$ then $-\frac{4\Delta t}{\Delta x^2} \left[(1-2\theta) + \frac{2\Delta t}{\Delta x}(1-\theta)\|(u_\Delta)^n\|_{\ell^\infty} \right] \leq 0$. The choice of a small $\Delta x \leq \omega_0$ satisfying (3.64) and the choice of τ (3.56) give $F_e \leq 0$.

□

We are now able to demonstrate by induction and by Grönwall Lemma 2.29 our main result for a regular initial datum : Theorem 3.7.

Remark 3.44. The induction method is used to control the norm $\|e^n\|_{\ell^\infty}$ in the factor coefficient G_a^{KdV} of Proposition 3.42. However, we can not perform the induction directly on ℓ_Δ^2 -norm, we have to take into account the contribution of Airy term via the dispersive operator $\mathcal{A}$. Thus we have to perform the induction method on $\mathcal{A}e^n$.

Section 3.5. Rate of convergence

Proof of Theorem 3.7. Let $\gamma \in [0, \frac{1}{2})$, we choose the Rusanov coefficient τ such that (3.7) is true, for every $n \in \mathbb{N}$. This choice is possible because of Theorem 3.15 which proves that the exact solution belongs to L_x^∞ for $t \in [0, T]$. We suppose the two CFL conditions (3.8) and eventually, we choose Δx small enough as in Proposition 3.42, and also small enough to satisfy

$$\Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}} \left(1 + \|u_0\|_{H^{\frac{1}{2}+\eta}}^2\right) \left(\frac{\|u_0\|_{H^6}}{\tau + \frac{1}{2}} + \|u_0\|_{H^4} + \|u_0\|_{H^{\frac{3}{2}+\eta}} \|u_0\|_{H^1}\right) \leq \frac{1}{\Delta x^\gamma}, \quad (3.67)$$

with $\Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}}$ defined in (3.10).

Let us prove by induction on $n \in \llbracket 0, N \rrbracket$ the inequality

$$\|e^n\|_{\ell^\infty} \leq \Delta x^{\frac{1}{2}-\gamma}. \quad (3.68)$$

Initialisation. For $n = 0$, Equation (3.68) is true because Expressions (3.2b) and (3.5) imply

$$e_j^0 = 0, \quad j \in \mathbb{Z}.$$

Heredity. Assume Estimate (3.68) is true for $n \geq 0$, let us prove this estimate for $n + 1$.

The induction hypothesis (3.68) at step n enables to control the term that seemed problematic in G_a^{KdV} of Proposition 3.42 : the norm $\|e^n\|_{\ell^\infty}$.

Once controlling the ℓ^∞ -norm, we use the first inequality of (3.16) to upper bound $\Delta t \|D_+ (u_\Delta)^n\|_{\ell^\infty}$ and $\Delta t \|D_+ (u_\Delta)^n\|_{\ell^\infty}^2$, Proposition 3.42 thus becomes (by splitting G_a^{KdV})

$$\|\mathcal{A}e^{n+1}\|_{\ell_\Delta^2}^2 \leq \|\mathcal{A}e^n\|_{\ell_\Delta^2}^2 \left\{1 + \Delta t G_b^n + \sum_{i=1}^2 \left(\int_{t^n}^{t^{n+1}} \|\partial_x u(t, \cdot)\|_{L_x^\infty}^i dt \right) G_{c,i}^n \right\} + \Delta t \|\epsilon^n\|_{\ell_\Delta^2}^2 \left\{1 + 4 \frac{\Delta t}{\Delta x} + \Delta t\right\},$$

with

$$G_b^n = \left[\|u_\Delta^n\|_{\ell^\infty}^2 \left(1 + \frac{\Delta t}{\Delta x}\right) + 1 + 2\tau^2 \frac{\Delta t}{\Delta x} \right] \leq \left[1 + \|u_\Delta\|_{\ell_n^\infty}^2 (1 + \frac{\Delta t}{\Delta x}) + 2 \frac{\Delta t}{\Delta x} \tau^2 \right]$$

and

$$G_{c,1}^n = \left[7 + \frac{\Delta t}{\Delta x} \left(2\tau + \frac{2}{3} \Delta x^{\frac{1}{2}-\gamma} + \frac{3}{2} \|(u_\Delta)^n\|_{\ell^\infty} \right) \right] \leq \left[7 + \frac{\Delta t}{\Delta x} \left(2\tau + \frac{2}{3} \Delta x^{\frac{1}{2}-\gamma} + \frac{3}{2} \|u_\Delta\|_{\ell_n^\infty} \right) \right]$$

and

$$G_{c,2}^n = \left[\sqrt{2} \frac{\sqrt{\Delta t}}{\sqrt{\Delta x}} + \frac{\Delta t^2}{\Delta x^2} \right].$$

Due to the CFL condition, we have, denoting by C a number independent of τ , u_Δ^n , Δt and Δx

$$G_b^n \leq C \left(1 + \tau^2 \left(1 + \frac{\Delta t}{\Delta x} \right) \right) =: G_b, \quad (3.69)$$

$$G_{c,1}^n \leq C \left(1 + \frac{\Delta t}{\Delta x} [1 + \tau] \right) =: G_{c,1}, \quad (3.70)$$

and

$$G_{c,2}^n = \left[\sqrt{2} \frac{\sqrt{\Delta t}}{\sqrt{\Delta x}} + \frac{\Delta t^2}{\Delta x^2} \right] =: G_{c,2}. \quad (3.71)$$

We can now apply the discrete Grönwall Lemma 2.29 (noticing that $e_j^0 = 0$, $j \in \mathbb{Z}$). It provides, for every $n \in \llbracket 0, N-1 \rrbracket$,

$$\|\mathcal{A}e^{n+1}\|_{\ell_\Delta^2}^2 \leq \exp \left(t^{n+1} G_b + \sum_{i=1}^2 \int_0^{t^{n+1}} \|\partial_x u(s, \cdot)\|_{L_x^\infty(\mathbb{R})}^i ds G_{c,i} \right) \sup_{n \in \llbracket 0, N \rrbracket} \|\epsilon^n\|_{\ell_\Delta^2}^2 T \left\{ 1 + 4 \frac{\Delta t}{\Delta x} + \Delta t \right\}. \quad (3.72)$$

Finally, Theorem 3.15 and Proposition 3.18 give, for $0 < \eta \leq 6 - \frac{3}{2}$,

$$\|\mathcal{A}e^{n+1}\|_{\ell_\Delta^2}^2 \leq \Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}}^2 \left(1 + \|u_0\|_{H^{\frac{1}{2}+\eta}}^2\right)^2 \left\{ \Delta t^2 \|u_0\|_{H^6}^2 + \Delta x^2 \left[\|u_0\|_{H^4}^2 + \|u_0\|_{H^{\frac{3}{2}+\eta}}^2 \|u_0\|_{H^1}^2 \right] \right\},$$

with

$$\begin{aligned} \Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}}^2 &= \exp \left(TG_b + \|u_0\|_{H^{\frac{3}{4}}} C e^{\kappa_{\frac{3}{4}} T} \left[G_{c,1} T^{\frac{3}{4}} + G_{c,2} T^{\frac{1}{2}} \right] \right) C^2 e^{2\kappa T} T \left\{ 1 + 4 \frac{\Delta t}{\Delta x} + \Delta t \right\} \\ &\leq \exp \left(C (1 + \tau^2) \left(1 + \frac{\Delta t^2}{\Delta x^2} \right) \left(T + (T^{\frac{3}{4}} + T^{\frac{1}{2}}) \|u_0\|_{H^{\frac{3}{4}}} e^{\kappa_{\frac{3}{4}} T} \right) \right) C^2 e^{2\kappa T} T \left\{ 1 + \frac{\Delta t}{\Delta x} \right\}, \end{aligned}$$

with C independent of u_0 and κ , $\kappa_{\frac{3}{4}}$ dependent only on $\|u_0\|_{L^2}$.

Since $\|e^{n+1}\|_{\ell_\Delta^2}^2 \leq \|\mathcal{A}e^{n+1}\|_{\ell_\Delta^2}^2$ (Proposition 3.33), the previous inequality gives

$$\begin{aligned} \|e^{n+1}\|_{\ell_\Delta^2}^2 &\leq \Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}}^2 \left(1 + \|u_0\|_{H^{\frac{1}{2}+\eta}}^2\right)^2 \left\{ \Delta t^2 \|u_0\|_{H^6}^2 + \Delta x^2 \left[\|u_0\|_{H^4}^2 + \|u_0\|_{H^{\frac{3}{2}+\eta}}^2 \|u_0\|_{H^1}^2 \right] \right\} \\ &\leq \Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}}^2 \left(1 + \|u_0\|_{H^{\frac{1}{2}+\eta}}^2\right)^2 \left(\frac{\|u_0\|_{H^6}^2}{(\tau + \frac{1}{2})^2} + \|u_0\|_{H^4}^2 + \|u_0\|_{H^{\frac{3}{2}+\eta}}^2 \|u_0\|_{H^1}^2 \right) \Delta x^2, \end{aligned} \quad (3.73)$$

where the last inequality is obtained thanks to the CFL condition (3.57b).

Conclusion. It remains to verify the induction hypothesis (3.68) at step $n + 1$. The definition of the ℓ_Δ^2 -norm, Identity (1.14), together with the inclusion $\ell^2 \subset \ell^\infty$, holds

$$\|e^n\|_{\ell^\infty} \leq \frac{\|e^n\|_{\ell_\Delta^2}}{\sqrt{\Delta x}}.$$

According to the upper bound (3.73), the ℓ^∞ -norm is bounded as follow

$$\|e^{n+1}\|_{\ell^\infty} \leq \Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}} \left(1 + \|u_0\|_{H^{\frac{1}{2}+\eta}}^2\right) \left(\frac{\|u_0\|_{H^6}}{\tau + \frac{1}{2}} + \|u_0\|_{H^4} + \|u_0\|_{H^{\frac{3}{2}+\eta}} \|u_0\|_{H^1} \right) \sqrt{\Delta x}.$$

The choice of a small Δx (in order to verify (3.67)) implies thus $\|e^{n+1}\|_{\ell^\infty} \leq \Delta x^{\frac{1}{2}-\gamma}$. The induction hypothesis is then true for $n + 1$. $\square$

Thus, we have proved Equation (3.9) with (3.10).

Remark 3.45. The choice of a time average in the definition of u_Δ , Equation (3.5), is dictated by the discrete Grönwall Lemma 2.29 on (3.72). Indeed, applying discrete Grönwall Lemma introduces the following term $\sum_{n=0}^N \Delta t \|D_+(u_\Delta)^n\|_{\ell^\infty}^i$ which is controlled thanks to the estimate (3.16), where the time integral plays a crucial role.

Regarding the space average in the definition of u_Δ , its necessity comes from controlling the sum on $j \in \mathbb{Z}$ in the consistency estimates (A.2).

3.6 Convergence for less smooth initial data

In this section, we relax the hypothesis $u_0 \in H^6(\mathbb{R})$ and adapt the previous proof for any solution in $H^{\frac{3}{4}}(\mathbb{R})$ to obtain Theorem 3.10. When u_0 is not smooth enough to verify $u_0 \in H^6(\mathbb{R})$,

Section 3.6. Convergence for less smooth initial data

we regularize it thanks to mollifiers $(\varphi^\delta)_{\delta>0}$, as explained in general introduction. Recall that we denote the mollifiers by $(\varphi^\delta)_{\delta>0}$, whose construction is based on χ a C^∞ -function such that $\chi \equiv 1$ on $[-\frac{1}{2}, \frac{1}{2}]$, χ is supported in $[-1, 1]$ and $\chi(\xi) = \chi(-\xi)$. We denote the exact solution of (3.1a) from u_0 by u , the exact solution from $u_0 \star \varphi^\delta$ by u^δ and the numerical solution from (3.11) by $([v^\delta]_j^n)_{(n,j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}}$.

Hereafter, we will use Lemma 2.33 and 2.34 on the approximation results between u^δ and u , which are mentioned for the first time in Subsection 2.2.2.1. We need moreover, the following result about the speed of decrease for the Sobolev norms of u^δ .

Proposition 3.46. Assume $m \geq s \geq 0$ and $u_0 \in H^s(\mathbb{R})$ then,

$$\sup_{t \in [0, T]} \|u^\delta(t, \cdot)\|_{H^m(\mathbb{R})} \leq C e^{\kappa_m T} \frac{\|u_0\|_{H^s(\mathbb{R})}}{\delta^{m-s}},$$

where C is a number which depends on m and κ_m depends on $\|u_0\|_{L^2}$ and m . Both are independent of δ .

Proof. We combine Theorem 3.15 and Lemma 2.33. □

3.6.1 Proof of convergence Theorem 3.10

Let $\alpha_0 > 0$ and $\beta_0 \in (0, 1)$, let us assume

$$\|u_\Delta^n\|_{\ell^\infty} + \alpha_0 \leq \tau, \quad n \in \mathbb{N}.$$

There exists $\omega_0 > 0$ such that

$$\Delta t \leq \omega_0, \quad \Delta x \leq \omega_0,$$

and such that the following CFL conditions be true

$$\begin{cases} 4(1 - 2\theta) \frac{\Delta t}{\Delta x^3} \leq 1 - \beta_0, \\ \left[\tau + \frac{1}{2} \right] \frac{\Delta t}{\Delta x} \leq 1 - \beta_0. \end{cases}$$

We eventually assume that $u_0 \in H^s(\mathbb{R})$ with $s \geq \frac{3}{4}$.

By using a triangle inequality, the global error is upper bounded by

$$\|e^n\|_{\ell_\Delta^2} \leq \Xi_1^n + \Xi_2^n,$$

with

$$\begin{aligned} \Xi_1^n &= \left[\sum_{j \in \mathbb{Z}} \Delta x \left(\frac{1}{\Delta x [\inf(t^{n+1}, T) - t^n]} \int_{t^n}^{\inf(t^{n+1}, T)} \int_{x_j}^{x_{j+1}} u(s, x) - u^\delta(s, x) dx ds \right)^2 \right]^{\frac{1}{2}} \\ &= \left(\sum_{j \in \mathbb{Z}} \Delta x \left((u_\Delta)_j^n - [u_\Delta^\delta]_j^n \right)^2 \right)^{\frac{1}{2}}, \end{aligned}$$

with Notation (3.4), and

$$\begin{aligned} \Xi_2^n &= \left[\sum_{j \in \mathbb{Z}} \Delta x \left(\frac{1}{\Delta x [\inf(t^{n+1}, T) - t^n]} \int_{t^n}^{\inf(t^{n+1}, T)} \int_{x_j}^{x_{j+1}} u^\delta(s, x) dx ds - (v^\delta)_j^n \right)^2 \right]^{\frac{1}{2}} \\ &= \left(\sum_{j \in \mathbb{Z}} \Delta x \left([u_\Delta^\delta]_j^n - (v^\delta)_j^n \right)^2 \right)^{\frac{1}{2}}. \end{aligned}$$

Let us first focus on term Ξ_1^n . The Cauchy-Schwarz inequality implies $\Xi_1^n \leq \sup_{t \in [0, T]} \|u(t, \cdot) - u^\delta(t, \cdot)\|_{L^2(\mathbb{R})}$, which imposes to study the difference between u and u^δ . Since u and u^δ are two solutions of the initial equation (3.1a), one has

$$\partial_t (u - u^\delta) + \partial_x^3 (u - u^\delta) + u \partial_x (u - u^\delta) + (u - u^\delta) \partial_x u^\delta = 0.$$

Multiplying by $(u - u^\delta)$, integrating the equation and changing u^δ in $u - (u - u^\delta)$ in the latest term yield

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}} \frac{(u(t, x) - u^\delta(t, x))^2}{2} dx - \int_{\mathbb{R}} \partial_x u(t, x) \frac{(u(t, x) - u^\delta(t, x))^2}{2} dx \\ + \int_{\mathbb{R}} (u(t, x) - u^\delta(t, x))^2 \partial_x [u(t, x) - (u(t, x) - u^\delta(t, x))] dx = 0, \end{aligned}$$

thus

$$\frac{d}{dt} \frac{\|u(t, \cdot) - u^\delta(t, \cdot)\|_{L^2(\mathbb{R})}^2}{2} \leq \frac{\|\partial_x u(t, \cdot)\|_{L^\infty(\mathbb{R})}}{2} \|u(t, \cdot) - u^\delta(t, \cdot)\|_{L^2(\mathbb{R})}^2.$$

The previous inequality looks like the 'weak-strong uniqueness' of DiPerna [44] or Dafermos [38, 39]. The $L^2(\mathbb{R})$ -norm of the difference $u - u^\delta$ is then upper bounded by

$$\begin{aligned} \|u(t, \cdot) - u^\delta(t, \cdot)\|_{L^2(\mathbb{R})}^2 &\leq \exp \left(\int_0^t \|\partial_x u(s, \cdot)\|_{L^\infty(\mathbb{R})} ds \right) \|u_0 - u_0^\delta\|_{L^2(\mathbb{R})}^2 \\ &\leq \exp \left(T^{\frac{3}{4}} C_{\frac{3}{4}} e^{\frac{\kappa_3}{4} T} \|u_0\|_{H^{\frac{3}{4}}} \right) \|u_0 - u_0^\delta\|_{L^2(\mathbb{R})}^2, \end{aligned}$$

where $\kappa_{\frac{3}{4}}$ and $C_{\frac{3}{4}}$ are defined in Theorem 3.15. Then

$$\Xi_1^n \leq \sup_{t \in [0, T]} \|u(t, \cdot) - u^\delta(t, \cdot)\|_{L^2(\mathbb{R})} \leq \exp \left(\frac{T^{\frac{3}{4}} C_{\frac{3}{4}} e^{\frac{\kappa_3}{4} T}}{2} \|u_0\|_{H^{\frac{3}{4}}} \right) \|u_0 - u_0^\delta\|_{L^2(\mathbb{R})}.$$

Lemma 2.34 implies

$$\Xi_1^n \leq C \delta^s \|u_0\|_{H^s(\mathbb{R})} \exp \left(\frac{T^{\frac{3}{4}} C_{\frac{3}{4}} e^{\frac{\kappa_3}{4} T}}{2} \|u_0\|_{H^{\frac{3}{4}}} \right). \quad (3.74)$$

In the other hand, the term Ξ_2^n corresponds to Estimate (3.73) derived in Subsection 3.5.3 with a smooth initial datum. It remains us to quantify the dependency of it upper bound with respect to δ . Thanks to Theorem 3.7, one has

$$\Xi_2^n \leq \Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0^\delta\|_{H^{\frac{3}{4}}}} \left(1 + \|u_0^\delta\|_{H^{\frac{1}{2} + \eta}}^2 \right) \left(\frac{\|u_0^\delta\|_{H^6}}{\tau + \frac{1}{2}} + \|u_0^\delta\|_{H^4} + \|u_0^\delta\|_{H^{\frac{3}{2} + \eta}} \|u_0^\delta\|_{H^1} \right) \Delta x,$$

with $\Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0^\delta\|_{H^{\frac{3}{4}}}}$ defined by (3.10). As u_0 belongs to $H^s(\mathbb{R})$ with $s \geq \frac{3}{4}$, then $\|u_0^\delta\|_{H^{\frac{3}{4}}} = \|u_0\|_{H^{\frac{3}{4}}}$ and $\|u_0^\delta\|_{H^{\frac{1}{2} + \eta}} = \|u_0\|_{H^{\frac{1}{2} + \eta}}$.

Lemma 3.47. For every $s \geq \frac{3}{4}$, there exists C , depending only on s and on $\|u_0\|_{L^2}$, such that, if $u_0 \in H^s(\mathbb{R})$,

$$\frac{\|u_0^\delta\|_{H^6}}{\tau + \frac{1}{2}} + \|u_0^\delta\|_{H^4} + \|u_0^\delta\|_{H^{\frac{3}{2} + \eta}} \|u_0^\delta\|_{H^1} \leq \frac{\|u_0\|_{H^s}}{\delta^{6-s}} C \left(\frac{1}{\tau + \frac{1}{2}} + 1 + \|u_0\|_{H^{\min(1, s)}} \right).$$

Section 3.6. Convergence for less smooth initial data

Proof. We apply Lemma 2.33 with $s = 6, 4, \frac{3}{2} + \eta, 1$ and the biggest power of δ is $\frac{1}{\delta^{6-s}}$. $\square$

Thus, an upper bound for Ξ_2^n is

$$\Xi_2^n \leq \Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}} \left(1 + \|u_0\|_{H^{\frac{1}{2}+\eta}}^2\right) \left(\frac{1}{\tau + \frac{1}{2}} + 1 + \|u_0\|_{H^{\min(1,s)}}\right) C \frac{\|u_0\|_{H^s}}{\delta^{6-s}} \Delta x.$$

However, for Theorem 3.7 to be applied, we need to choose a small Δx such that condition (3.67) is true. With the above lemma, this condition is satisfied if

$$\Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}} \left(1 + \|u_0\|_{H^{\frac{1}{2}+\eta}}^2\right) \left(\frac{1}{\tau + \frac{1}{2}} + 1 + \|u_0\|_{H^{\min(1,s)}}\right) C \frac{\|u_0\|_{H^s}}{\delta^{6-s}} \leq \frac{1}{\Delta x^\gamma}, \quad (3.75)$$

what is of course possible to ensure (and is assumed in the following). Hence, the convergence error $(e_j^n)_{(n,j)}$ is upper bounded by

$$\begin{aligned} \|e^n\|_{\ell_\Delta^2} \leq C & \left[\Lambda_{T, \frac{\Delta t}{\Delta x}, \|u_0\|_{H^{\frac{3}{4}}}} \left(1 + \|u_0\|_{H^{\frac{1}{2}+\eta}}^2\right) \left(\frac{1}{\tau + \frac{1}{2}} + 1 + \|u_0\|_{H^{\min(1,s)}}\right) \right. \\ & \left. + \exp \left(\frac{T^{\frac{3}{4}} C_{\frac{3}{4}} e^{\frac{\kappa_{\frac{3}{4}} T}{2}} \|u_0\|_{H^{\frac{3}{4}}}}{2} \right) \right] \|u_0\|_{H^s} \left[\frac{\Delta x}{\delta^{6-s}} + \delta^s \right], \quad (3.76) \end{aligned}$$

for $n \in \llbracket 0, N \rrbracket$.

The key point is to find the optimal δ , in other words, the parameter δ which makes both terms δ^s , coming from Ξ_1^n and $\frac{\Delta x}{\delta^{6-s}}$ coming from Ξ_2^n in (3.76) equal while respecting the constraint (3.75). It is exactly the same guidelines of Sections 2.2.2 or 2.3.2, but we have an additional constraint here. Defining $\delta = \Delta x^a$ summarizes the problem in the following system

$$\begin{cases} \text{Find } a \text{ such that : } \Delta x^{as} = \frac{\Delta x}{\Delta x^{a(6-s)}}, \\ \text{under the constraint : } \frac{1}{\Delta x^{a(6-s)}} < \frac{1}{\Delta x^\gamma}. \end{cases}$$

Three cases have to be considered :

- if $\frac{3}{4} \leq s \leq 6 - 6\gamma$, the constraint is binding and we have to choose a which transforms the constraint inequality in an equality : $a = \frac{\gamma}{6-s}$. In that case, the rate of convergence is given by the smallest term between Δx^{as} and $\frac{\Delta x}{\Delta x^{a(6-s)}}$ i.e. $\Delta x^{\frac{\gamma s}{6-s}}$.
- If $6 - 6\gamma \leq s \leq 6$, $a = \frac{1}{6}$ enables both terms Δx^{as} and $\frac{\Delta x}{\Delta x^{a(6-s)}}$ to be equal without violating the constraint. This choice of a gives a rate of convergence of $\Delta x^{\frac{s}{6}}$.
- If $s \geq 6$, the result of Theorem 3.7 applies.

Since γ is in $[0, \frac{1}{2})$ (cf. Lemma 3.29, Proposition 3.42 and induction hypothesis (3.68)), we take the optimal γ : $\gamma = \frac{1}{2} - \eta$ with η small and $\eta > 0$. The conclusion of the theorem is a straightforward consequence.

Remark 3.48. The choice of δ is independent of the regularity s of the initial datum, if $3 \leq s \leq 6$.

Remark 3.49. Contrary to what happens for linear dispersive equations with less smooth initial data (Section 2.2.2 and Section 2.3.2 in Chapter 2), we have to use the numerical solution from u_0^δ and not from u_0 . Namely, the scheme in Theorem 3.10 is (3.2a)–(3.11) and not (3.2a)–(3.2b). We will explain in this remark that difference.

The majoration of Ξ_1^n comes from the stability of the continuous EDP, the majoration of Ξ_2^n comes from the convergence error in smooth case and the majoration of Ξ_3^n comes from the stability of the numerical scheme. However, one has proved, for KdV equation, the stability of the scheme provided $u \in L^\infty([0, T], H^{\frac{3}{4}}(\mathbb{R}))$. An upper bound for Ξ_3^n thus requires for the numerical solution $([v^\delta]_j^n)_{(n,j)}$ to be in the discrete equivalent of $L^\infty([0, T], H^{\frac{3}{4}}(\mathbb{R}))$. Since this discrete regularity is not verified, no upper bound of Ξ_3^n is available. To deal only with Ξ_1^n and Ξ_2^n , we have to initialize the numerical scheme with u_0^δ .

3.7 Numerical results

In this section, the previous results are illustrated numerically by some examples and the numerical convergence rates are computed for the KdV equation.

3.7.1 Convergence rates

Through the rest of the section, the computations are performed with an implicit scheme $\theta = 1$ in order to avoid the dispersive CFL condition. Our purpose is to gauge the relevance of our theoretical results on the rate of convergence with respect to Δx . To this end, the time step is chosen according to the hyperbolic CFL condition. More precisely, τ is numerically chosen such that $\tau^n = \max_{k \in \llbracket 0, n-1 \rrbracket, j \in \llbracket 0, J \rrbracket} |v_j^k|$ and $\Delta t^n = \frac{\Delta x}{\tau^n}$. This choice seems surprising related to the CFL

of Theorems 3.7 and 3.10 but, as explain in Remark 3.8, the condition $[\tau + \frac{1}{2}]\Delta t < \Delta x$ seems technical and may be replaced by the classical one $\tau\Delta t \leq \Delta x$. Eventually, we fix the final time $T = 0.1$.

We can not simulate numerical solutions on $\mathbb{Z}$ as done in the theoretical results. We have to take into account numerical boundaries : we use periodic initial data. We fix the space domain to $[0, L]$ with $L = 50$ (except for cnoidal wave where $L = 1$) and fix $J = \lfloor \frac{L}{\Delta x} \rfloor$.

Remark 3.50. Notice that the theoretical results do not apply rigorously since the solutions do not belong to $H^s(\mathbb{R})$ because of their periodicity.

The convergence rate r_J is computed with

$$r_J = \frac{\log(E_J) - \log(E_{2J})}{\log(2)}.$$

When the exact solution is known (e.g. for the cnoidal-wave solution), the variable E_J corresponds to the following difference

$$E_J = \sup_{n \in \llbracket 0, N \rrbracket} \left\| (e_j^n)_{j \in \llbracket 0, J \rrbracket} \right\|_{\ell_\Delta^2} = \sup_{n \in \llbracket 0, N \rrbracket} \left\| (v_j^n)_{j \in \llbracket 0, J \rrbracket} - ([u_\Delta]_j^n)_{j \in \llbracket 0, J \rrbracket} \right\|_{\ell_\Delta^2},$$

with $(v_j^n)_{j \in \llbracket 0, J \rrbracket}$ the numerical solution computed with J cells in space and $([u_\Delta]_j^n)_{j \in \llbracket 0, J \rrbracket}$ the J -piecewise constant function from the analytical solution. However, for test cases with an unknown exact solution, the convergence error is computed from two numerical solutions with different meshes and E_J is replaced by the following $\tilde{E}_J$

$$\tilde{E}_J = \sup_{n \in \llbracket 0, N \rrbracket} \left\| (v_j^n)_{j \in \llbracket 0, J \rrbracket} - (\tilde{v}_j^n)_{j \in \llbracket 0, J \rrbracket} \right\|.$$

Section 3.7. Numerical results

Previously, $\tilde{v}_j^n := w_{2j}^n$ for all $j \in \llbracket 0, J \rrbracket$, where $(w_j^n)_{j \in \llbracket 0, 2J \rrbracket}$ is a numerical solution computed with a refined grid with $2J$ space meshes. In that case, $(\tilde{v}_j^n)_{j \in \llbracket 0, J \rrbracket}$, computed from the refined numerical solution $(w_j^n)_{j \in \llbracket 0, 2J \rrbracket}$, plays the role of the exact one $([u_\Delta]_j^n)_{j \in \llbracket 0, J \rrbracket}$.

Remark 3.51. We have considered exactly $2J$ space meshes for the refined numerical solution. Moreover, in all cases, 0 is a grid node, therefore, the error $\tilde{E}_J$ is nothing but the difference between every two consecutive values for $(w_j^n)_{j \in \llbracket 0, 2J \rrbracket}$ and every value for $(v_j^n)_{j \in \llbracket 0, J \rrbracket}$.

3.7.2 Smooth initial data

To assess the optimality of Theorem 3.7, the corresponding test cases are carried out with two smooth periodic initial data, either the cosinusoidal initial datum

$$u_0(x) = \cos\left(\frac{2\pi}{L}x\right),$$

or the so-called cnoidal-wave initial datum. This cnoidal-wave solution represents a periodic solitary wave solution of the Korteweg-de Vries equation whose analytical expression is known as follow :

$$u(t, x) = \frac{1}{\mu^{\frac{1}{5}}} \text{acn}^2\left(4K(m)\left(\mu^{\frac{2}{5}}\left(x - \frac{L}{2}\right) - v\mu^{\frac{1}{5}}t\right)\right),$$

where $\mu = \frac{1}{24^2}$ and $\text{cn}(z) = \text{cn}(z : m)$ is the Jacobi elliptic function with modulus $m \in (0, 1)$ (we choose $m = 0.9$) and the parameters have the values $a = 192m\mu K(m)^2$ and $v = 64\mu(2m - 1)K(m)^2$. $K(m)$ is the complete elliptic integral of the first kind (cf [13]).

Both results are gathered in Figure 3.1 for cosinusoidal solution and Figure 3.2 for cnoidal-wave solution. We display the values of r with respect to J in the left table and post the corresponding graph in logarithmic scale on the right. The first order is confirmed for both initial data whether in tables or in graphs.

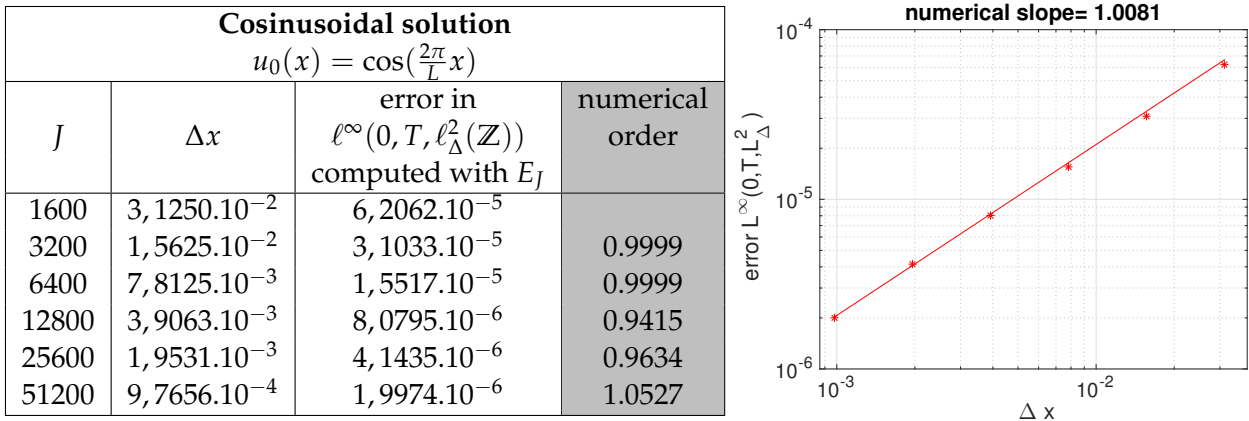


FIGURE 3.1 – Experimental rate of convergence for cosinusoidal solution

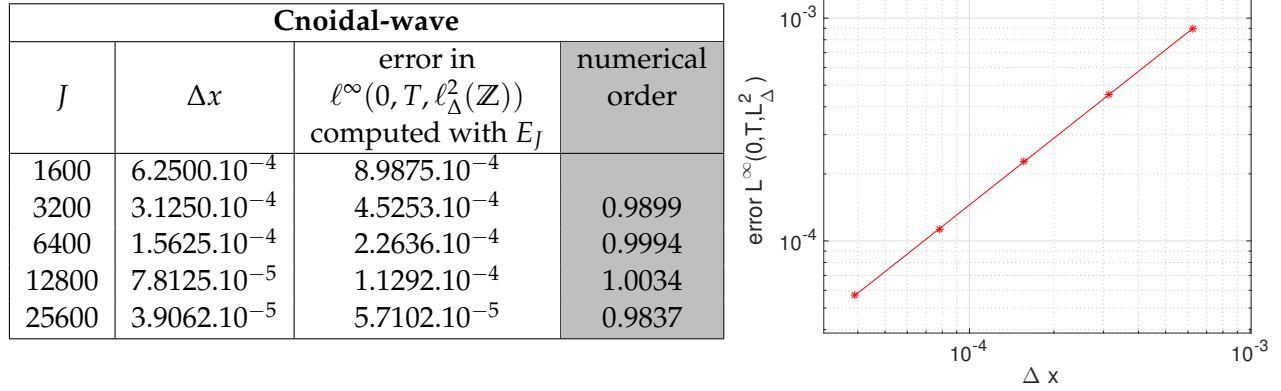


FIGURE 3.2 – Experimental rate of convergence for cnoidal-wave solution

3.7.3 For a regularity between $H^3(\mathbb{R})$ and $H^{10}(\mathbb{R})$

To illustrate numerically Theorem 3.10, we initialize the scheme with a less regular initial datum. We test two kinds of periodic initial data in $H^s(\mathbb{R})$, with $s \geq 3$. The first one is computed from some integrations of the indicator function and leads to initial data in $H^s(\mathbb{R})$ with s a half-integer. The second one is computed from the periodized square root function and leads to initial data in $H^s(\mathbb{R})$ with s an integer.

Tests achieved from the indicator function. Since the indicator function $\mathbb{1}_{[0, \frac{L}{2}]}$ belongs to $H^s([0, L])$ for all $s < \frac{1}{2}$, an idea to construct a periodic function in $H^{s+\ell}([0, L])$, with $s < \frac{1}{2}$ and $\ell \in \mathbb{N}^*$ is to integrate ℓ times the periodic indicator function. For instance, after a first integration, the initial datum

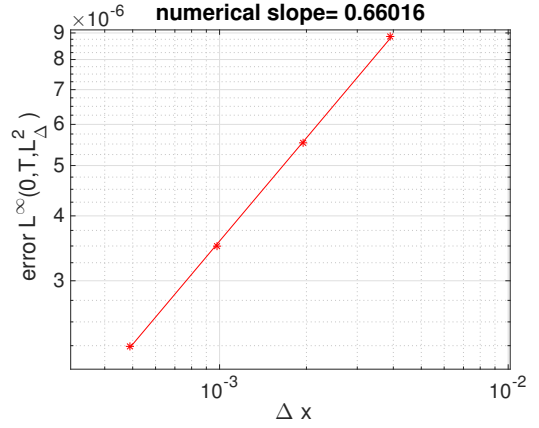
$$u_0(x) = x\mathbb{1}_{[0, \frac{L}{2}]} + (L - x)\mathbb{1}_{[\frac{L}{2}, L]}$$

is periodic and "almost" in $H^{\frac{3}{2}}([0, L])$. By reiterating the process of periodization and integration, we obtain initial data in $H^s([0, L])$, with $s = \frac{7}{2}^-, \frac{9}{2}^-, \frac{11}{2}^- \dots$. The numerical results are rather closed to the theoretical ones as summarized in the following table.

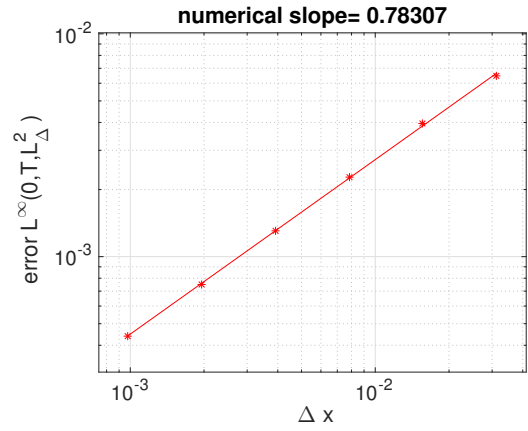
Sobolev regularity	$s = \frac{7}{2}^-$	$s = \frac{9}{2}^-$	$s = \frac{11}{2}^-$
Experimental convergence rate	0.66016	0.78307	0.97340
Theoretical convergence rate	0.58333	0.75000	0.91667

We detail some tests below : $s = \frac{7}{2}^-$ in Figure 3.3, and $s = \frac{9}{2}^-$ in Figure 3.4. For each figure, the graph on the right illustrates the error in $\ell^\infty(0, T, \ell_\Delta^2(\mathbb{Z}))$ with respect to the mesh size Δx in logarithmic scale. Those errors and those mesh sizes are detailed in the table on the left. On the graph, we have displayed the points cloud but also the least squares regression and the numerical slope of that regression line is considered to be the experimental convergence rate. For more simplicity, we denote $H^{s-}(\mathbb{R})$ the space $H^{s-\ell}(\mathbb{R})$ with $\ell > 0$ as small as possible.

$u_0 \in H^{\frac{7}{2}-}([0, L])$			
J	Δx	error in $\ell^\infty(0, T, \ell_\Delta^2(\mathbb{Z}))$ computed with $\tilde{E}_J$	numerical order
12800	$3.9063 \cdot 10^{-3}$	$8.8562 \cdot 10^{-6}$	0.6796 0.6637 0.6360
25600	$1.9531 \cdot 10^{-3}$	$5.5291 \cdot 10^{-6}$	
51200	$9.7656 \cdot 10^{-4}$	$3.4903 \cdot 10^{-6}$	
102400	$4.8828 \cdot 10^{-4}$	$2.2460 \cdot 10^{-6}$	


 FIGURE 3.3 – Experimental rate of convergence for $u_0 \in H^{\frac{7}{2}-}([0, L])$

$u_0 \in H^{\frac{9}{2}-}([0, L])$			
J	Δx	error in $\ell^\infty(0, T, \ell_\Delta^2(\mathbb{Z}))$ computed with $\tilde{E}_J$	numerical order
1600	$3.1250 \cdot 10^{-2}$	$6.5105 \cdot 10^{-3}$	0.71941 0.80574 0.78909 0.80504 0.764470
3200	$1.5625 \cdot 10^{-2}$	$3.9541 \cdot 10^{-3}$	
6400	$7.8125 \cdot 10^{-3}$	$2.2620 \cdot 10^{-3}$	
12800	$3.9063 \cdot 10^{-3}$	$1.3091 \cdot 10^{-3}$	
25600	$1.9531 \cdot 10^{-3}$	$7.4923 \cdot 10^{-4}$	
51200	$9.7656 \cdot 10^{-4}$	$4.4105 \cdot 10^{-4}$	


 FIGURE 3.4 – Experimental rate of convergence for $u_0 \in H^{\frac{9}{2}-}([0, L])$

Remark 3.52. The relative error between the experimental rate and the theoretical one is sometimes significant, for example, this relative error is more than 12% in the case $s = \frac{7}{2}-$. However, the theoretical rate is an asymptotic result for Δx and Δt small enough. As seen in the table on the left of Figure 3.3, the numerical order appears to be decreasing when Δx decreases, which agrees closely with the theoretical results.

Tests achieved from the square root function. Since the square root function is in $H^{1-}([0, L])$, we construct a $H^{s-}([0, L])$ function by integrating the square root function $s - 1$ times. However, we need, in addition, a periodic initial datum, this is why we add the beginning of a Taylor expansion for the function and its derivatives up to $(s - 1)$ -th to be continuous and periodic. More precisely, we search the coefficients $b_i, i \in \llbracket 1, s \rrbracket$ such that the function

$$x^{s-1+\frac{1}{2}} - b_1 x - \frac{b_2}{2} x^2 - \frac{b_3}{3!} x^3 \dots - \frac{b_s}{s!} x^s$$

and all its derivatives up to $(s - 1)$ -th be equal for $x = 0$ and for $x = L$. To find those coefficients, we just have to solve the matrix system

$$\begin{pmatrix} L & \frac{L^2}{2} & \frac{L^3}{3!} & \frac{L^4}{4!} & \dots & \frac{L^s}{s!} \\ 0 & L & \frac{L^2}{2} & \frac{L^3}{3!} & \dots & \frac{L^{s-1}}{(s-1)!} \\ 0 & 0 & L & \frac{L^2}{2} & \dots & \frac{L^{s-2}}{(s-2)!} \\ & & & \ddots & & \\ 0 & 0 & 0 & 0 & \dots & L \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_s \end{pmatrix} = \begin{pmatrix} L^{s-\frac{1}{2}} \\ (s-\frac{1}{2})L^{s-\frac{3}{2}} \\ (s-\frac{1}{2})(s-\frac{3}{2})L^{s-\frac{5}{2}} \\ \vdots \\ (s-\frac{1}{2})(s-\frac{3}{2})(s-\frac{5}{2}) \dots \frac{3}{2}L^{1/2} \end{pmatrix}.$$

The following tests correspond to $s \in \llbracket 4, 10 \rrbracket$.

The comparison between the experimental rates of convergence and the theoretical ones are very closed and gathered in the following table.

Sobolev regularity	$s = 4$	$s = 5$	$s = 6$	$s = 7$	$s = 8$	$s = 9$	$s = 10$
Experimental convergence rate	0.67225	0.86032	0.98708	0.99485	1.0060	1.0148	1.0062
Theoretical convergence rate	0.66667	0.83333	1.0000	1.0000	1.0000	1.0000	1.0000

Some results are developed below ($u_0 \in H^{4-}(\mathbb{R})$ in Figure 3.5 and $u_0 \in H^{5-}(\mathbb{R})$ in Figure 3.6).

$u_0 \in H^{4-}([0, L])$			
J	Δx	error in $\ell^\infty(0, T, \ell_\Delta^2(\mathbb{Z}))$ computed with $\tilde{E}_J$	numerical order
1600	$3.1250 \cdot 10^{-2}$	$4.6454 \cdot 10^{-3}$	0.72476 0.71307 0.65474 0.66290 0.67172
3200	$1.5625 \cdot 10^{-2}$	$2.8109 \cdot 10^{-3}$	
6400	$7.8125 \cdot 10^{-3}$	$1.7147 \cdot 10^{-3}$	
12800	$3.9063 \cdot 10^{-3}$	$1.0892 \cdot 10^{-3}$	
25600	$1.9531 \cdot 10^{-3}$	$6.8793 \cdot 10^{-4}$	
51200	$9.7656 \cdot 10^{-4}$	$4.3185 \cdot 10^{-4}$	

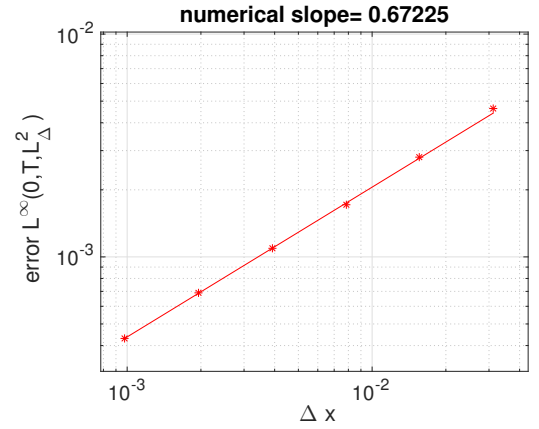


FIGURE 3.5 – Experimental rate of convergence for $u_0 \in H^{4-}([0, L])$

$u_0 \in H^{5-}([0, L])$			
J	Δx	error in $\ell^\infty(0, T, \ell_\Delta^2(\mathbb{Z}))$ computed with $\tilde{E}_J$	numerical order
1600	$3.1250 \cdot 10^{-2}$	$1.1066 \cdot 10^{-2}$	0.86208 0.88054 0.82941 0.86182 0.88424
3200	$1.5625 \cdot 10^{-2}$	$6.0878 \cdot 10^{-3}$	
6400	$7.8125 \cdot 10^{-3}$	$3.3067 \cdot 10^{-3}$	
12800	$3.9063 \cdot 10^{-3}$	$1.8609 \cdot 10^{-3}$	
25600	$1.9531 \cdot 10^{-3}$	$1.0240 \cdot 10^{-3}$	
51200	$9.7656 \cdot 10^{-4}$	$5.5475 \cdot 10^{-4}$	

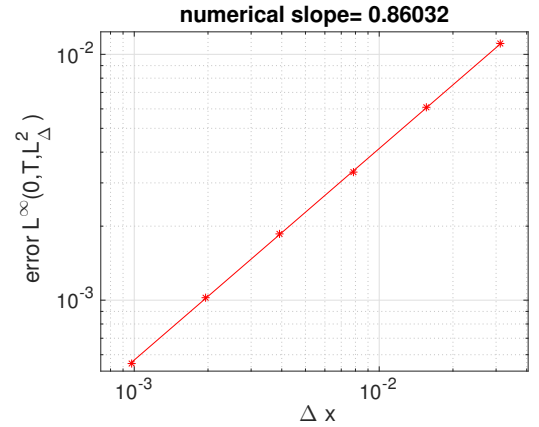


FIGURE 3.6 – Experimental rate of convergence for $u_0 \in H^{5-}([0, L])$

3.7.4 For a regularity between $L^2(\mathbb{R})$ and $H^3(\mathbb{R})$

Theoretically, the necessity to bound $\int_0^T \|\partial_x u(s, \cdot)\|_{L^\infty(\mathbb{R})}^i ds$ in (3.72) forces to choose $s \geq \frac{3}{4}$. In addition, the necessity to bound $\|e^n\|_{\ell^\infty}$ in G_a^{KdV} in (3.51a) in order to apply the Grönwall Lemma 2.29 leads to choose Δx such as Equation (3.67) be true which forces to have the constraint

Section 3.7. Numerical results

$\frac{1}{\delta^{6-s}} < \frac{1}{\Delta x^\gamma}$ in (3.75). However, those restrictions may be only technical and the rate of convergence seems to be $\Delta x^{\frac{s}{6}}$ for all $s \in [0, 3)$. This is based on the following numerical results.

Tests achieved from the indicator function. The integration from the indicator function enables us to obtain examples of functions in $H^s(\mathbb{R})$ with $s = \frac{1}{2}^-$, $\frac{3}{2}^-$, $\frac{5}{2}^-$, whose results are listed below.

Sobolev regularity	$s = \frac{1}{2}^-$	$s = \frac{3}{2}^-$	$s = \frac{5}{2}^-$
Experimental convergence rate	0.08795	0.25500	0.42595
"Conjectured" convergence rate	0.08333	0.25000	0.41667

More precisions are given in Figures 3.7 (for $u_0 \in H^{\frac{1}{2}^-}(\mathbb{R})$), 3.8 (for $u_0 \in H^{\frac{3}{2}^-}(\mathbb{R})$) and 3.9 (for $u_0 \in H^{\frac{5}{2}^-}(\mathbb{R})$).

$u_0 \in H^{\frac{1}{2}^-}([0, L])$			
J	Δx	error in $\ell^\infty(0, T, \ell_\Delta^2(\mathbb{Z}))$ computed with $\tilde{E}_J$	numerical order
3200	$1.5625 \cdot 10^{-2}$	$1.0567 \cdot 10^{-2}$	0.0864
6400	$7.8125 \cdot 10^{-3}$	$9.8843 \cdot 10^{-3}$	
12800	$3.9063 \cdot 10^{-3}$	$9.2992 \cdot 10^{-3}$	
25600	$1.9531 \cdot 10^{-3}$	$8.7490 \cdot 10^{-3}$	
51200	$9.7656 \cdot 10^{-4}$	$8.2289 \cdot 10^{-3}$	
102400	$4.8828 \cdot 10^{-4}$	$7.7468 \cdot 10^{-3}$	

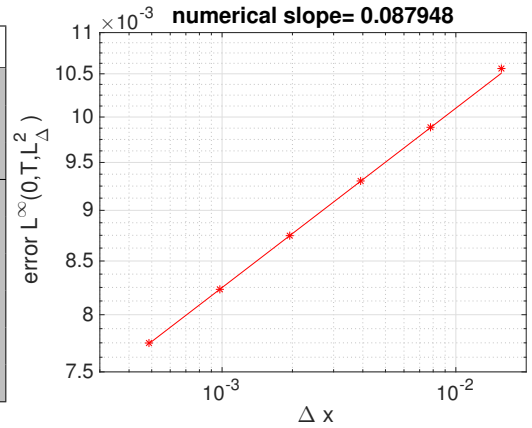


FIGURE 3.7 – Experimental rate of convergence for $u_0 \in H^{\frac{1}{2}^-}([0, L])$

$u_0 \in H^{\frac{3}{2}^-}([0, L])$			
J	Δx	error in $\ell^\infty(0, T, \ell_\Delta^2(\mathbb{Z}))$ computed with $\tilde{E}_J$	numerical order
3200	$1.5625 \cdot 10^{-2}$	$2.6584 \cdot 10^{-4}$	0.2660
6400	$7.8125 \cdot 10^{-3}$	$2.2108 \cdot 10^{-4}$	
12800	$3.9063 \cdot 10^{-3}$	$1.8624 \cdot 10^{-4}$	
25600	$1.9531 \cdot 10^{-3}$	$1.5632 \cdot 10^{-4}$	
51200	$9.7656 \cdot 10^{-4}$	$1.3074 \cdot 10^{-4}$	
102400	$4.8828 \cdot 10^{-4}$	$1.0949 \cdot 10^{-4}$	

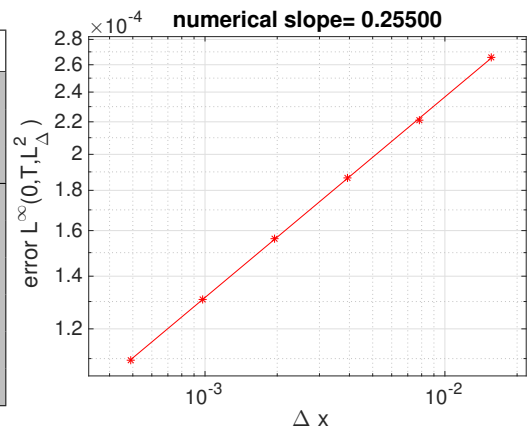
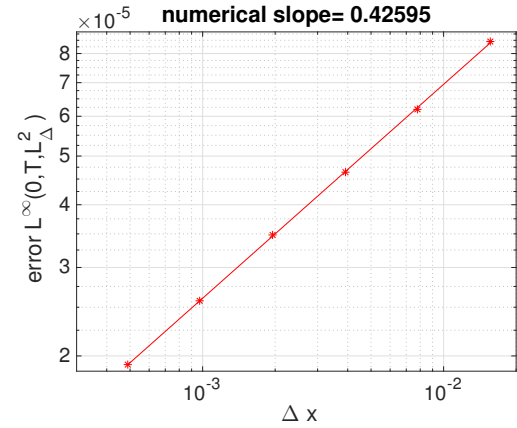


FIGURE 3.8 – Experimental rate of convergence for $u_0 \in H^{\frac{3}{2}^-}([0, L])$

$u_0 \in H^{\frac{5}{2}-}([0, L])$			
J	Δx	error in $\ell^\infty(0, T, \ell_\Delta^2(\mathbb{Z}))$ computed with $\tilde{E}_J$	numerical order
3200	$1.5625 \cdot 10^{-2}$	$8.4546 \cdot 10^{-5}$	0.4482 0.4133 0.4214 0.4297 0.4260
6400	$7.8125 \cdot 10^{-3}$	$6.1968 \cdot 10^{-5}$	
12800	$3.9063 \cdot 10^{-3}$	$4.6531 \cdot 10^{-5}$	
25600	$1.9531 \cdot 10^{-3}$	$3.4745 \cdot 10^{-5}$	
51200	$9.7656 \cdot 10^{-4}$	$2.5795 \cdot 10^{-5}$	
102400	$4.8828 \cdot 10^{-4}$	$1.9199 \cdot 10^{-5}$	

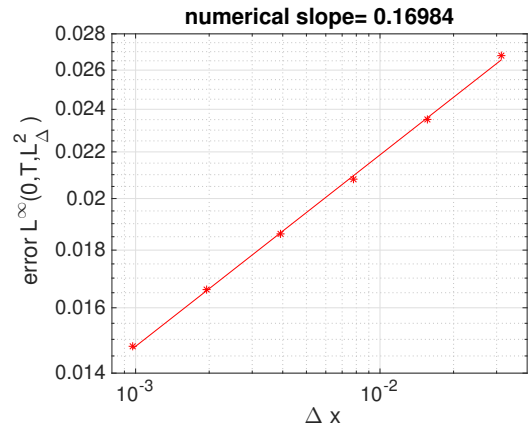

 FIGURE 3.9 – Experimental rate of convergence for $u_0 \in H^{\frac{5}{2}-}([0, L])$

Tests achieved from the square root function. As before, the second test consists in periodizing the square root function to create an initial datum in $H^s(\mathbb{R})$ with $s = 1, 2, 3$.

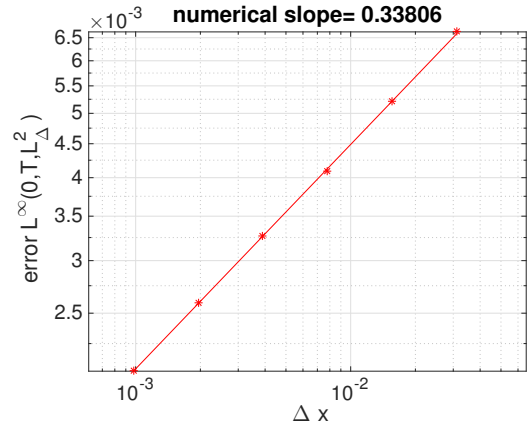
Sobolev regularity	$s = 1$	$s = 2$	$s = 3$
Experimental convergence rate	0.16984	0.33806	0.50173
"Conjectured" convergence rate	0.16667	0.33333	0.50000

Some precisions for those tests are given in Figures 3.10, 3.11 and 3.12.

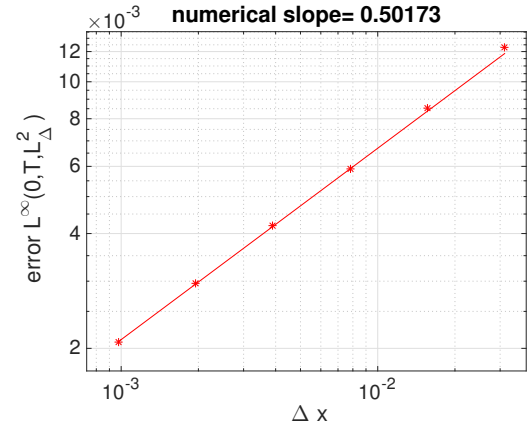
$u_0 \in H^{1-}([0, L])$			
J	Δx	error in $\ell^\infty(0, T, \ell_\Delta^2(\mathbb{Z}))$ computed with $\tilde{E}_J$	numerical order
1600	$3.1250 \cdot 10^{-2}$	$2.6762 \cdot 10^{-2}$	0.18748 0.17660 0.16119 0.16360 0.16701
3200	$1.5625 \cdot 10^{-2}$	$2.3501 \cdot 10^{-2}$	
6400	$7.8125 \cdot 10^{-3}$	$2.0793 \cdot 10^{-2}$	
12800	$3.9063 \cdot 10^{-3}$	$1.8595 \cdot 10^{-2}$	
25600	$1.9531 \cdot 10^{-3}$	$1.6602 \cdot 10^{-2}$	
51200	$9.7656 \cdot 10^{-4}$	$1.4787 \cdot 10^{-2}$	


 FIGURE 3.10 – Experimental rate of convergence for $u_0 \in H^{1-}([0, L])$

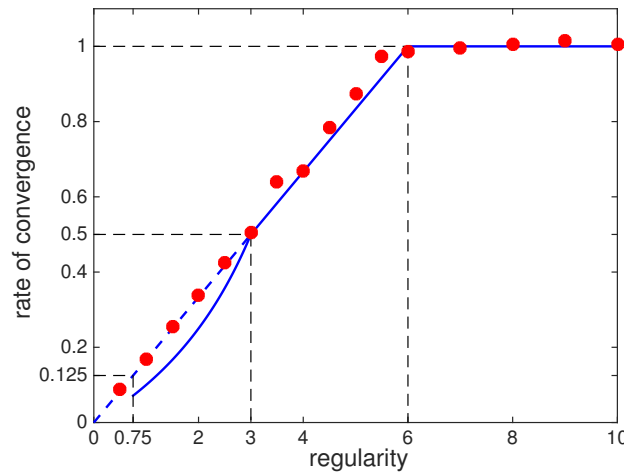
$u_0 \in H^2-([0, L])$			
J	Δx	error in $\ell^\infty(0, T, \ell_\Delta^2(\mathbb{Z}))$ computed with $\tilde{E}_J$	numerical order
1600	$3.1250 \cdot 10^{-2}$	$6.6322 \cdot 10^{-3}$	0.34779
3200	$1.5625 \cdot 10^{-2}$	$5.2115 \cdot 10^{-3}$	
6400	$7.8125 \cdot 10^{-3}$	$4.0950 \cdot 10^{-3}$	
12800	$3.9063 \cdot 10^{-3}$	$3.2699 \cdot 10^{-3}$	
25600	$1.9531 \cdot 10^{-3}$	$2.5937 \cdot 10^{-3}$	
51200	$9.7656 \cdot 10^{-4}$	$2.0449 \cdot 10^{-3}$	


 FIGURE 3.11 – Experimental rate of convergence for $u_0 \in H^2-([0, L])$

$u_0 \in H^3-([0, L])$			
J	Δx	error in $\ell^\infty(0, T, \ell_\Delta^2(\mathbb{Z}))$ computed with $\tilde{E}_J$	numerical order
1600	$3.1250 \cdot 10^{-2}$	$1.2295 \cdot 10^{-2}$	0.52599
3200	$1.5625 \cdot 10^{-2}$	$8.5386 \cdot 10^{-3}$	
6400	$7.8125 \cdot 10^{-3}$	$5.8964 \cdot 10^{-3}$	
12800	$3.9063 \cdot 10^{-3}$	$4.1913 \cdot 10^{-3}$	
25600	$1.9531 \cdot 10^{-3}$	$2.9576 \cdot 10^{-3}$	
51200	$9.7656 \cdot 10^{-4}$	$2.0778 \cdot 10^{-3}$	


 FIGURE 3.12 – Experimental rate of convergence for $u_0 \in H^3-([0, L])$

We summarize the theoretical and numerical results in Figure 3.13. The blue line corresponds to the theoretical rate of convergence, the dashed line matches the conjectured rate and the red dots stand for the numerical rates of convergence. Both are intertwined, which validates the rate of convergence of $\frac{\min(s, 6)}{6}$ with s the Sobolev regularity of the initial value.


 FIGURE 3.13 – Rates of convergence according to the Sobolev regularity of u_0 . – Rates proved in this chapter (solid line) versus experimental rates (dots)

3.8 Afterword : go one step further in the Burgers and Airy structure of (KdV)

As explain in Section 3.1, the KdV equation is a "combination" between the Burgers equation and the Airy one. The previous convergence study could be used to recover (granted in an unusual way) the classical CFL conditions for Airy and Burgers equations and hence enables to understand the contribution of both terms.

3.8.1 The Airy equation

We focus in this subsection on the dispersive term of KdV equation, the Airy equation

$$\partial_t u(t, x) + \partial_x^3 u(t, x) = 0, \quad (t, x) \in [0, T] \times \mathbb{R}.$$

We consider the right winded scheme (2.4a) with (2.5a) (with $p = 1$) with initialization (2.4b). By performing exactly the same proof than the one for (KdV), we recover the Airy equivalent of Proposition 3.25, which enables to highlight the dispersive CFL condition (by an unusual way).

Proposition 3.53. *Let $(e_j^n)_{(n,j)}$ and $(\epsilon_j^n)_{(n,j)}$ be respectively the convergence and the consistency error for the θ -linear scheme (2.4a)-(2.4b) with (2.5a) apply to Airy equation. Then, the following discrete relative entropy inequality is true, for $\mathcal{A}$ defined by (3.46)*

$$\begin{aligned} \|\mathcal{A}e^{n+1}\|_{\ell_\Delta^2}^2 &\leq \|\mathcal{A}e^n\|_{\ell_\Delta^2}^2 + \Delta t \|e^n - (1 - \theta)\Delta t D_+ D_+ D_-(e)^n\|_{\ell_\Delta^2}^2 + \Delta t \|\epsilon^n\|_{\ell_\Delta^2}^2 \{\Delta t + 1\} \\ &\quad + \Delta t C_e^{\text{Airy}} \|D_+ D_c(e)^n\|_{\ell_\Delta^2}^2 + \Delta t C_f^{\text{Airy}} \|D_+ D_+ D_-(e)^n\|_{\ell_\Delta^2}^2, \end{aligned}$$

where, the constants $C_\bullet^{\text{Airy}}$ are defined as follow

$$C_e^{\text{Airy}} := -\Delta x, \tag{3.77}$$

$$C_f^{\text{Airy}} := \Delta t(1 - 2\theta) - \frac{\Delta x^3}{4}. \tag{3.78}$$

Remark 3.54. *The same dispersive CFL conditions are recovered, both with Fourier analysis (cf Table 2.1, with $p = 1$) and within the relative entropy method, which consolidates our intuition that this second method is a good alternative for equations which combine hyperbolic and dispersive terms.*

3.8.2 The Burgers equation

Unsurprisingly, the dispersive operator $\mathcal{A}$ comes from the Airy term. It is implied in the next subsection that the induction method comes from the Burgers term. To prove this hint, let us consider the non viscous Burgers equation

$$\partial_t u(t, x) + \partial_x \left(\frac{u^2}{2} \right) (t, x) = 0, \quad (t, x) \in [0, T] \times \mathbb{R},$$

discretized thanks to the Rusanov scheme

$$\frac{v_j^{n+1} - v_j^n}{\Delta t} + D_c \left(\frac{v^2}{2} \right)_j^n = \frac{\tau \Delta x}{2} D_+ D_-(v)_j^n, \quad (n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z},$$

with initial condition

$$v_j^0 = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u_0(y) dy, \quad j \in \mathbb{Z}.$$

Section 3.8. Afterword : go one step further in the Burgers and Airy structure of (KdV)

Remark 3.55. Notice that the regularity of the initial data is not always propagated at any time : Burgers equation can create discontinuous solution, named shock, even from a smooth initial datum. However, in this subsection, we focus only on smooth exact solution and consider the final time T small enough such that $T < T_{\text{blow up}}$ with $T_{\text{blow up}}$ the blow up time, if $T_{\text{blow up}}$ exists. The exacte solution is taken to be equal to the strong solution with a sufficient regularity, for example $u \in C^2([0, T], \mathbb{R})$.

3.8.2.1 The well-known hyperbolic CFL

The convergence error satisfies

$$e_j^{n+1} = e_j^n - \Delta t D_c \left(\frac{e}{2} \right)_j^n - \Delta t D_c (u_\Delta e)_j^n + \frac{\tau \Delta t \Delta x}{2} D_+ D_- (e)_j^n - \Delta t \epsilon_j^n, \quad (n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}, \quad (3.79)$$

and verifies the following property, which is the Burgers equivalent of Proposition 3.25.

Proposition 3.56. Let $(e_j^n)_{(n,j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}}$ be the convergence error satisfying Equation (3.79). For every $(n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}$, ones has

$$\begin{aligned} \|e^{n+1}\|_{\ell_\Delta^2}^2 &\leq \|e^n\|_{\ell_\Delta^2}^2 + \Delta t \left(1 + C_a^{\text{Burgers}} \right) \|e^n\|_{\ell_\Delta^2}^2 + \Delta t \|\epsilon^n\|_{\ell_\Delta^2}^2 \left\{ 1 + \Delta t + 4 \frac{\Delta t}{\Delta x} \right\} \\ &\quad + \Delta t C_b^{\text{Burgers}} \|D_+ (e)^n\|_{\ell_\Delta^2}^2 + \Delta t^2 C_c^{\text{Burgers}} \|D_c (e)^n\|_{\ell_\Delta^2}^2. \end{aligned}$$

The variables $C_\bullet^{\text{Burgers}}$ are defined below

$$\begin{aligned} C_a^{\text{Burgers}} &:= \|(u_\Delta)^n\|_{\ell_\infty}^2 + \|D_+ (u_\Delta)^n\|_{\ell^\infty} \left(1 + \frac{\Delta t}{\Delta x} \left[2\tau + \frac{2}{3} \|e^n\|_{\ell^\infty} + \frac{3}{2} \|(u_\Delta)^n\|_{\ell^\infty} \right] \right) \\ &\quad + \frac{\Delta t^2}{\Delta x^2} \|D_+ (u_\Delta)^n\|_{\ell^\infty}^2 + \frac{\Delta t}{\Delta x} (\|(u_\Delta)^n\|_{\ell^\infty}^2 + 2\tau^2), \end{aligned} \quad (3.80a)$$

$$C_b^{\text{Burgers}} := \left(\frac{\Delta x}{6} D_+ (e)_j^n - \tau \right) (\Delta x - \tau \Delta t), \quad (3.80b)$$

$$C_c^{\text{Burgers}} := \|e^n\|_{\ell^\infty}^2 [1 + \Delta x] + \|(u_\Delta)^n\|_{\ell^\infty}^2 - \tau^2 + 2\|e^n\|_{\ell^\infty} \|(u_\Delta)^n\|_{\ell^\infty} + \frac{2\tau}{3} \|e^n\|_{\ell^\infty}. \quad (3.80c)$$

We recognize in C_a^{Burgers} and C_b^{Burgers} , the factor C_a^{KdV} and C_b^{KdV} (without θ -terms which provided from the Airy part). The term C_c^{Burgers} is exactly the same as C_c^{KdV} .

Remark 3.57. We have to upper bound the norm $\|e^n\|_{\ell^\infty}$ in C_a^{Burgers} -term. We could use an induction method as the one developed for (KdV) in Subsection 3.5.3, however, the Burgers equation and the Rusanov scheme verify the well-known maximum principle :

$$\min_{x \in \mathbb{R}} (u_0(x)) \leq u(t, x) \leq \max_{x \in \mathbb{R}} (u_0(x)), \quad \forall (t, x) \in [0, T] \times \mathbb{R},$$

and

$$\min_{x \in \mathbb{R}} (u_0(x)) \leq v_j^n \leq \max_{x \in \mathbb{R}} (u_0(x)), \quad \forall (n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}.$$

Therefore, we know a natural bound for the convergence error : $\|e^n\|_{\ell^\infty} \leq 2\|u_0\|_{L^\infty}$.

Remark 3.58. The hyperbolic CFL condition $\tau \Delta t \leq \Delta x$ is imposed to obtain $C_b^{\text{Burgers}} \leq 0$ and the Rusanov coefficient τ is chosen such that $\|u_0\|_{L^\infty} < \tau$ to obtain $C_c^{\text{Burgers}} \leq 0$. Once again, the maximum principle enables to impose the classical condition $\|u_0\|_{L^\infty} < \tau$ instead of $\|u_\Delta^n\|_{\ell^\infty} < \tau$. Note that the hyperbolic CFL is the classical one, without $\frac{1}{2}$ as for KdV equation in (3.8). The additional $\frac{1}{2}$ comes from the cross terms between Airy part and Burgers part when we square Equation (3.36) in the proof of Proposition 3.25

We will just digress in this subsection and move away from the convergence proof, to caution that preliminaries lemmas of Subsection 3.4.1 are also useful to compute explicitly numerical entropy flux. All the discrete integrations by part of the Subsection 3.4.1 enable us to explicit the numerical entropy flux $\mathcal{G}$ associated with the quadratic entropy $\eta_0(u) = \frac{u^2}{2}$.

Remark 3.59. We recall that the continuous entropy flux g_0 associated to this quadratic entropy η_0 is $g_0(u) := \frac{u^3}{3}$.

Proposition 3.60. Assume the CFL conditions

$$\begin{cases} 2\|u_0\|_{L^\infty(\mathbb{R})}^2 \leq \tau^2, \\ \tau\Delta t \leq \Delta x, \end{cases} \quad (3.81a)$$

$$(3.81b)$$

be true then, the numerical entropy flux associated with a quadratic entropy is

$$\mathcal{G}(v, u) := \left(1 + \frac{\tau\Delta t}{\Delta x}\right) \frac{v^3}{12} + \left(1 - \frac{\tau\Delta t}{\Delta x}\right) \left(\frac{v(u)^2}{4} + \frac{\tau uv}{2}\right) - \tau \left(1 - \frac{\tau\Delta t}{2\Delta x}\right) \frac{v^2}{2} + \frac{\tau\Delta t}{\Delta x} \left(\frac{\tau}{4} + \frac{u}{6}\right) u^2.$$

More precisely, the numerical solution $(v_j^n)_{(n,j)}$ verifies

$$\frac{(v_j^{n+1})^2 - (v_j^n)^2}{2\Delta t} + \frac{\mathcal{G}(v_{j+1}^n, v_j^n) - \mathcal{G}(v_j^n, v_{j-1}^n)}{\Delta x} \leq 0, \quad (n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}.$$

Remark 3.61. Those numerical entropy flux $\mathcal{G}$ is consistent with the (continuous) entropy flux g_0 . Indeed

$$\begin{aligned} \mathcal{G}(u, u) &= \left(1 + \frac{\tau\Delta t}{\Delta x}\right) \frac{u^3}{12} + \left(1 - \frac{\tau\Delta t}{\Delta x}\right) \left(\frac{u^3}{4} + \frac{\tau u^2}{2}\right) - \tau \left(1 - \frac{\tau\Delta t}{2\Delta x}\right) \frac{u^2}{2} + \frac{\tau\Delta t}{\Delta x} \left(\frac{\tau}{4} + \frac{u}{6}\right) u^2 \\ &= u^3 \left[\frac{1}{12} + \frac{\tau\Delta t}{12\Delta x} + \frac{1}{4} - \frac{\tau\Delta t}{4\Delta x} + \frac{\tau\Delta t}{6\Delta x} \right] + u^2 \left[\frac{\tau}{2} - \frac{\tau^2\Delta t}{2\Delta x} - \tau \left(\frac{1}{2} - \frac{\tau\Delta t}{4\Delta x} \right) + \frac{\tau^2\Delta t}{4\Delta x} \right] \\ &= \frac{1}{3} u^3 \\ &= g_0(u). \end{aligned}$$

Before proving Proposition 3.60, we have to enforce some preliminaries lemmas of Subsection 3.4.1. Indeed, since we do not sum over $j \in \mathbb{Z}$ in the present subsection, some terms of preliminaries lemmas do not cancel each other any more, and we have to be more precise to determine them.

We recall the shift operator $\mathcal{S}^\ell$, introduced in Subsection 2.2.1.3, defined by

$$\left(\mathcal{S}^1 a\right)_j := a_{j+1}, \quad \left(\mathcal{S}^{-1} a\right)_j := a_{j-1}.$$

First of all, the ℓ^2 -norm identities of Lemma 3.20 are changed as follows.

Lemma 3.62. For $(a_j)_{j \in \mathbb{Z}}$ a sequence in $\ell_\Delta^2(\mathbb{Z})$, then

$$\left[D_c \left(\frac{a^2}{2} \right)_j \right]^2 = \left[D_c(a)_j \right]^2 \left(\frac{a_{j+1} + a_{j-1}}{2} \right)^2, \quad (3.82)$$

and

$$\left[D_+ D_- (a)_j \right]^2 = \frac{4}{\Delta x^2} \left[D_+(a)_j \right]^2 + \frac{4}{\Delta x^3} D_+ \left(a \mathcal{S}^{-1} a \right)_j - \frac{4}{\Delta x^2} \left[D_c(a)_j \right]^2 - \frac{4}{\Delta x^3} D_c(a^2)_j. \quad (3.83)$$

Proof. All the proofs of this section are gathered in Appendix B.2. $\square$

The discrete integrations by parts of Corollary 3.23 are specified in the following lemma.

Lemma 3.63. Assume $(a_j)_{j \in \mathbb{Z}}$ be a sequence in $\ell_\Delta^2(\mathbb{Z})$, then it holds

$$a_j D_c \left(\frac{a^2}{2} \right)_j = -\frac{\Delta x^2}{12} [D_+(a)_j]^3 + \frac{1}{12} D_+(a^3)_j + \frac{1}{4} D_+(aS^{-1}a^2)_j. \quad (3.84)$$

Moreover,

$$\begin{aligned} D_+ D_-(a)_j D_c \left(\frac{a^2}{2} \right)_j &= -\frac{2}{3} [D_c(a)_j]^3 + \frac{2}{3\Delta x^2} D_c(a^3)_j + \frac{1}{6} [D_+(a)_j]^3 \\ &\quad - \frac{1}{6\Delta x^2} D_+(a^3)_j - \frac{1}{2\Delta x^2} D_+(aS^{-1}a^2)_j. \end{aligned} \quad (3.85)$$

Last but not least,

$$a_j D_+ D_-(a)_j = -[D_+(a)_j]^2 + \frac{1}{\Delta x} D_+(a^2)_j - \frac{1}{\Delta x} D_+(aS^{-1}a)_j. \quad (3.86)$$

Proof. Those inequalities are proved in Appendix B.2. $\square$

Remark 3.64. Integrating previous lemmas enables us to find Lemma 3.20 and Corollary 3.23 of Subsection 3.4.1.

We are now able to compute the numerical entropy flux $\mathcal{G}$.

Proof of Proposition 3.60. As to obtain the discrete relative entropy inequality, we square numerical scheme. The resulting equality is similar to the one in the beginning of the proof of Proposition 3.25, but without sums, without ϵ_j^n and without u_Δ :

$$(v_j^{n+1})^2 = (\text{rhs}_a^n)^2 + (\text{rhs}_b^n)^2 \quad (3.87)$$

with

$$\begin{aligned} (\text{rhs}_a^n)^2 &:= (v_j^n)^2 + \Delta t^2 \left[D_c \left(\frac{v^2}{2} \right)_j \right]^2 + \frac{\tau^2 \Delta t^2 \Delta x^2}{4} [D_+ D_-(v)_j^n]^2, \\ (\text{rhs}_b^n)^2 &:= -2\Delta t v_j^n D_c \left(\frac{v^2}{2} \right)_j + \tau \Delta x \Delta t v_j^n D_+ D_-(v)_j^n - \tau \Delta t^2 \Delta x D_c \left(\frac{v^2}{2} \right)_j D_+ D_-(v)_j^n. \end{aligned}$$

Lemma 3.62 enables us to simplify $(\text{rhs}_a^n)^2$ term

$$\begin{aligned} (\text{rhs}_a^n)^2 &= (v_j^n)^2 + \Delta t^2 [D_c(v)_j^n]^2 \left\{ \left(\frac{v_{j+1}^n + v_{j-1}^n}{2} \right)^2 - \tau^2 \right\} + \tau^2 \Delta t^2 [D_+(v)_j^n]^2 \\ &\quad + \frac{\tau^2 \Delta t^2}{\Delta x} D_+(vS^{-1}v)_j^n - \frac{\tau^2 \Delta t^2}{\Delta x} D_c(v^2)_j^n. \end{aligned}$$

Lemma 3.63 gives

$$\begin{aligned} (\text{rhs}_b^n)^2 &= \Delta x \Delta t [D_+(v)_j^n]^2 \left\{ \frac{\Delta x}{6} D_+(v)_j^n - \tau - \frac{\tau \Delta t}{6} D_+(v)_j^n \right\} + \frac{\Delta t}{2} D_+(vS^{-1}v^2)_j^n \left\{ \frac{\tau \Delta t}{\Delta x} - 1 \right\} \\ &\quad + \tau \Delta t D_+(v^2)_j^n - \tau \Delta t D_+(vS^{-1}v)_j^n + \frac{2\tau \Delta t^2 \Delta x}{3} [D_c(v)_j^n]^3 - \frac{2\tau \Delta t^2}{3\Delta x} D_c(v^3)_j^n + \frac{\Delta t}{6} D_+(v^3)_j^n \left\{ \frac{\tau \Delta t}{\Delta x} - 1 \right\}. \end{aligned}$$

Equation (3.87) becomes

$$\begin{aligned} \frac{(v_j^{n+1})^2 - (v_j^n)^2}{2\Delta t} &= D_+ \left(\left\{ \frac{\tau\Delta t}{\Delta x} - 1 \right\} \left\{ \frac{\tau v \mathcal{S}^{-1}v}{2} + \frac{v \mathcal{S}^{-1}v^2}{4} + \frac{v^3}{12} \right\} \right)_j^n + \frac{\tau}{2} D_+ (v^2)_j^n \\ &\quad - D_c \left(\frac{\tau\Delta t}{\Delta x} v^2 \left\{ \frac{v}{3} + \frac{\tau}{2} \right\} \right)_j^n + \frac{\Delta x}{2} [D_+(v)_j^n]^2 \left\{ \frac{\Delta x}{6} D_+(v)_j^n - \tau \right\} \left\{ 1 - \frac{\tau\Delta t}{\Delta x} \right\} \\ &\quad + \frac{\Delta t}{2} [D_c(v)_j^n]^2 \left\{ \left(\frac{v_{j+1}^n + v_{j-1}^n}{2} \right)^2 - \tau^2 + \frac{2\tau\Delta x}{3} D_c(v)_j^n \right\}. \end{aligned} \quad (3.88)$$

However, since $D_c = \frac{D_+ + D_-}{2}$, the D_c term can be simplified into

$$\begin{aligned} D_c \left(\frac{\tau\Delta t}{\Delta x} v^2 \left\{ \frac{v}{3} + \frac{\tau}{2} \right\} \right)_j^n &= D_+ \left(\frac{\tau\Delta t}{\Delta x} v^2 \left\{ \frac{v}{6} + \frac{\tau}{4} \right\} \right)_j^n + D_- \left(\frac{\tau\Delta t}{\Delta x} v^2 \left\{ \frac{v}{6} + \frac{\tau}{4} \right\} \right)_j^n \\ &= D_+ \left(\frac{\tau\Delta t}{\Delta x} v^2 \left\{ \frac{v}{6} + \frac{\tau}{4} \right\} \right)_j^n + D_+ \left(\frac{\tau\Delta t}{\Delta x} \mathcal{S}^{-1}v^2 \left\{ \frac{\mathcal{S}^{-1}v}{6} + \frac{\tau}{4} \right\} \right)_j^n \end{aligned}$$

and the numerical flux writes

$$\begin{aligned} &D_+ \left(\left\{ \frac{\tau\Delta t}{\Delta x} - 1 \right\} \left\{ \frac{\tau v \mathcal{S}^{-1}v}{2} + \frac{v \mathcal{S}^{-1}v^2}{4} + \frac{v^3}{12} \right\} \right)_j^n + \frac{\tau}{2} D_+ (v^2)_j^n - D_c \left(\frac{\tau\Delta t}{\Delta x} v^2 \left\{ \frac{v}{3} + \frac{\tau}{2} \right\} \right)_j^n \\ &= D_+ \left(\left\{ -\frac{\tau\Delta t}{\Delta x} - 1 \right\} \frac{v^3}{12} \right)_j^n + D_+ \left(\left\{ \frac{\tau\Delta t}{\Delta x} - 1 \right\} \left\{ \frac{\tau v \mathcal{S}^{-1}v}{2} + \frac{v \mathcal{S}^{-1}v^2}{4} \right\} \right)_j^n + D_+ \left(v^2 \frac{\tau}{2} \left\{ 1 - \frac{\tau\Delta t}{2\Delta x} \right\} \right)_j^n \\ &\quad - D_+ \left(\frac{\tau\Delta t}{\Delta x} \mathcal{S}^{-1}v^2 \left\{ \frac{\mathcal{S}^{-1}v}{6} + \frac{\tau}{4} \right\} \right)_j^n. \end{aligned}$$

Thus, Equation (3.88) is transformed into

$$\begin{aligned} \frac{(v_j^{n+1})^2 - (v_j^n)^2}{2\Delta t} + D_+ (\mathcal{G}(v, \mathcal{S}^{-1}v))_j^n &= \frac{\Delta x}{2} [D_+(v)_j^n]^2 \left\{ \frac{\Delta x}{6} D_+(v)_j^n - \tau \right\} \left\{ 1 - \frac{\tau\Delta t}{\Delta x} \right\} \\ &\quad + \frac{\Delta t}{2} [D_c(v)_j^n]^2 \left\{ \left(\frac{v_{j+1}^n + v_{j-1}^n}{2} \right)^2 - \tau^2 + \frac{2\tau\Delta x}{3} D_c(v)_j^n \right\}, \end{aligned} \quad (3.89)$$

with

$$\begin{aligned} \mathcal{G}(v, \mathcal{S}^{-1}v) &:= \left\{ 1 + \frac{\tau\Delta t}{\Delta x} \right\} \frac{v^3}{12} + \left\{ 1 - \frac{\tau\Delta t}{\Delta x} \right\} \left\{ \frac{\tau v \mathcal{S}^{-1}v}{2} + \frac{v (\mathcal{S}^{-1}v)^2}{4} \right\} - v^2 \frac{\tau}{2} \left\{ 1 - \frac{\tau\Delta t}{2\Delta x} \right\} \\ &\quad + \frac{\tau\Delta t}{\Delta x} (\mathcal{S}^{-1}v)^2 \left\{ \frac{\mathcal{S}^{-1}v}{6} + \frac{\tau}{4} \right\}. \end{aligned}$$

Let us focus on the right hand side of (3.89). CFL conditions (3.81a)-(3.81b) together with the maximum principle imply the non positivity of the coefficient behind $[D_+(v)_j^n]^2$. For the second

term, one has

$$\begin{aligned}
 \left(\frac{v_{j+1}^n + v_{j-1}^n}{2} \right)^2 - \tau^2 + \frac{2\tau\Delta x}{3} D_c(v)_j^n &\leq \|v^n\|_{\ell^\infty}^2 - \tau^2 + \frac{\tau v_{j+1}^n - \tau v_{j-1}^n}{3} \\
 &\leq \|v^n\|_{\ell^\infty}^2 - \tau^2 + \frac{\tau^2}{3} + \frac{\|v^n\|_{\ell^\infty}^2}{3} \\
 &\leq \frac{2}{3} [2\|v^n\|_{\ell^\infty}^2 - \tau^2] \\
 &\leq 0.
 \end{aligned}$$

□

CHAPITRE 4

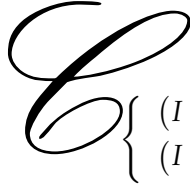
Energy estimates for a finite difference scheme on the $abcd$ system

The $abcd$ system is studied numerically according to a specific finite difference scheme. The non linear term is discretized with respect to a Rusanov scheme and a center scheme is used for the BBM- and KdV-type dispersion. To establish the convergence rate of this scheme, we develop a method which generalizes the one used for (KdV) equation in Chapter 3. The main step consists on establishing an energy estimate which mimics the continuous one in [27]. We demonstrate the first order accurate in time and space and illustrate our theoretical results with some experimental simulations done on solitary waves.

Summary

4.1	Introduction (en français)	115
4.2	One step towards the $abcd$ system : the KdV-BBM equation	120
4.2.1	Left and right winded numerical schemes	120
4.2.1.1	The case of a non negative dispersion ($a \geq 0$)	121
4.2.1.2	The case of a negative dispersion ($a < 0$)	124
4.2.2	For a centered scheme	124
4.3	Non zero BBM-dispersive term in both equations : $abcd$ system with $b, d > 0$	125
4.3.1	Discrete energy inequality	126
4.3.2	Convergence of the scheme	130
4.4	KdV equation coupled with BBM equation : $abcd$ system with $bd = 0$	131
4.4.1	Burgers-type estimates	132
4.4.2	The case $b = d = 0$	135
4.4.3	The case $b = 0, d > 0$	137
4.4.4	The case $d = 0, b > 0$	142
4.5	Experimental results	147
4.5.1	Linear case	147
4.5.2	Case $b > 0, d > 0$	148
4.5.3	Case $bd = 0$	151
4.6	Afterword : traveling waves collision	152

4.1 Introduction (en français)



considérons le système $abcd$ avec coefficients unitaires

$$\begin{cases} (I - b\partial_x^2) \partial_t \eta + \partial_x u + a\partial_x^3 u + \partial_x (\eta u) = 0, \\ (I - d\partial_x^2) \partial_t u + \partial_x \eta + c\partial_x^3 \eta + \partial_x \left(\frac{u^2}{2}\right) = 0, \end{cases} \quad (t, x) \in [0, T] \times \mathbb{R}, \quad (abcd)$$

avec la contrainte

$$a + b + c + d = \frac{1}{3} - Bo, \quad \text{avec } Bo \text{ le nombre de Bond.} \quad (4.1)$$



Schémas numériques pour le système $abcd$. Le but de ce chapitre est de généraliser l'étude numérique faite sur l'équation de (KdV) au chapitre 3 à un système de deux équations couplées : le système $abcd$. La partie non linéaire est, comme pour (KdV), discrétisée au moyen d'un schéma de Rusanov. Nous noterons par la suite τ_1 et τ_2 les deux coefficients de viscosité de Rusanov. Par contre, une discrétisation centrée pour la partie dispersive a été préférée à la discrétisation décentrée de l'équation de (KdV) afin d'éviter les problèmes de stabilité en fonction du décentrement, déjà évoqués au chapitre 2. Pour ne conserver qu'une condition CFL hyperbolique, nous avons implicité les termes de dispersion et de transport, ou utilisé un schéma de Crank-Nicolson pour certains cas.

Soient Δt et Δx les pas de temps et d'espace. Soit (η, u) la solution régulière du système (abcd), comme précédemment, nous considérons leur moyenne sur une maille en espace et en temps $\eta_\Delta, u_\Delta \in \ell^\infty([0, N]; \ell_\Delta^2(\mathbb{Z}))$ définies par

$$\begin{cases} (\eta_\Delta)_j^n := \frac{1}{[\inf(T, t^{n+1}) - t^n] \Delta x} \int_{t^n}^{\inf(T, t^{n+1})} \int_{x_j}^{x_{j+1}} \eta(s, y) dy ds, \\ (u_\Delta)_j^n := \frac{1}{[\inf(T, t^{n+1}) - t^n] \Delta x} \int_{t^n}^{\inf(T, t^{n+1})} \int_{x_j}^{x_{j+1}} u(s, y) dy ds, \end{cases} \quad (n, j) \in [1, N] \times \mathbb{Z},$$

avec l'initialisation suivante

$$\begin{cases} (\eta_\Delta)_j^0 := \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} \eta_0(y) dy, \\ (u_\Delta)_j^0 := \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u_0(y) dy, \end{cases} \quad j \in \mathbb{Z}.$$

La solution du schéma numérique, défini ci-dessous, sera notée $(\eta_j^n, u_j^n)_{(n,j) \in [0, N] \times \mathbb{Z}}$ et les erreurs de consistance, définies ci-dessous aussi, notées $((\epsilon_1)_j^n, (\epsilon_2)_j^n)_{(n,j) \in [0, N] \times \mathbb{Z}}$. Finalement, nous définissons les deux erreurs de convergence par

$$\begin{cases} e_j^n = \eta_j^n - (\eta_\Delta)_j^n & \text{et} & (E^\pm)_j^n = e_j^{n+1} \pm e_j^n, \\ f_j^n = u_j^n - (u_\Delta)_j^n & \text{et} & (F^\pm)_j^n = f_j^{n+1} \pm f_j^n. \end{cases} \quad (4.3)$$

L'équivalent discret de l'énergie $\mathcal{E}_s^{abcd}$ (1.10) sera choisi tel que

$$\begin{aligned} \mathcal{E}(e, f) := & \|e\|_{\ell_\Delta^2}^2 + (b - c) \|D_+ e\|_{\ell_\Delta^2}^2 + b(-c) \|D_+ D_- e\|_{\ell_\Delta^2}^2 \\ & + \|f\|_{\ell_\Delta^2}^2 + (d - a) \|D_+ f\|_{\ell_\Delta^2}^2 + d(-a) \|D_+ D_- f\|_{\ell_\Delta^2}^2. \end{aligned} \quad (4.4)$$

Remarque 4.1. Nous pouvons remarquer sur le cas continu que la présence des termes b et/ou d permet de régulariser le système. En effet, l'équation devient dans ce cas une EDO

$$\partial_t \eta = (I - b\partial_x^2)^{-1} (I + a\partial_x^2) \partial_x u + (I - b\partial_x^2)^{-1} \partial_x (\eta u),$$

où $(I - b\partial_x^2)^{-1} \partial_x (\eta u)$ correspond à un produit de convolution. Cette distinction est conservée au niveau discret puisque notre étude est scindée en deux parties selon la valeur des paramètres b et d .

Section 4.1. Introduction (en français)

CAS $b > 0$ ET $d > 0$: Considérons $(\eta_0, u_0) \in H^{s_{bc}}(\mathbb{R}) \times H^{s_{ad}}(\mathbb{R})$ avec s suffisamment grand pour assurer une certaine régularité à la solution et s_{bc} , s_{ad} définis par (1.11). Le schéma numérique considéré est le suivant

$$\begin{cases} \frac{1}{\Delta t} (I - bD_+D_-) (\eta_j^{n+1} - \eta_j^n) + \frac{1}{2} (I + aD_+D_-) \left(D_c(u)_j^n + D_c(u)_j^{n+1} \right) + D_c(\eta u)_j^n = 0, \\ \frac{1}{\Delta t} (I - dD_+D_-) (u_j^{n+1} - u_j^n) + \frac{1}{2} (I + cD_+D_-) \left(D_c(\eta)_j^n + D_c(\eta)_j^{n+1} \right) + \frac{1}{2} D_c(u^2)_j^n = 0, \end{cases} \quad (4.5)$$

pour tout $n \geq 0$ et $j \in \mathbb{Z}$. Nous initialisons le schéma par la donnée initiale moyennée sur une maille en espace, comme expliqué en introduction générale,

$$\begin{cases} \eta_j^0 = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} \eta_0(y) dy, \\ u_j^0 = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u_0(y) dy, \end{cases} \quad j \in \mathbb{Z}. \quad (4.6)$$

Les erreurs de consistance sont dans ce cas définies par

$$\begin{cases} \frac{1}{\Delta t} (I - bD_+D_-) ((\eta_\Delta)_j^{n+1} - (\eta_\Delta)_j^n) + \frac{1}{2} (I + aD_+D_-) \left(D_c(u_\Delta)_j^n + D_c(u_\Delta)_j^{n+1} \right) \\ \quad + D_c(\eta_\Delta u_\Delta)_j^n =: (\epsilon_1)_j^n, \\ \frac{1}{\Delta t} (I - dD_+D_-) ((u_\Delta)_j^{n+1} - (u_\Delta)_j^n) + \frac{1}{2} (I + cD_+D_-) \left(D_c(\eta_\Delta)_j^n + D_c(\eta_\Delta)_j^{n+1} \right) \\ \quad + \frac{1}{2} D_c(u_\Delta^2)_j^n =: (\epsilon_2)_j^n, \end{cases} \quad (4.7)$$

où η_Δ et u_Δ sont définis par (4.2).

Remarque 4.2. Comme précisé en remarque 4.1, le fait d'avoir $b > 0$ et $d > 0$ assure, même au niveau discret, une certaine régularité et le schéma sera stable même sans le terme de viscosité de Rusanov $\frac{\tau}{2} \Delta x D_+D_- (\cdot)$.

Ce cas ($b > 0$ et $d > 0$) assure un contrôle des dérivées discrètes jusqu'à l'ordre deux et ainsi un contrôle de l'énergie discrète (4.4). Le raisonnement continu de [27] peut se reproduire à l'identique au niveau discret et nous permet de démontrer notre premier résultat de convergence, résumé dans le théorème suivant.

Théorème 4.3. Convergence du schéma dans le cas $b > 0$ et $d > 0$

Soit $a \leq 0, b > 0, c \leq 0$ et $d > 0$, considérons $s > 6, (\eta_0, u_0) \in H^{s_{bc}}(\mathbb{R}) \times H^{s_{ad}}(\mathbb{R})$ et $T > 0$ tel que $(\eta, u) \in \mathcal{C}([0, T], H^{s_{bc}}(\mathbb{R}) \times H^{s_{ad}}(\mathbb{R}))$ soit la solution sur $[0, T]$ du système (abcd) partant de la donnée initiale (η_0, u_0) . Soit $N \in \mathbb{N}$, il existe une constante ω_0 dépendant seulement de $\sup_{t \in [0, T]} \|(\eta(t), u(t))\|_{H^{s_{bc}} \times H^{s_{ad}}}$ telle que si le nombre de pas de temps N et le pas d'espace Δx sont choisis tels que

$$\Delta t = T/N \leq \omega_0 \quad \text{et} \quad \Delta x \leq \omega_0$$

et si nous considérons le schéma numérique (4.5) avec l'initialisation (4.6), alors ce schéma numérique est d'ordre un en temps et en espace. Plus précisément, les erreurs de convergence définies par (4.3) satisfont l'inégalité suivante

$$\sup_{n \in \llbracket 0, N \rrbracket} \mathcal{E}(e^n, f^n) \leq C_{abcd} \left\{ \Delta t^2 + \Delta x^2 \right\}, \quad (4.8)$$

où C_{abcd} dépend des paramètres a, b, c, d , de $\sup_{t \in [0, T]} \|(\eta(t), u(t))\|_{H^{s_{bc}} \times H^{s_{ad}}}$ et de T et $\mathcal{E}$ est définie par (4.4).

Remarque 4.4. Il n'y a donc pas de condition CFL nécessaire pour assurer la convergence lorsque $b > 0$ et $d > 0$.

CAS $bd = 0$: Comme pour le cas précédent, nous considérons $(\eta_0, u_0) \in H^{s_{bc}}(\mathbb{R}) \times H^{s_{ad}}(\mathbb{R})$ avec s suffisamment grand et s_{bc}, s_{ad} définis par (1.11). Le schéma numérique avec lequel nous travaillons est le suivant

$$\left\{ \begin{array}{l} \frac{1}{\Delta t} (I - bD_+ D_-) (\eta_j^{n+1} - \eta_j^n) + (I + aD_+ D_-) D_c (u_j^{n+1}) + D_c (\eta u)_j^n \\ \qquad \qquad \qquad = \frac{1}{2} (1 - \text{sgn}(b)) \tau_1 \Delta x D_+ D_- (\eta)_j^n, \\ \frac{1}{\Delta t} (I - dD_+ D_-) (u_j^{n+1} - u_j^n) + (I + cD_+ D_-) D_c (\eta)_j^{n+1} + \frac{1}{2} D_c (u^2)_j^n \\ \qquad \qquad \qquad = \frac{1}{2} (1 - \text{sgn}(d)) \tau_2 \Delta x D_+ D_- (u)_j^n, \end{array} \right. \quad (4.9)$$

avec $n \geq 0$ et $j \in \mathbb{Z}$ et la même condition initiale que précédemment, définie par (4.6). Les erreurs de consistance sont donc définies par

$$\left\{ \begin{array}{l} \frac{1}{\Delta t} (I - bD_+ D_-) ((\eta_\Delta)_j^{n+1} - (\eta_\Delta)_j^n) + (I + aD_+ D_-) D_c (u_\Delta)_j^{n+1} + D_c (\eta_\Delta u_\Delta)_j^n \\ \qquad \qquad \qquad = (\epsilon_1)_j^n + \frac{1}{2} (1 - \text{sgn}(b)) \tau_1 \Delta x D_+ D_- (\eta_\Delta)_j^n, \\ \frac{1}{\Delta t} (I - dD_+ D_-) ((u_\Delta)_j^{n+1} - (u_\Delta)_j^n) + (I + cD_+ D_-) D_c (\eta_\Delta)_j^{n+1} + \frac{1}{2} D_c (u_\Delta^2)_j^n \\ \qquad \qquad \qquad = (\epsilon_2)_j^n + \frac{1}{2} (1 - \text{sgn}(d)) \tau_2 \Delta x D_+ D_- (u_\Delta)_j^n, \end{array} \right. \quad (4.10)$$

où η_Δ et u_Δ sont toujours définis par (4.2). Comme pour l'équation de (KdV), les coefficients de Rusanov τ_1 et τ_2 ainsi que le pas de temps Δt pourraient dépendre de n . Cependant, pour plus de clarté, nous n'écrivons pas cette dépendance.

Pour ce cas, nous nous intéressons aux paramètres

$$a \leq 0, c \leq 0, b \geq 0, d \geq 0, \text{ avec } bd = 0$$

en excluant les cinq cas suivants :

$$\left\{ \begin{array}{l} a = b = 0, d > 0, c < 0, \\ a = d = 0, b > 0, c < 0, \\ b = d = 0, c < 0, a < 0, \\ a = b = c = d = 0, \\ a = b = d = 0, c < 0. \end{array} \right. \quad (4.11)$$

Remarque 4.5. De haut en bas, les deux premiers cas étaient déjà exclus dans l'étude continue de [27], voir (1.9). Les cas $b = d = 0, c < 0, a < 0$ et $a = b = c = d = 0$ semblent particulièrement délicats à cause du manque de symétrie de l'opérateur hyperbolique du système (abcd) (cependant, [97] pourrait consister un début de preuve). Pour finir, le cas $a = b = d = 0, c < 0$ est lui aussi particulier car l'énergie ne contrôle pas assez de dérivées discrètes sur η pour pouvoir borner la version discrète du terme non linéaire $\partial_x(\eta u)$ (cependant, il serait intéressant d'utiliser les récents résultats de [25] pour notre problème de convergence).

Le résultat de convergence devient dans ce cas

Théorème 4.6. *Convergence du schéma dans le cas $bd = 0$*

Soit $a \leq 0, b \geq 0, c \leq 0$ et $d \geq 0$ avec $bd = 0$ en excluant les cas (4.11), considérons $s > 9, (\eta_0, u_0) \in H^{s_{bc}}(\mathbb{R}) \times H^{s_{ad}}(\mathbb{R})$ et $T > 0$ tel que $(\eta, u) \in \mathcal{C}([0, T], H^{s_{bc}}(\mathbb{R}) \times H^{s_{ad}}(\mathbb{R}))$ soit la solution sur $[0, T]$ du système (abcd) partant de la donnée initiale (η_0, u_0) . Choisissons $\alpha_0 > 0, \tau_1 > 0$ et $\tau_2 > 0$ tels que

$$\|u_\Delta^n\|_{\ell^\infty} + \alpha_0 \leq \tau_1 \quad \text{et} \quad \|u_\Delta^n\|_{\ell^\infty} + \alpha_0 \leq \tau_2, \quad n \in \llbracket 0, N \rrbracket.$$

Soit $\beta_0 \in]0, 1[$, il existe une constante $\omega_0 > 0$ (dépendant de $\alpha_0, \beta_0, \tau_1, \tau_2$ et $\sup_{t \in [0, T]} \|(\eta(t), u(t))\|_{H^{s_{bc}} \times H^{s_{ad}}}$)

telle que si le nombre de pas de temps $N \in \mathbb{N}$ et le pas d'espace Δx sont choisis tels que

$$\Delta t = T/N \leq \omega_0 \quad \text{et} \quad \Delta x \leq \omega_0$$

et

$$\max\{(1 - \operatorname{sgn}(b)) \tau_1, (1 - \operatorname{sgn}(d)) \tau_2\} \Delta t \leq (1 - \beta_0) \Delta x, \quad (4.12)$$

si nous considérons le schéma numérique (4.9) avec les viscosités τ_1 et τ_2 , choisies précédemment ainsi que l'initialisation (4.6), alors ce schéma numérique est d'ordre un en temps et en espace. Plus précisément, les erreurs de convergence définies par (4.3) satisfont l'inégalité suivante

$$\sup_{n \in \llbracket 0, N \rrbracket} \mathcal{E}(e^n, f^n) \leq C_{abcd} \Delta x^2,$$

où C_{abcd} dépend des paramètres a, b, c, d , de $\sup_{t \in [0, T]} \|(\eta(t), u(t))\|_{H^{s_{bc}} \times H^{s_{ad}}}$ et de T .

Remarque 4.7. Dans le cas où l'un des paramètres b et/ou d est nul, une condition CFL de type hyperbolique (4.12) est nécessaire.



Méthode suivie. Nous adaptons la méthode développée sur l'équation de (KdV) au chapitre 3 afin de prouver les deux principaux théorèmes précédents. Cette méthode consiste à dériver une inégalité d'entropie discrète entre la solution numérique $(v_j^n)_{(n,j)}$ et la solution exacte moyennée u_Δ . Pour ce faire, nous introduisons un opérateur dispersif, voir (3.46),

$$\mathcal{A} := I + \theta \Delta t D_+ D_+ D_-$$

et montrons par récurrence que sous condition CFL, nous obtenons bien cet équivalent discret recherché, équation (3.58) :

$$\|\mathcal{A}e^{n+1}\|_{\ell_\Delta^2}^2 \leq \|\mathcal{A}e^n\|_{\ell_\Delta^2}^2 \{1 + \Delta t C\} + C \Delta t \|e^n\|_{\ell_\Delta^2}^2.$$

L'utilisation du lemme de Grönwall 2.29 permet de boucler la récurrence et de prouver l'ordre de convergence. Revenons quelques instants sur cet opérateur dispersif, nous avons montré qu'il permettait de contrôler l'énergie ℓ_Δ^2 à la Proposition 3.36

$$\|\mathcal{A}e\|_{\ell_\Delta^2}^2 = \|e\|_{\ell_\Delta^2}^2 + \theta \Delta t \Delta x \|D_+ D_-(e)\|_{\ell_\Delta^2}^2 + \theta^2 \Delta t^2 \|D_+ D_+ D_-(e)\|_{\ell_\Delta^2}^2.$$

Cet opérateur dispersif est donc constitué de l'énergie $\|\cdot\|_{\ell^2}$ (terme d'ordre 0) et de termes d'ordre 2 en Δt^2 et $\Delta x \Delta t$. Seule l'énergie (d'ordre 0) nous intéresse pour prouver l'ordre de convergence et nous utilisons la majoration $\|\cdot\|_{\ell^2}^2 \leq \|\mathcal{A} \cdot\|_{\ell_\Delta^2}^2$ pour conclure.

Généralisons ce processus dans un premier temps à l'équation de (KdV-BBM), comme cela sera développé en section 4.2. L'opérateur dispersif est naturellement modifié pour prendre en compte la dispersion de type BBM et devient dans ce cas, voir équation (4.14),

$$\mathcal{A}^{a,b} := (I - b D_+ D_-) + a \theta \Delta t D_+ D_+ D_-.$$

D'après l'équation (4.17) que nous rappelons ci-dessous,

$$\begin{aligned} \|\mathcal{A}^{a,b}e\|_{\ell_\Delta^2}^2 &= \|e\|_{\ell_\Delta^2}^2 + b^2 \|D_+ D_-(e)\|_{\ell_\Delta^2}^2 + a^2 \Delta t^2 \theta^2 \|D_+ D_+ D_-(e)\|_{\ell_\Delta^2}^2 + 2b \|D_+(e)\|_{\ell_\Delta^2}^2 \\ &\quad + a \Delta t \theta \Delta x \|D_+ D_-(e)\|_{\ell_\Delta^2}^2 + ab \Delta t \theta \Delta x \|D_+ D_+ D_-(e)\|_{\ell_\Delta^2}^2, \end{aligned}$$

cet opérateur permet de contrôler l'énergie (termes d'ordre 0)

$$\|\cdot\|_{\ell_\Delta^2}^2 + 2b \|D_+(\cdot)\|_{\ell_\Delta^2}^2 + b^2 \|D_+ D_-(\cdot)\|_{\ell_\Delta^2}^2.$$

Remarque 4.8. Nous retrouvons bien le cas de l'énergie ℓ_Δ^2 dans le cas de (KdV) i.e $b = 0$.

Passons maintenant au système (abcd). L'énergie du système correspond à $\mathcal{E}$ définie en (4.4) et l'opérateur dispersif est donc beaucoup plus volumineux. Nous travaillerons donc directement avec l'énergie $\mathcal{E}$ et chercherons une estimation d'énergie semblable à

$$(1 - C\Delta t)\mathcal{E}(e^{n+1}, f^{n+1}) \leq (1 + C\Delta t)\mathcal{E}(e^n, f^n) + C\Delta t\|\epsilon_1^n\|_{\ell_\Delta^2}^2 + C\Delta t\|\epsilon_2^n\|_{\ell_\Delta^2}^2.$$

Selon les paramètres a, b, c, d , le terme $\|\epsilon_1^n\|_{\ell_\Delta^2}^2$ peut être modifié en $\|\epsilon_1^n\|_{\ell_\Delta^2}^2 + \|D_+(\epsilon_1)^n\|_{\ell_\Delta^2}^2$ (c'est le cas par exemple quand $cd \neq 0$ ou $c \neq 0, b = d = 0$), idem pour ϵ_2^n (par exemple, pour les cas $ab \neq 0$ ou $a \neq 0, b = d = 0$). La conclusion de la preuve de convergence suit les mêmes étapes clés que pour (KdV) (au chapitre 3) : l'utilisation du lemme de Grönwall combiné avec un raisonnement par récurrence. Cependant, nous utiliserons une version «plus forte» du lemme de Grönwall discret : la version rappelée au lemme 4.27. Nous synthétisons les grandes étapes de la preuve de convergence dans le cas du système (abcd) ci-après.

Étapes clés : preuve de convergence pour (abcd)

- Étude de la consistance
 - Obtention d'une estimation d'énergie sous condition CFL hyperbolique
 - Inégalité finale par le lemme de Grönwall «version forte»
-

Par la suite, nous détaillerons surtout l'étape d'obtention d'une estimation d'énergie pour l'énergie $\mathcal{E}$.



Organisation du chapitre. La première section, section 4.2 est consacrée à une étude préliminaire sur (KdV-BBM) avant d'étudier par la suite le système (abcd). Nous discutons en sous-section 4.2.1.2 de ce que deviennent ces résultats pour le cas d'un terme de dispersion de type KdV négatif ou dans le cas d'un schéma totalement centré en sous-section 4.2.2. Cela constitue un premier pas vers le système (abcd), pour lequel nous avons choisi un schéma centré pour discrétiser les termes de dispersion.

Le reste du chapitre est consacré, quant à lui, au système (abcd) et contient les preuves des principaux résultats : les théorèmes 4.3 et 4.6.

La section 4.3 regroupe la preuve (complète) du théorème 4.3 dans le cas $bd > 0$. Comme expliqué précédemment, dans ce cas établir l'estimation d'énergie pour l'erreur de convergence peut être fait en suivant à la lettre l'approche de [27] pour le système continu, c'est l'objet de la sous-section 4.3.1. La seconde étape clé de la preuve (avec la récurrence et le lemme fort de Grönwall) est détaillée en sous-section 4.3.2.

La section 4.4 regroupe la preuve du théorème 4.6 dans le cas $bd = 0$. Seule l'étape de dérivation d'estimées d'énergie est détaillée. Un premier résultat technique sur la partie hyperbolique du système est démontrée en sous section 4.4.1. Il s'agit essentiellement d'une réécriture d'inégalités prouvées en filigrane au chapitre 3 sur (KdV). Nous détaillons ensuite en sous-sections 4.4.2, 4.4.3 et 4.4.4 les preuves d'énergie selon un éventail de choix possibles pour a, b, c et d .

Pour finir, et dans le but de valider nos résultats théoriques concernant l'ordre de convergence, nous comparons les solutions numériques à certaines solutions exactes de type solitons, établies par Chen dans [30] ou [14] en section 4.5.

Remarque 4.9. L'objectif de ce chapitre n'est pas de quantifier l'ordre de convergence en fonction de la régularité, comme nous avons pu le faire au chapitre 3. Nous nous focalisons plutôt sur les estimations d'énergie sur $\mathcal{E}$ (étape 2). Nous supposons donc le temps T suffisamment petit pour assurer l'existence d'une solution exacte régulière pour toutes les configurations des paramètres a, b, c et d considérés.



4.2 One step towards the *abcd* system : the KdV-BBM equation

We deal with a generalization of (KdV) equation with an additional dispersion term of BBM-type : the Korteweg-de Vries-Benjamin-Bona-Mahony equation (KdV-BBM) which combines both a BBM-type dispersive operator $\partial_x^2 \partial_t$ and a KdV-type operator ∂_x^3

$$(I - b\partial_x^2)\partial_t u + \partial_x u + a\partial_x^3 u + \partial_x \left(\frac{u^2}{2} \right) = 0, \quad (\text{KdV-BBM})$$

where $a \in \mathbb{R}$ and $b \geq 0$. Contrary to the (KdV) equation studied in Chapter 3, the KdV dispersive coefficient a is not signed.

We adapt the numerical scheme of Chapter 3 to this equation and derive CFL condition stability and rate of convergence. As expected, only a hyperbolic CFL is need with a non zero BBM dispersion and the dispersive "Airy" CFL is necessary (according the θ -scheme used) when the BBM dispersion cancels. The scheme is first order accurate both in space and time.

Remark 4.10. *It is physically relevant to study the case of a negative dispersion in Korteweg-de Vries equation. Indeed, before normalize to one the coefficient in KdV equation, the dispersive coefficient is equal to*

$$\frac{1}{2} \left(\frac{1}{3} - Bo \right) \mu$$

with $Bo > 0$ the Bond number and $\mu > 0$ the shallowness parameter, as seen in (1.4). Therefore, a negative dispersion corresponds to a large Bond number ($Bo > \frac{1}{3}$) i.e a significant surface tension.

Remark 4.11. *In order to distinguish both cases with a negative or positive dispersion, we denote (KdV⁻) the KdV type equation with a negative dispersion*

$$\partial_t u + \partial_x \left(\frac{u^2}{2} \right) + a\partial_x^3 u = 0, \quad \text{with } a < 0. \quad (\text{KdV}^-)$$

When the dispersion parameter a is non negative, we still denote (KdV), the corresponding equation, which would correspond to a (KdV⁺) equation.

It is relatively easy to switch from one sign of dispersion to an other. To this end, we define the following change of variables

$$v(t, x) := -u(t, -x), \quad (t, x) \in \mathbb{R}^2.$$

The function u is solution of (KdV) if and only if v is solution of (KdV⁻).

The study in this section has a double aim : on the one hand, we seek to get a better understanding of the discretized BBM-dispersive operator (and as we will see, proving convergence results becomes much more easy with the BBM dispersive b term than for the KdV equation, in Chapter 3). On the other hand, it can be seen as a shy attempt towards attacking the resolution of the *abcd* system (detailed in Sections 4.3 and 4.4).

Remark 4.12. *To more closely mimic the *abcd* system we add an advection term in the KdV-BBM equation, not present in the study of KdV in Chapter 3 for instance.*

4.2.1 Left and right winded numerical schemes

We first work with left winded (respectively right winded) scheme according to the sign of the KdV dispersive coefficient a . Indeed, the decentralization of the scheme depends on the sign of the dispersion : $a \geq 0$ corresponds to the studied cases so far and is linked with a right winded scheme, whereas $a < 0$ corresponds to (KdV⁻) and is linked with a left winded numerical scheme.

4.2.1.1 The case of a non negative dispersion ($a \geq 0$)

We discretize (KdV-BBM) according to the numerical scheme, for any $(n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}$

$$(I - bD_+D_-)e_j^{n+1} + a\Delta t\theta D_+D_+D_-e_j^{n+1} = (I - bD_+D_-)e_j^n - \Delta tD_c(e)_j^n - \Delta tD_c(eu_\Delta)_j^n - \Delta tD_c\left(\frac{e^2}{2}\right)_j^n - a(1-\theta)\Delta tD_+D_+D_-e_j^n + \frac{\tau}{2}\Delta t\Delta xD_+D_-e_j^n - \Delta t\epsilon_j^n. \quad (4.13)$$

As $a \geq 0$, we have adapted the discretization used for KdV and chosen a center discretization both for the BBM dispersive b term and the advection term $D_c(e)_j^n$.

Theorem 4.13.

Choose $\alpha_0 > 0$ and $\beta_0 \in (0, 1)$, there exists $\omega_0 > 0$ such that

$$\Delta t = T/N \leq \omega_0, \quad \Delta x \leq \omega_0,$$

and

• if $b \neq 0$: $a\Delta t(1 - 2\theta) \leq b\Delta x$,

$$\bullet \text{ if } b = 0 : \begin{cases} \left[\tau + \frac{1}{2}\mathbb{1}_{\{a \neq 0\}} \right] \Delta t \leq (1 - \beta_0)\Delta x, \text{ (if } a = 0 \text{ we may take } \tau\Delta t \leq \Delta x) \\ \|1 + u_\Delta^n\|_{\ell^\infty} + \alpha_0 \leq \tau, \\ a\Delta t(1 - 2\theta) \leq \frac{\Delta x^3}{4}(1 - \beta_0). \end{cases}$$

Then, the numerical scheme (4.13) with the Rusanov coefficient τ and the time step Δt and the space size Δx chosen as previously is of first order in ℓ_Δ^2 -norm.

Remark 4.14. • Both cases $a = 0$ and $a > 0$ are included in the previous CFL condition.

• When $b \neq 0$, we can choose $\tau = 0$ because we have no minoration condition on τ . Indeed, the term $-bD_+D_-$ brings enough control.

Remark 4.15. The case $b = 0$ is not detailed here, because it corresponds more or less to the KdV equation (with the added advection term $\Delta tD_c(e)_j^n$). The proof is so similar to the one of KdV + advection (if $a \neq 0$) or Burgers + advection (if $a = 0$), treated in Chapter 3.

We are looking for the case $b \neq 0$ here, which corresponds to a simpler case than KdV equation because of the laplacian (coming from the BBM dispersive operator $\partial_x^2 \partial_t u$) in the equation.

The discrete relative entropy inequality

We will follow essentially the same guidelines than the ones developed for KdV equation in Subsection 3.4.2 and Section 3.5. First, we change a little the definition of dispersive operators $\mathcal{A}$ and $\mathcal{B}$ (Equations (3.46) and (3.47)) to take into account the BBM dispersive term : bD_+D_-

Definition 4.16. Let $(e_j)_{j \in \mathbb{Z}}$ a sequence in $\ell_\Delta^2(\mathbb{Z})$, we define the dispersive operators

$$\left(\mathcal{A}^{a,b}e\right)_j := (I - bD_+D_-)e_j + a\theta\Delta tD_+D_+D_-(e)_j \quad (4.14)$$

and

$$\left(\mathcal{B}^{a,b}e\right)_j := (I - bD_+D_-)e_j - a(1 - \theta)\Delta tD_+D_+D_-e_j. \quad (4.15)$$

The operator I corresponds to identity operator.

We are now able to state the discrete inequality entropy for the (KdV-BBM) equation.

Section 4.2. One step towards the abcd system : the KdV-BBM equation

Proposition 4.17. Let $(e_j^n)_{(j,n) \in \mathbb{Z} \times \llbracket 0, N \rrbracket}$ be the convergence error defined by (4.3) with respect to Scheme (4.13). For every $\theta \in [0, 1]$, $\Delta t > 0$, $\Delta x > 0$, for all $n \in \llbracket 0, N \rrbracket$, one has

$$\begin{aligned} \left\| \left(\mathcal{A}^{a,b} e \right)^{n+1} \right\|_{\ell_\Delta^2} &\leq \left\| \left(\mathcal{B}^{a,b} e \right)^n \right\|_{\ell_\Delta^2} + \Delta t \left(\|D_c(e)^n\|_{\ell_\Delta^2} (\|1 + u_\Delta^n\|_{\ell^\infty} + \|e^n\|_{\ell^\infty}) \right. \\ &\quad \left. + \|e^n\|_{\ell_\Delta^2} \|D_c(u_\Delta)^n\|_{\ell^\infty} \right) + \Delta t \|e^n\|_{\ell_\Delta^2}. \end{aligned} \quad (4.16)$$

Proof. We recall that $\tau = 0$ when $b \neq 0$.

The numerical scheme rewrites with dispersive operators $\mathcal{A}^{a,b}$ and $\mathcal{B}^{a,b}$

$$\left(\mathcal{A}^{a,b} e \right)_j^{n+1} = \left(\mathcal{B}^{a,b} e \right)_j^n - \Delta t D_c(e)_j^n - \Delta t D_c(e u_\Delta)_j^n - \Delta t D_c \left(\frac{e^2}{2} \right)_j^n - \Delta t \epsilon_j^n.$$

By a triangular inequality, one has

$$\left\| \left(\mathcal{A}^{a,b} e \right)^{n+1} \right\|_{\ell_\Delta^2} \leq \left\| \left(\mathcal{B}^{a,b} e \right)^n \right\|_{\ell_\Delta^2} + \Delta t \left\| D_c \left(\left(1 + u_\Delta + \frac{e}{2} \right) e \right)^n \right\|_{\ell_\Delta^2} + \Delta t \|e^n\|_{\ell_\Delta^2}.$$

However, we can develop the middle term thanks to Relation (3.20)

$$\begin{aligned} &\|D_c \left(\left(1 + u_\Delta + \frac{e}{2} \right) e \right)^n\|_{\ell_\Delta^2} \\ &= \|D_c(e)^n + D_c(u_\Delta)^n (\mathcal{S}^1 e)^n + D_c(e)^n (\mathcal{S}^{-1} u_\Delta)^n + D_c(e)^n \frac{(\mathcal{S}^1 e)^n}{2} + D_c(e)^n \frac{(\mathcal{S}^{-1} e)^n}{2}\|_{\ell_\Delta^2} \\ &\leq \|D_c(e)^n\|_{\ell_\Delta^2} (\|1 + u_\Delta^n\|_{\ell^\infty} + \|e^n\|_{\ell^\infty}) + \|e^n\|_{\ell_\Delta^2} \|D_c(u_\Delta)^n\|_{\ell^\infty}. \end{aligned}$$

The shift operator $\mathcal{S}^\ell$ (with $\ell \in \{-1, 1\}$) has already been introduced in Subsection 2.2.1.3. □

Properties of discrete dispersive operators $\mathcal{A}^{a,b}$ and $\mathcal{B}^{a,b}$

First and foremost, the dispersive operators $\mathcal{A}^{a,b}$ and $\mathcal{B}^{a,b}$ are able to control the ℓ_Δ^2 -norm, as follows.

Proposition 4.18. Let $(e)_{j \in \mathbb{Z}}$ be a ℓ_Δ^2 -sequence. For a positive BBM dispersion ($b > 0$) one has

$$\|e\|_{\ell_\Delta^2} \leq \|\mathcal{A}^{a,b} e\|_{\ell_\Delta^2} \quad \text{and} \quad \|D_+(e)\|_{\ell_\Delta^2} \leq \frac{1}{\sqrt{2b}} \|\mathcal{A}^{a,b} e\|_{\ell_\Delta^2}.$$

Remark 4.19. Since $\|D_c(e)\|_{\ell_\Delta^2} \leq \|D_+(e)\|_{\ell_\Delta^2}$, then $\frac{1}{\sqrt{2b}} \|\mathcal{A}^{a,b} e\|_{\ell_\Delta^2}$ is also an upper bound for the central derivative operator D_c .

Proof. We compute the ℓ_Δ^2 -norm of $(\mathcal{A}^{a,b} e)_j$

$$\begin{aligned} \|\mathcal{A}^{a,b} e\|_{\ell_\Delta^2}^2 &= \|e\|_{\ell_\Delta^2}^2 + b^2 \|D_+ D_-(e)\|_{\ell_\Delta^2}^2 + a^2 \Delta t^2 \theta^2 \|D_+ D_+ D_-(e)\|_{\ell_\Delta^2}^2 - 2b \sum_{j \in \mathbb{Z}} \Delta x e_j D_+ D_-(e)_j \\ &\quad + 2a \Delta t \theta \sum_{j \in \mathbb{Z}} \Delta x e_j D_+ D_+ D_-(e)_j - 2ab \Delta t \theta \sum_{j \in \mathbb{Z}} \Delta x D_+ D_+ D_-(e)_j D_+ D_-(e)_j. \end{aligned}$$

The discrete integration by parts (3.27) together with Relation (3.29) give

$$\begin{aligned} \|\mathcal{A}^{a,b} e\|_{\ell_\Delta^2}^2 &= \|e\|_{\ell_\Delta^2}^2 + b^2 \|D_+ D_-(e)\|_{\ell_\Delta^2}^2 + a^2 \Delta t^2 \theta^2 \|D_+ D_+ D_-(e)\|_{\ell_\Delta^2}^2 + 2b \|D_+(e)\|_{\ell_\Delta^2}^2 \\ &\quad + a \Delta t \theta \Delta x \|D_+ D_-(e)\|_{\ell_\Delta^2}^2 + ab \Delta t \theta \Delta x \|D_+ D_+ D_-(e)\|_{\ell_\Delta^2}^2. \end{aligned} \quad (4.17)$$

Proposition 4.18 is a straightforward consequence (we recall that a and b are non negative). □

Then, we are able to associate $\mathcal{B}^{a,b}$ to $\mathcal{A}^{a,b}$ under a dispersive CFL (exactly as in Proposition 3.37)

Proposition 4.20. *Under the CFL condition*

$$a\Delta t(1 - 2\theta) \leq b\Delta x, \quad (4.18)$$

then, for any sequence $(e_j)_{j \in \mathbb{Z}}$

$$\|\mathcal{B}^{a,b}e\|_{\ell_\Delta^2}^2 \leq \|\mathcal{A}^{a,b}e\|_{\ell_\Delta^2}^2.$$

Proof. First of all, we compute the ℓ_Δ^2 -norm of $\mathcal{B}^{a,b}e$, which gives us

$$\begin{aligned} \|\mathcal{B}^{a,b}e\|_{\ell_\Delta^2}^2 &= \|(I - bD_+D_-)e\|_{\ell_\Delta^2}^2 - 2a(1 - \theta)\Delta t \sum_{j \in \mathbb{Z}} \Delta x (I - bD_+D_-)e_j D_+D_+D_-(e)_j \\ &\quad + a^2(1 - \theta)^2\Delta t^2 \|D_+D_+D_-(e)\|_{\ell_\Delta^2}^2. \end{aligned} \quad (4.19)$$

By discrete integrations by parts (Relations (3.27) and (3.29)), it holds

$$\begin{aligned} -2a(1 - \theta)\Delta t \sum_{j \in \mathbb{Z}} \Delta x (I - bD_+D_-)e_j D_+D_+D_-(e)_j &= -a(1 - \theta)\Delta t \Delta x \|D_+D_-(e)\|_{\ell_\Delta^2}^2 \\ &\quad - ab(1 - \theta)\Delta t \Delta x \|D_+D_+D_-(e)\|_{\ell_\Delta^2}^2. \end{aligned}$$

By developing $\|(I - bD_+D_-)e\|_{\ell_\Delta^2}^2$, one has

$$\begin{aligned} \|\mathcal{B}^{a,b}e\|_{\ell_\Delta^2}^2 &= \|e\|_{\ell_\Delta^2}^2 + 2b\|D_+(e)\|_{\ell_\Delta^2}^2 \\ &\quad + \|D_+D_-(e)\|_{\ell_\Delta^2}^2 \{b^2 + a\theta\Delta t\Delta x - a\Delta t\Delta x\} \\ &\quad + \|D_+D_+D_-(e)\|_{\ell_\Delta^2}^2 \{ab\theta\Delta t\Delta x + a^2\theta^2\Delta t^2 + a^2(1 - 2\theta)\Delta t^2 - ab\Delta t\Delta x\}. \end{aligned}$$

The CFL condition $a(1 - 2\theta)\Delta t \leq b\Delta x$, the non negativity of a and Identity (4.17) imply

$$\|\mathcal{B}^{a,b}e\|_{\ell_\Delta^2}^2 \leq \|\mathcal{A}^{a,b}e\|_{\ell_\Delta^2}^2.$$

□

Convergence of the scheme

We prove Theorem 4.13 in this paragraph. We details only the final discrete entropy inequality, since the guidelines are the same as those of KdV equation : Grönwall Lemma 2.29 coupled with an induction hypothesis on the smallness of e^n .

Proposition 4.18 transforms Equation (4.16) into

$$\|\mathcal{A}^{a,b}e^{n+1}\|_{\ell_\Delta^2}^2 \leq \|\mathcal{B}^{a,b}e^n\|_{\ell_\Delta^2}^2 + \frac{\Delta t}{\sqrt{2b}} \|\mathcal{A}^{a,b}e^n\|_{\ell_\Delta^2} \left(\|1 + u_\Delta^n\|_{\ell^\infty} + \|e^n\|_{\ell^\infty} + \sqrt{2b} \|D_c u_\Delta^n\|_{\ell^\infty} \right) + \Delta t \|\epsilon^n\|_{\ell_\Delta^2}^2.$$

Young's inequality gives

$$\begin{aligned} \|\mathcal{A}^{a,b}e^{n+1}\|_{\ell_\Delta^2}^2 &\leq \|\mathcal{B}^{a,b}e^n\|_{\ell_\Delta^2}^2 (1 + 2\Delta t) + \frac{\Delta t}{2b} \left(\|1 + u_\Delta^n\|_{\ell^\infty} + \|e^n\|_{\ell^\infty} + \sqrt{2b} \|D_c u_\Delta^n\|_{\ell^\infty} \right)^2 (2 + \Delta t) \|\mathcal{A}^{a,b}e^n\|_{\ell_\Delta^2}^2 \\ &\quad + \Delta t \|\epsilon^n\|_{\ell_\Delta^2}^2 (2 + \Delta t). \end{aligned}$$

Proposition 4.20 enables us to conclude

$$\begin{aligned} \|\mathcal{A}^{a,b}e^{n+1}\|_{\ell_\Delta^2}^2 &\leq \|\mathcal{A}^{a,b}e^n\|_{\ell_\Delta^2}^2 \left[1 + \Delta t \left(2 + \frac{1}{2b} (\|1 + u_\Delta^n\|_{\ell^\infty} + \|e^n\|_{\ell^\infty} \right. \right. \\ &\quad \left. \left. + \sqrt{2b}\|D_c u_\Delta^n\|_{\ell^\infty} \right)^2 (2 + \Delta t) \right] + \Delta t \|e^n\|_{\ell_\Delta^2}^2 (2 + \Delta t). \end{aligned}$$

The conclusion comes from the Grönwall Lemma 2.29 together with an induction hypothesis $\|e^n\|_{\ell^\infty} \leq C\Delta x^{\frac{1}{2}-\gamma}$, with $\gamma \in (0, \frac{1}{2})$.

4.2.1.2 The case of a negative dispersion ($a < 0$)

We have to deal with the case $a < 0$ in order to better fit to $abcd$ system. Indeed, according to the general introduction, the $abcd$ system is well-posed only in two cases (1.7)-(1.8). We have chosen to deal with only the first configuration : $a \leq 0, c \leq 0, b \geq 0$ and $d \geq 0$. For that case, we have to use a left winded dispersion *i.e.* $D_+D_-D_-$ instead of $D_+D_+D_-$. The numerical scheme writes then

$$\begin{aligned} (I - bD_+D_-)e_j^{n+1} + a\Delta t\theta D_+D_-D_-(e)_j^{n+1} &= (I - bD_+D_-)e_j^n - \Delta t D_c(e)_j^n - \Delta t D_c(eu_\Delta)_j^n \\ &\quad - \Delta t D_c\left(\frac{e^2}{2}\right)_j^n - a(1 - \theta)\Delta t D_+D_-D_-(e)_j^n + \frac{\tau}{2}\Delta t \Delta x D_+D_-(e)_j^n - \Delta t \epsilon_j^n. \end{aligned} \quad (4.20)$$

The proof follows exactly the same guidelines (this is why we do not detail it) and the conclusions are the same provided the CFL becomes

- if $b \neq 0$: $|a|\Delta t(1 - 2\theta) \leq b\Delta x$,
- if $b = 0$: $\begin{cases} \left[\tau + \frac{1}{2}\mathbb{1}_{\{a \neq 0\}} \right] \Delta t \leq \Delta x(1 - \beta_0), \\ \|1 + u_\Delta^n\|_{\ell^\infty} + \alpha_0 \leq \tau, \\ |a|\Delta t(1 - 2\theta) \leq \frac{\Delta x^3}{4}(1 - \beta_0). \end{cases}$

Remark 4.21. By defining $f_j^n := -e_{-j}^n$, the sequence $(f_j^n)_{j \in \mathbb{Z}}$ satisfies the non negative dispersion scheme (4.13) with $-a$ instead of a and $-1 - (u_\Delta)_j^n$ instead of the advection speed $1 + (u_\Delta)_j^n$ provided $(e_j^n)_{j \in \mathbb{Z}}$ satisfies (4.20). Switching from a positive dispersion to a negative one is then obvious provided we choose the well decentralization for the discrete dispersive operator. This change of variables exactly corresponds to the discrete equivalent of the continuous change of variables to switch from (KdV) to (KdV⁻).

4.2.2 For a centered scheme

Since we will use a centered scheme for the $abcd$ system, a first step is to focus now on the centered scheme for the KdV-BBM equation :

$$\begin{aligned} (I - bD_+D_-)e_j^{n+1} + a\theta\Delta t D_+D_-D_c(e)_j^{n+1} &= (I - bD_+D_-)e_j^n - \Delta t D_c(e)_j^n - \Delta t D_c(eu_\Delta)_j^n \\ &\quad - \Delta t D_c\left(\frac{e^2}{2}\right)_j^n - a(1 - \theta)\Delta t D_+D_-D_c(e)_j^n + \frac{\tau}{2}\Delta t \Delta x D_+D_-(e)_j^n - \Delta t \epsilon_j^n. \end{aligned} \quad (4.21)$$

Since the scheme is centered, we can consider either $a \geq 0$ or $a < 0$ without changing the strategy of the proof and the scheme.

Theorem 4.22.

Choose $\alpha_0 > 0$, $\beta_0 \in (0, 1)$, there exists $\omega_0 > 0$ such that

$$\Delta t \leq \omega_0, \quad \Delta x \leq \omega_0,$$

and

- if $b \neq 0$: $|a|(1 - 2\theta) \leq 0$,
- if $b = 0$: $\begin{cases} \left[\tau + \frac{1}{2} \mathbb{1}_{\{a \neq 0\}} \right] \Delta t \leq \Delta x(1 - \beta_0), \\ \|1 + u_\Delta^n\|_{\ell^\infty} + \alpha_0 \leq \tau, \\ |a|(1 - 2\theta) \leq 0. \end{cases}$

Choose the numerical scheme (4.21) with the Rusanov coefficient τ , Δt and Δx as previously. Then the numerical dispersive operator

$$\|\mathcal{A}_c^{ab} \cdot\|_{\ell_\Delta^2}^2 := \|\cdot\|_{\ell_\Delta^2}^2 + 2b\|D_+(\cdot)\|_{\ell_\Delta^2}^2 + b^2\|D_+D_-(\cdot)\|_{\ell_\Delta^2}^2 + a^2\Delta t^2\theta^2\|D_+D_-D_c(\cdot)\|_{\ell_\Delta^2}^2$$

satisfies

$$\|\mathcal{A}_c^{ab} e^n\|_{\ell_\Delta^2} \leq C(\Delta t + \Delta x),$$

with C independent of Δt , Δx and e^n . Thus, the numerical scheme (4.21) is of first order in space.

Proof. The proof is adapted from Chapter 3 for $b = 0$ and from Subection 4.2.1.1 for $b \neq 0$. $\square$

The $abcd$ system

We are interested to derive discrete energy estimates for the $abcd$ system, by following the main steps used for KdV in Chapter 3 or KdV-BBM equations. The discrete dispersive operator $\mathcal{A}^{abcd}$ (equivalent respectively of $\mathcal{A}$ in Chapter 3 or $\mathcal{A}^{a,b}$ of Section 4.2) is not used on behalf of its length and we study directly the energy $\mathcal{E}$ instead of $\mathcal{A}^{abcd}$ for simplicity. It corresponds to the discrete equivalent of the conserved continuous energy $\mathcal{E}_s^{abcd}$ (1.10) of [27]. A distinction is made between two cases according to the BBM dispersive terms of the system.

When $b > 0$ and $d > 0$, both equations have a BBM - type operator and we consider the numerical scheme (4.5), with a Crank-Nicolson discretization for the KdV dispersive term.

When $bd = 0$, in certain respects, we have a limited degree of freedom and we used an implicit KdV dispersive term, Scheme (4.9).

Remark 4.23. *The discretization used to design the numerical schemes (Crank-Nicolson scheme or implicit scheme) are chosen in order to ensure a not too restrictive CFL condition, only a hyperbolic one is needed as we will see later.*

4.3 Non zero BBM-dispersive term in both equations : $abcd$ system with $b, d > 0$

We see from numerical scheme (4.5) that the equations governing the convergence error (e_j^n, f_j^n) are the following :

$$\begin{cases} (I - bD_+D_-)(E^-)_j^n + \frac{\Delta t}{2}(I + aD_+D_-)D_c(F^+)_j^n + \Delta t D_c(eu_\Delta)_j^n \\ \quad + \Delta t D_c(\eta_\Delta f)_j^n + \Delta t D_c(ef)_j^n = -\Delta t(\epsilon_1)_j^n, \\ (I - dD_+D_-)(F^-)_j^n + \frac{\Delta t}{2}(I + cD_+D_-)D_c(E^+)_j^n + \Delta t D_c(fu_\Delta)_j^n \\ \quad + \frac{\Delta t}{2}D_c(f^2)_j^n = -\Delta t(\epsilon_2)_j^n. \end{cases} \quad (4.22)$$

We recall the main result of this section.

Section 4.3. Non zero BBM-dispersive term in both equations : $abcd$ system with $b, d > 0$

Theorem 4.24. *Convergence rate in the case $b > 0, d > 0$*

Let $a \leq 0, b > 0, c \leq 0$ and $d > 0$. Consider $s > 6, (\eta_0, u_0) \in H^{s_{bc}}(\mathbb{R}) \times H^{s_{ad}}(\mathbb{R})$ and $T > 0$ such that $(\eta, u) \in \mathcal{C}([0, T], H^{s_{bc}}(\mathbb{R}) \times H^{s_{ad}}(\mathbb{R}))$ is the solution on $[0, T]$ of the system (abcd). Let $N \in \mathbb{N}$, there exists a positive constant ω_0 depending on $\sup_{t \in [0, T]} \|(\eta(t), u(t))\|_{H^{s_{bc}} \times H^{s_{ad}}}$ such that if the number of time steps N and the space discretization step Δx are chosen such that

$$\Delta t = \frac{T}{N} \leq \omega_0 \text{ and } \Delta x \leq \omega_0,$$

if we consider the numerical scheme (4.5) along with the initial data (4.6) as well as the approximation $((\eta_\Delta)_j^n, (u_\Delta)_j^n)_{(n,j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}}$ defined by (4.2), then, the numerical scheme (4.5) is first order convergent i.e. the convergence error defined in (4.3) satisfies :

$$\sup_{n \in \llbracket 0, N \rrbracket} \mathcal{E}(e^n, f^n) \leq C_{abcd} \left\{ \Delta t^2 + \Delta x^2 \right\}, \quad (4.23)$$

where C_{abcd} depends on the parameters a, b, c, d , on $\sup_{t \in [0, T]} \|(\eta(t), u(t))\|_{H^{s_{bc}} \times H^{s_{ad}}}$ and on T and $\mathcal{E}$ is defined by (4.4).

Remark 4.25. *As we have chosen the center scheme for the space discretization, we have even the second order convergence rate in space*

$$\sup_{n \in \llbracket 0, N \rrbracket} \mathcal{E}(e^n, f^n) \leq C_{abcd} \left\{ \Delta t^2 + \Delta x^4 \right\}.$$

4.3.1 Discrete energy inequality

Proposition 4.26. *There exist two constants B_1 and B_2 depending on a, b, c, d and on the ℓ^∞ -norm of $((\eta_\Delta)_j^n, (u_\Delta)_j^n)$ and $(D_c(\eta_\Delta)_j^n, D_c(u_\Delta)_j^n)$ and proportional to $\max\{\|e^n\|_{\ell^\infty}, \|f^n\|_{\ell^\infty}\}$, such that, for all $(n, j) \in \llbracket 0, N-1 \rrbracket \times \mathbb{Z}$*

$$\mathcal{E}(e^{n+1}, f^{n+1}) \leq 2\Delta t \|\epsilon^n\|_{\ell_\Delta^2}^2 + (1 + \Delta t B_1) \mathcal{E}(e^n, f^n) + \Delta t B_2 \mathcal{E}(e^{n+1}, f^{n+1}), \quad (4.24)$$

with (e_j^n, f_j^n) the two convergence errors defined by (4.3) and $\|\epsilon^n\|_{\ell_\Delta^2} = \max\{\|(\epsilon_1)^n\|_{\ell_\Delta^2}, \|(\epsilon_2)^n\|_{\ell_\Delta^2}\}$.

Proof. As we announced in the introduction, we will establish energy estimates imitating the approach from the continuous case. Let $n \in \llbracket 0, N \rrbracket$ and observe that by multiplying the first equation of (4.22) by $(I + cD_+D_-)(E^+)_j^n$, the second by $(I + aD_+D_-)(F^+)_j^n$ and adding up the results, we find that :

$$\|\text{LHS}^{abcd}\|_{\ell_\Delta^2}^2 = \|\text{RHS}^{abcd}\|_{\ell_\Delta^2}^2 \quad (4.25)$$

with

$$\begin{aligned} \|\text{LHS}^{abcd}\|_{\ell_\Delta^2}^2 := & \sum_{j \in \mathbb{Z}} \Delta x (I - bD_+D_-)(E^-)_j^n (I + cD_+D_-)(E^+)_j^n \\ & + \sum_{j \in \mathbb{Z}} \Delta x (I - dD_+D_-)(F^-)_j^n (I + aD_+D_-)(F^+)_j^n \\ & + \frac{\Delta t}{2} \sum_{j \in \mathbb{Z}} \Delta x \left[(I + aD_+D_-)D_c(F^+)_j^n (I + cD_+D_-)(E^+)_j^n \right. \\ & \left. + (I + cD_+D_-)D_c(E^+)_j^n (I + aD_+D_-)(F^+)_j^n \right] \end{aligned}$$

and

$$\begin{aligned}
 ||\text{RHS}^{abcd}||_{\ell_\Delta^2}^2 := & -\Delta t \Delta x \sum_{j \in \mathbb{Z}} \left[D_c (eu_\Delta)_j^n + D_c (\eta_\Delta f)_j^n \right] (I + cD_+ D_-) (E^+)_j^n \\
 & - \Delta t \Delta x \sum_{j \in \mathbb{Z}} D_c (fu_\Delta)_j^n (I + aD_+ D_-) (F^+)_j^n \\
 & - \Delta t \Delta x \sum_{j \in \mathbb{Z}} D_c (ef)_j^n (I + cD_+ D_-) (E^+)_j^n \\
 & - \frac{\Delta t}{2} \Delta x \sum_{j \in \mathbb{Z}} D_c (f^2)_j^n (I + aD_+ D_-) (F^+)_j^n \\
 & - \Delta t \Delta x \sum_{j \in \mathbb{Z}} \left[(\epsilon_1)_j^n (I + cD_+ D_-) (E^+)_j^n + (\epsilon_2)_j^n (I + aD_+ D_-) (F^+)_j^n \right].
 \end{aligned}$$

Left hand side $||\text{LHS}^{abcd}||_{\ell_\Delta^2}^2$. Notice that

$$\sum_{j \in \mathbb{Z}} \Delta x \mathcal{L} (E^-)_j^n \mathcal{L} (E^+)_j^n = ||\mathcal{L} e^{n+1}||_{\ell_\Delta^2}^2 - ||\mathcal{L} e^n||_{\ell_\Delta^2}^2, \quad (4.26)$$

with $\mathcal{L}$ any linear operator (for instance, $\mathcal{L} = D_\bullet^\ell$ a derivative operator of order ℓ , as in our case).

• With this in mind, developing $\sum_{j \in \mathbb{Z}} \Delta x (I - bD_+ D_-) (E^-)_j^n (I + cD_+ D_-) (E^+)_j^n$ holds

$$\begin{aligned}
 \sum_{j \in \mathbb{Z}} \Delta x (I - bD_+ D_-) (E^-)_j^n (I + cD_+ D_-) (E^+)_j^n = & ||e^{n+1}||_{\ell_\Delta^2}^2 - ||e^n||_{\ell_\Delta^2}^2 - \sum_{j \in \mathbb{Z}} \Delta x b D_+ D_- (E^-)_j^n (E^+)_j^n \\
 & + \sum_{j \in \mathbb{Z}} \Delta x c D_+ D_- (E^+)_j^n (E^-)_j^n + b(-c) \left[||D_+ D_-(e)^{n+1}||_{\ell_\Delta^2}^2 - ||D_+ D_-(e)^n||_{\ell_\Delta^2}^2 \right].
 \end{aligned}$$

Relation (3.27) together with (4.26) enable us to simplify

$$\begin{aligned}
 \sum_{j \in \mathbb{Z}} \Delta x (I - bD_+ D_-) (E^-)_j^n (I + cD_+ D_-) (E^+)_j^n = & ||e^{n+1}||_{\ell_\Delta^2}^2 - ||e^n||_{\ell_\Delta^2}^2 \\
 & + (b - c) \left[||D_+(e)^{n+1}||_{\ell_\Delta^2}^2 - ||D_+(e)^n||_{\ell_\Delta^2}^2 \right] + b(-c) \left[||D_+ D_-(e)^{n+1}||_{\ell_\Delta^2}^2 - ||D_+ D_-(e)^n||_{\ell_\Delta^2}^2 \right].
 \end{aligned} \quad (4.27)$$

• Similarly, we obtain

$$\begin{aligned}
 \sum_{j \in \mathbb{Z}} \Delta x (I - dD_+ D_-) (F^-)_j^n (I + aD_+ D_-) (F^+)_j^n = & ||f^{n+1}||_{\ell_\Delta^2}^2 - ||f^n||_{\ell_\Delta^2}^2 \\
 & + (d - a) \left[||D_+(f)^{n+1}||_{\ell_\Delta^2}^2 - ||D_+(f)^n||_{\ell_\Delta^2}^2 \right] + d(-a) \left[||D_+ D_-(f)^{n+1}||_{\ell_\Delta^2}^2 - ||D_+ D_-(f)^n||_{\ell_\Delta^2}^2 \right].
 \end{aligned} \quad (4.28)$$

• Both operators D_c and $I + cD_+ D_-$ commute, thus, by (3.27), it becomes

$$\begin{aligned}
 & \frac{\Delta t}{2} \sum_{j \in \mathbb{Z}} \Delta x \left[(I + aD_+ D_-) D_c (F^+)_j^n (I + cD_+ D_-) (E^+)_j^n \right. \\
 & \quad \left. + (I + cD_+ D_-) D_c (E^+)_j^n (I + aD_+ D_-) (F^+)_j^n \right] \\
 = & \frac{\Delta t}{2} \sum_{j \in \mathbb{Z}} \Delta x \left[-(I + aD_+ D_-) (F^+)_j^n (I + cD_+ D_-) D_c (E^+)_j^n \right. \\
 & \quad \left. + (I + cD_+ D_-) D_c (E^+)_j^n (I + aD_+ D_-) (F^+)_j^n \right] \\
 = & 0.
 \end{aligned}$$

Right hand side $||\text{RHS}^{abcd}||_{\ell_\Delta^2}^2$. We treat the five terms which occur in the right hand side of (4.25).

• Using (3.20), we first write

$$\begin{aligned}
 \Delta t \left\| D_c (eu_\Delta)_j^n \right\|_{\ell_\Delta^2} &= \Delta t \left\| \mathcal{S}^{-1} e^n D_c(u_\Delta)^n + \mathcal{S}^1 u_\Delta^n D_c(e)^n \right\|_{\ell_\Delta^2} \\
 &\leq \Delta t \|D_c(u_\Delta)^n\|_{\ell^\infty} \left\| \mathcal{S}^{-1} e^n \right\|_{\ell_\Delta^2} + \Delta t \left\| \mathcal{S}^1 u_\Delta^n \right\|_{\ell^\infty} \|D_c(e)^n\|_{\ell_\Delta^2} \\
 &\leq \Delta t \max \left\{ 1, \frac{1}{\sqrt{b-c}} \right\} \left(\|D_c(u_\Delta)^n\|_{\ell^\infty} \|e^n\|_{\ell_\Delta^2} + \|u_\Delta^n\|_{\ell^\infty} \sqrt{b-c} \|D_c(e)^n\|_{\ell_\Delta^2} \right) \\
 &\leq \Delta t \max \{ \|u_\Delta^n\|_{\ell^\infty}, \|D_c(u_\Delta)^n\|_{\ell^\infty} \} \max \left\{ 1, \frac{1}{\sqrt{b-c}} \right\} \left(\|e^n\|_{\ell_\Delta^2} + \sqrt{b-c} \|D_c(e)^n\|_{\ell_\Delta^2} \right) \\
 &\leq \Delta t \max \{ \|u_\Delta^n\|_{\ell^\infty}, \|D_c(u_\Delta)^n\|_{\ell^\infty} \} \max \left\{ 1, \frac{1}{\sqrt{b-c}} \right\} 2\sqrt{\mathcal{E}(e^n, f^n)}.
 \end{aligned}$$

Proceeding in a similar fashion with the other terms we arrive, thanks to the Cauchy-Schwarz inequality, at (we recall that $c \leq 0$)

$$\begin{aligned}
 -\Delta t \Delta x \sum_{j \in \mathbb{Z}} \left[D_c (eu_\Delta)_j^n + D_c (\eta_\Delta f)_j^n \right] (I + cD_+ D_-) (E^+)_j^n \\
 \leq \Delta t \left(\|D_c (eu_\Delta)^n\|_{\ell_\Delta^2} + \|D_c (\eta_\Delta f)^n\|_{\ell_\Delta^2} \right) \left(\left\| (E^+)^n \right\|_{\ell_\Delta^2} - c \left\| D_+ D_- (E^+)^n \right\|_{\ell_\Delta^2} \right).
 \end{aligned}$$

By Equation (4.3) on E^+ , one has

$$\begin{aligned}
 \left\| (E^+)^n \right\|_{\ell_\Delta^2} - c \left\| D_+ D_- (E^+)^n \right\|_{\ell_\Delta^2} &\leq \|e^{n+1}\|_{\ell_\Delta^2} + \|e^n\|_{\ell_\Delta^2} - c \|D_+ D_- (e)^{n+1}\|_{\ell_\Delta^2} - c \|D_+ D_- (e)^n\|_{\ell_\Delta^2} \\
 &\leq \max \left\{ 1, \sqrt{-\frac{c}{b}} \right\} 2 \left[\sqrt{\mathcal{E}(e^{n+1}, f^{n+1})} + \sqrt{\mathcal{E}(e^n, f^n)} \right].
 \end{aligned}$$

These together give

$$\begin{aligned}
 -\Delta t \Delta x \sum_{j \in \mathbb{Z}} \left[D_c (eu_\Delta)_j^n + D_c (\eta_\Delta f)_j^n \right] (I + cD_+ D_-) (E^+)_j^n \\
 \leq \Delta t C_{1,1} \mathcal{E}(e^n, f^n) + \Delta t C_{2,1} \mathcal{E}(e^{n+1}, f^{n+1}), \quad (4.29)
 \end{aligned}$$

where $C_{1,1}$ and $C_{2,1}$ can be written, for example, as a numerical constants multiplied with :

$$\begin{aligned}
 \max \left\{ 1, \sqrt{-\frac{c}{b}} \right\} \left[\max \{ \|u_\Delta^n\|_{\ell^\infty}, \|D_c(u_\Delta)^n\|_{\ell^\infty} \} \max \left\{ 1, \frac{1}{\sqrt{b-c}} \right\} \right. \\
 \left. + \max \{ \|\eta_\Delta^n\|_{\ell^\infty}, \|D_c(\eta_\Delta)^n\|_{\ell^\infty} \} \max \left\{ 1, \frac{1}{\sqrt{d-a}} \right\} \right]. \quad (4.30)
 \end{aligned}$$

• We treat $-\Delta t \Delta x \sum_{j \in \mathbb{Z}} D_c (fu_\Delta)_j^n (I + aD_+ D_-) (F^+)_j^n$ in same spirit as above in order to obtain that

$$-\Delta t \Delta x \sum_{j \in \mathbb{Z}} D_c (fu_\Delta)_j^n (I + aD_+ D_-) (F^+)_j^n \leq \Delta t C_{1,2} \mathcal{E}(e^n, f^n) + \Delta t C_{2,2} \mathcal{E}(e^{n+1}, f^{n+1}), \quad (4.31)$$

where $C_{1,2}$ and $C_{2,2}$ are multiples of :

$$\max \{ \|u_\Delta^n\|_{\ell^\infty}, \|D_c(u_\Delta)^n\|_{\ell^\infty} \} \max \left\{ 1, \frac{1}{\sqrt{d-a}} \right\} \max \left\{ 1, \sqrt{-\frac{a}{d}} \right\}. \quad (4.32)$$

• In order to treat other terms, we first observe that

$$\begin{aligned}
 \|D_c (ef)^n\|_{\ell_\Delta^2} &= \left\| \mathcal{S}^{-1} e^n D_c(f)^n + \mathcal{S}^1 f^n D_c(e)^n \right\|_{\ell_\Delta^2} \leq \|e^n\|_{\ell^\infty} \|D_c(f)^n\|_{\ell_\Delta^2} + \|f^n\|_{\ell^\infty} \|D_c(e)^n\|_{\ell_\Delta^2} \\
 &\leq \max \left\{ \frac{1}{\sqrt{b-c}}, \frac{1}{\sqrt{d-a}} \right\} \left(\|e^n\|_{\ell^\infty} \sqrt{d-a} \|D_c(f)^n\|_{\ell_\Delta^2} + \|f^n\|_{\ell^\infty} \sqrt{b-c} \|D_c(e)^n\|_{\ell_\Delta^2} \right).
 \end{aligned}$$

Thus, we obtain

$$-\Delta t \Delta x \sum_{j \in \mathbb{Z}} D_c (ef)_j^n (I + cD_+ D_-) (E^+)_j^n \leq \Delta t C_{1,3} \mathcal{E} (e^n, f^n) + \Delta t C_{2,3} \mathcal{E} (e^{n+1}, f^{n+1}), \quad (4.33)$$

where $C_{1,3}$ respectively $C_{2,3}$ are proportional with

$$\max \left\{ \frac{1}{\sqrt{b-c}}, \frac{1}{\sqrt{d-a}} \right\} \max \left\{ 1, \sqrt{-\frac{c}{b}} \right\} \max \{ \|e^n\|_{\ell^\infty}, \|f^n\|_{\ell^\infty} \}. \quad (4.34)$$

• The same holds for $-\frac{\Delta t}{2} \Delta x \sum_{j \in \mathbb{Z}} D_c (f^2)_j^n (I + aD_+ D_-) (F^+)_j^n$, namely, we get that

$$-\frac{\Delta t}{2} \Delta x \sum_{j \in \mathbb{Z}} D_c (f^2)_j^n (I + aD_+ D_-) (F^+)_j^n \leq \Delta t C_{1,4} \mathcal{E} (e^n, f^n) + \Delta t C_{2,4} \mathcal{E} (e^{n+1}, f^{n+1}), \quad (4.35)$$

where $C_{1,4}$ respectively $C_{2,4}$ are proportional with

$$\frac{1}{\sqrt{d-a}} \max \left\{ 1, \sqrt{-\frac{a}{d}} \right\} \|f^n\|_{\ell^\infty}. \quad (4.36)$$

• Let us treat the last term. We denote

$$\|\epsilon^n\|_{\ell_\Delta^2} = \max \left\{ \|\epsilon_1^n\|_{\ell_\Delta^2}, \|\epsilon_2^n\|_{\ell_\Delta^2} \right\},$$

and we write that, thanks to Cauchy-Schwarz inequality (we recall that $c \leq 0$ and $a \leq 0$)

$$\begin{aligned} & -\Delta t \Delta x \sum_{j \in \mathbb{Z}} \left[(\epsilon_1)_j^n (I + cD_+ D_-) (E^+)_j^n + (\epsilon_2)_j^n (I + aD_+ D_-) (F^+)_j^n \right] \\ & \leq \Delta t \|\epsilon_1^n\|_{\ell_\Delta^2} \left(\|(E^+)^n\|_{\ell_\Delta^2} - c \|D_+ D_- (E^+)^n\|_{\ell_\Delta^2} \right) + \Delta t \|\epsilon_2^n\|_{\ell_\Delta^2} \left(\|(F^+)^n\|_{\ell_\Delta^2} - a \|D_+ D_- (F^+)^n\|_{\ell_\Delta^2} \right). \end{aligned}$$

By applied Young inequality, we recover the ℓ_Δ^2 -norm of e^n and f^n .

$$\begin{aligned} & -\Delta t \Delta x \sum_{j \in \mathbb{Z}} \left[(\epsilon_1)_j^n (I + cD_+ D_-) (E^+)_j^n + (\epsilon_2)_j^n (I + aD_+ D_-) (F^+)_j^n \right] \\ & \leq 2\Delta t \|\epsilon^n\|_{\ell_\Delta^2}^2 + \Delta t \left(\|e^n\|_{\ell_\Delta^2}^2 + \|f^n\|_{\ell_\Delta^2}^2 + c^2 \|D_+ D_- e^n\|_{\ell_\Delta^2}^2 + a^2 \|D_+ D_- f^n\|_{\ell_\Delta^2}^2 \right) \\ & \quad + \Delta t \left(\|e^{n+1}\|_{\ell_\Delta^2}^2 + \|f^{n+1}\|_{\ell_\Delta^2}^2 + c^2 \|D_+ D_- e^{n+1}\|_{\ell_\Delta^2}^2 + a^2 \|D_+ D_- f^{n+1}\|_{\ell_\Delta^2}^2 \right) \\ & \leq 2\Delta t \|\epsilon^n\|_{\ell_\Delta^2}^2 + \Delta t \max \left\{ 1, \frac{-c}{b}, \frac{-a}{d} \right\} \mathcal{E} (e^n, f^n) + \Delta t \max \left\{ 1, \frac{-c}{b}, \frac{-a}{d} \right\} \mathcal{E} (e^{n+1}, f^{n+1}). \quad (4.37) \end{aligned}$$

Gathering the informations from (4.25), (4.27), (4.28), (4.29), (4.31), (4.33), (4.35) and (4.37), we obtain the existence of two constants B_1 and B_2 that depend on a, b, c, d , and on the ℓ^∞ -norm of $((\eta_\Delta)_j^n, (u_\Delta)_j^n)$ and $(D_c(\eta_\Delta)_j^n, D_c(u_\Delta)_j^n)$ (dependence which we can track using relations (4.30), (4.32), (4.34) respectively (4.36)) and that are proportional to $\max \{ \|e^n\|_{\ell^\infty}, \|f^n\|_{\ell^\infty} \}$ such that

$$\mathcal{E} (e^{n+1}, f^{n+1}) - \mathcal{E} (e^n, f^n) \leq 2\Delta t \|\epsilon^n\|_{\ell_\Delta^2}^2 + \Delta t B_1 \mathcal{E} (e^n, f^n) + \Delta t B_2 \mathcal{E} (e^{n+1}, f^{n+1})$$

which rewrites as Equation (4.24). □

4.3.2 Convergence of the scheme

We begin this section with a stronger discrete version of Grönwall's lemma, a proof of which can be found in [32] :

Lemma 4.27. *Let $(u^n)_{n \in \mathbb{N}}, (v^n)_{n \in \mathbb{N}}, (w^n)_{j \in \mathbb{N}}$ be sequences of real numbers with $w^n \geq 0$ for all $n \in \mathbb{N}$ which satisfies*

$$u^n \leq v^n + \sum_{k=0}^{n-1} w^k u^k,$$

for all $n \in \mathbb{N}$. Then, for any $n \in \mathbb{N}$, we have :

$$u^n \leq \max_{k \in \llbracket 1, n \rrbracket} v^k \prod_{j=0}^{n-1} (1 + w^j).$$

This subsection is devoted to the proof of Theorem 4.24.

Proof of Theorem 4.24. Exactly as for the KdV case, we remove the dependency of $\|e^n\|_{\ell^\infty}$ and $\|f^n\|_{\ell^\infty}$ in Proposition 4.26 by induction. Let us arbitrary fix $n \in \llbracket 0, N-1 \rrbracket$. Suppose the strong induction hypothesis

$$\|e^k\|_{\ell^\infty} \leq 1 \quad \text{and} \quad \|f^k\|_{\ell^\infty} \leq 1, \quad (4.38)$$

for all $k \in \llbracket 0, n \rrbracket$.

This is obvious true for $n = 0$, since $e_j^0 = f_j^0 = 0$, for all $j \in \mathbb{Z}$. Let us prove that $\|e^{n+1}\|_{\ell^\infty} \leq 1$ and $\|f^{n+1}\|_{\ell^\infty} \leq 1$.

Inequality (4.24) is thus available for any $k \in \llbracket 0, n \rrbracket$ and constants B_1 and B_2 may be upper bounded by C_1 and C_2 independent of $\|e^k\|_{\ell^\infty}$ and $\|f^k\|_{\ell^\infty}$. One has, for all $k \in \llbracket 0, n \rrbracket$

$$\mathcal{E}(e^{k+1}, f^{k+1}) \leq \frac{2\Delta t}{1 - \Delta t C_2} \|\epsilon^k\|_{\ell_\Delta^2}^2 + \left(1 + \frac{\Delta t (C_1 + C_2)}{1 - \Delta t C_2}\right) \mathcal{E}(e^k, f^k).$$

Namely, it becomes

$$\mathcal{E}(e^{k+1}, f^{k+1}) - \mathcal{E}(e^k, f^k) \leq \frac{2\Delta t}{1 - \Delta t C_2} \|\epsilon^k\|_{\ell_\Delta^2}^2 + \frac{\Delta t (C_1 + C_2)}{1 - \Delta t C_2} \mathcal{E}(e^k, f^k).$$

Thus, taking the sum of all these inequalities and noticing that $e_j^0 = f_j^0 = 0$, we end up with

$$\mathcal{E}(e^{n+1}, f^{n+1}) \leq \frac{2(n+1)\Delta t}{1 - \Delta t C_2} \max_{k \in \llbracket 0, n \rrbracket} \|\epsilon^k\|_{\ell_\Delta^2}^2 + \frac{\Delta t (C_1 + C_2)}{1 - \Delta t C_2} \sum_{k=0}^n \mathcal{E}(e^k, f^k).$$

Applying the strong discrete Grönwall inequality, Lemma 4.27, and using the fact that the consistency error is order one accurate in time and space, see Section A.2, we get

$$\begin{aligned} \mathcal{E}(e^{n+1}, f^{n+1}) &\leq \frac{2(n+1)\Delta t}{1 - \Delta t C_2} \exp\left(\frac{(n+1)\Delta t (C_1 + C_2)}{1 - \Delta t C_2}\right) \max_{k \in \llbracket 0, n \rrbracket} \|\epsilon^k\|_{\ell_\Delta^2}^2 \\ &\leq \frac{TC(\eta_0, u_0)}{1 - \Delta t C_2} \exp\left(\frac{T(C_1 + C_2)}{1 - \Delta t C_2}\right) \left\{(\Delta x)^2 + (\Delta t)^2\right\}, \end{aligned}$$

where $C(\eta_0, u_0)$ is some constant depending on the initial data (η_0, u_0) . Thus, for $\Delta x, \Delta t$ small enough and using the inequality :

$$\|e^{n+1}\|_{\ell^\infty} \leq C \|e^{n+1}\|_{\ell_\Delta^2}^{\frac{1}{2}} \|D_+(e)^{n+1}\|_{\ell_\Delta^2}^{\frac{1}{2}},$$

with C a constant, we get that for sufficiently small Δx and Δt :

$$\|e^{n+1}\|_{\ell^\infty} \leq 1 \text{ and } \|f^{n+1}\|_{\ell^\infty} \leq 1. \quad (4.39)$$

We can assure that the inductive hypothesis (4.38) holds for all $k \in \llbracket 0, n+1 \rrbracket$. Obviously, this allows to close the estimates and to provide the desired bound.

□

Remark 4.28. In reality, the Taylor expansion of Section A.2 could be developed a step further to cancel terms in factor of Δx^2 and Δx^3 , thanks to the use of a center scheme. It enables us to recover the second order in space

$$\mathcal{E}(e^{n+1}, f^{n+1}) \leq \frac{TC(\eta_0, u_0)}{1 - \Delta t C_2} \exp\left(\frac{T(C_1 + C_2)}{1 - \Delta t C_2}\right) \{(\Delta x)^4 + (\Delta t)^2\},$$

cf Remark 4.25.

4.4 KdV equation coupled with BBM equation : $abcd$ system with $bd = 0$

As announced in the introduction, in this section we aim at providing a proof for our second main result. We recall that we will treat the case where the parameters verify :

$$a \leq 0, c \leq 0, b \geq 0, d \geq 0, \text{ with } bd = 0,$$

excluding the five cases :

$$\left\{ \begin{array}{l} a = b = 0, d > 0, c < 0, \\ a = d = 0, b > 0, c < 0, \\ b = d = 0, c < 0, a < 0, \\ a = b = c = d = 0, \\ a = b = d = 0, c < 0. \end{array} \right. \quad (4.40)$$

Theorem 4.29. Convergence rate in the case $bd = 0$

Let $a \leq 0, b \geq 0, c \leq 0$ and $d \geq 0$ with $bd = 0$, excluding the cases (4.40). Consider $s > 9$, $(\eta_0, u_0) \in H^{s_{bc}}(\mathbb{R}) \times H^{s_{ad}}(\mathbb{R})$ and $T > 0$ such that $(\eta, u) \in \mathcal{C}([0, T], H^{s_{bc}} \times H^{s_{ad}})$ is the solution on $[0, T]$ of the system (abcd). Choose $\alpha_0 > 0, \tau_1 > 0, \tau_2 > 0$ such that :

$$\|u_\Delta^n\|_{\ell^\infty} + \alpha_0 \leq \tau_1 \text{ and } \|u_\Delta^n\|_{\ell^\infty} + \alpha_0 \leq \tau_2, \quad n \in \mathbb{N}.$$

Let $\beta_0 \in (0, 1)$, there exists $\omega_0 > 0$ (depending on $\alpha_0, \beta_0, \tau_1, \tau_2$ and on $\sup_{t \in [0, T]} \|(\eta(t), u(t))\|_{H^{s_{bc}} \times H^{s_{ad}}}$) such that if the number of time steps $N \in \mathbb{N}$ and the space discretization step Δx are chosen in order to verify

$$\Delta t = T/N \leq \omega_0, \Delta x \leq \omega_0,$$

and

$$\max\{(1 - \text{sgn}(b)) \tau_1, (1 - \text{sgn}(d)) \tau_2\} \Delta t \leq (1 - \beta_0) \Delta x,$$

if we consider the numerical scheme (4.9) along with the initial data (4.6) and the numerical viscosities τ_1 and τ_2 as well as the approximation $\left((\eta_\Delta)_j^n, (u_\Delta)_j^n\right)_{(n,j) \in \llbracket 1, N \rrbracket \times \mathbb{Z}}$ defined by (4.2), then the numerical scheme (4.9) is first order convergent i.e. the convergence error defined in (4.3) satisfies :

$$\sup_{n \in \llbracket 0, N \rrbracket} \mathcal{E}(e^n, f^n) \leq C_{abcd} (\Delta x)^2,$$

where C_{abcd} depends on the parameters a, b, c, d , on $\sup_{t \in [0, T]} \|(\eta(t), u(t))\|_{H^{s_{bc}} \times H^{s_{ad}}}$ and on T .

Section 4.4. KdV equation coupled with BBM equation : $abcd$ system with $bd = 0$

Remark 4.30. In the first case, $b > 0$ and $d > 0$, the discrete energy $\mathcal{E}$ corresponds to a H^2 -discrete norm on both u and η , if a and c are non zero. In the case where $b = 0$ or $d = 0$, it corresponds to a H^1 -discrete norm and possibly even a L^2 -discrete norm according to the value of parameters a and c .

As opposed to the previous result, the proof of Theorem 4.29 is rather sensitive to the different values of the a, b, c, d parameters.

In this section, we only detail the derivation of the energy inequality for $\mathcal{E}$ (the equivalent of the discrete entropy inequality of KdV equation (3.37)). This inequality is summarized as follows.

Proposition 4.31. Assume $\|e^n\|_{\ell^\infty}, \|f^n\|_{\ell^\infty}, |\operatorname{sgn}(a)| \|D_+(f)^n\|_{\ell^\infty} \leq \Delta x^{\frac{1}{2}-\gamma}$, with $\gamma \in (0, \frac{1}{2})$ and $|\operatorname{sgn}(cd)| \|D_+ D_-(f)^n\|_{\ell^\infty} \leq 1$, for all $n \in \llbracket 0, N \rrbracket$. Then the following energy estimate holds true

$$\begin{aligned} (1 - \max \{ \operatorname{sgn}(b), \operatorname{sgn}(d) \} C \Delta t) \mathcal{E} (e^{n+1}, f^{n+1}) &\leq (1 + C \Delta t) \mathcal{E} (e^n, f^n) \\ &+ (\Delta t + \max \{ |\operatorname{sgn}(c)|, 1 - \operatorname{sgn}(b) \} \Delta t^2) C \|\epsilon_1^n\|_{\ell_\Delta^2}^2 \\ &+ (\Delta t + \max \{ |\operatorname{sgn}(a)|, 1 - \operatorname{sgn}(d) \} \Delta t^2) C \|\epsilon_2^n\|_{\ell_\Delta^2}^2 \\ &+ (\Delta t + \Delta t^2) \max \{ |\operatorname{sgn}(cd)|, |\operatorname{sgn}(c)| (1 - \max \{ \operatorname{sgn}(b), \operatorname{sgn}(d) \}) \} C \|D_+ (\epsilon_1)^n\|_{\ell_\Delta^2}^2 \\ &+ (\Delta t + \Delta t^2) \max \{ |\operatorname{sgn}(ab)|, |\operatorname{sgn}(a)| (1 - \max \{ \operatorname{sgn}(b), \operatorname{sgn}(d) \}) \} C \|D_+ (\epsilon_2)^n\|_{\ell_\Delta^2}^2. \end{aligned} \quad (4.41)$$

with $C > 0$ depending on $\|\eta_\Delta^n\|_{\ell^\infty}, \|D_c(\eta_\Delta)^n\|_{\ell^\infty}, \|u_\Delta^n\|_{\ell^\infty}, \|D_c(u_\Delta)^n\|_{\ell^\infty}$ and $\|D_+ D_c(u_\Delta)^n\|_{\ell^\infty}$.

In order to close the estimates and ensure the convergence proof (as the one made in Subsection 4.3.2, for the case $b > 0$ and $d > 0$) we perform as usual an induction hypothesis on the smallness of $\|e^n\|_{\ell^\infty}$ and/or $\|f^n\|_{\ell^\infty}$ and/or $\|D_+(f)^n\|_{\ell^\infty}$ according the cases. It is sufficient to assume by induction

$$\|e^n\|_{\ell^\infty}, \|f^n\|_{\ell^\infty}, |\operatorname{sgn}(a)| \|D_+(f)^n\|_{\ell^\infty} \leq \Delta x^{\frac{1}{2}-\gamma}, \quad \text{with } \gamma \in (0, \frac{1}{2}). \quad (4.42)$$

Hypothesis (4.42) is sufficient to assure the hypothesis of Proposition 4.33, Proposition 4.34 and Proposition 4.31. The energy estimate (4.41) is thus satisfied and the convergence rate (Theorem 4.29) is a consequence of the discrete strong Grönwall inequality, Lemma 4.27 (all the previous guidelines are detailed in Subsection 4.3.2 for the case $b > 0$ and $d > 0$).

The proofs of Proposition 4.31 are strongly based on inequalities implicitly proved for Burgers terms of KdV equation in Chapter 3 and recalled in the following subsection.

4.4.1 Burgers-type estimates

Let us state the first consequences coming from Burgers estimates.

Proposition 4.32. Let $(u_j)_{j \in \mathbb{Z}} \in \ell^\infty(\mathbb{Z})$, such that $\|D_+(u)\|_{\ell^\infty(\mathbb{Z})} < +\infty$ and $\lambda \geq 0$ respectively $\beta \in (0, 1)$. Fix $\alpha > 0$ and τ such that

$$\|u\|_{\ell^\infty} + \alpha \leq \tau.$$

Then, there exists a sufficiently small positive number ε_0 such that the following holds true. Consider two positive reals $\Delta t, \Delta x$ such that

$$\frac{\tau \Delta t}{\Delta x} \leq 1 - \beta, \quad \Delta x \leq \varepsilon_0$$

and $(e_j)_{j \in \mathbb{Z}} \in \ell^2(\mathbb{Z})$, such that

$$\lambda \|e\|_{\ell^\infty} \leq \Delta x^{\frac{1}{2}-\gamma}, \quad \text{with } \gamma \in (0, \frac{1}{2}).$$

Then, there exists a positive constant C depending on the ℓ^∞ -norms of $(u_j)_{j \in \mathbb{Z}}$ and $(D_+(u)_j)_{j \in \mathbb{Z}}$ such that

$$\left\| e - \Delta t D_c \left(\left(u + \lambda \frac{e}{2} \right) e \right) + \frac{\tau}{2} \Delta x \Delta t D_+ D_- (e) \right\|_{\ell_\Delta^2} \leq (1 + C \Delta t) \|e\|_{\ell_\Delta^2}.$$

Proof. We define

$$\mathcal{B}^{Burgers} e_j := e_j - \Delta t D_c \left(e \left(u + \lambda \frac{e}{2} \right) \right)_j + \frac{\tau}{2} \Delta x \Delta t D_+ D_- (e)_j.$$

We compute the ℓ_Δ^2 -norm :

$$\begin{aligned} \|\mathcal{B}^{Burgers} e\|_{\ell_\Delta^2}^2 &= \|e\|_{\ell_\Delta^2}^2 + \sum_{j \in \mathbb{Z}} \Delta x \Delta t^2 \lambda^2 \left[D_c \left(\frac{e^2}{2} \right)_j \right]^2 + \sum_{j \in \mathbb{Z}} \Delta x \Delta t^2 \left[D_c (ue)_j \right]^2 \\ &+ \sum_{j \in \mathbb{Z}} \frac{\tau^2 \Delta t^2 \Delta x^3}{4} [D_+ D_- (e)_j]^2 - \sum_{j \in \mathbb{Z}} 2 \Delta x \Delta t \lambda e_j D_c \left(\frac{e^2}{2} \right)_j - \sum_{j \in \mathbb{Z}} 2 \Delta x \Delta t e_j D_c (ue)_j \\ &+ \sum_{j \in \mathbb{Z}} \tau \Delta x^2 \Delta t e_j D_+ D_- (e)_j + \sum_{j \in \mathbb{Z}} 2 \Delta x \Delta t^2 \lambda D_c \left(\frac{e^2}{2} \right)_j D_c (ue)_j \\ &- \sum_{j \in \mathbb{Z}} \tau \Delta x^2 \Delta t^2 \lambda D_c \left(\frac{e^2}{2} \right)_j D_+ D_- (e)_j - \sum_{j \in \mathbb{Z}} \tau \Delta x^2 \Delta t^2 D_c (ue)_j D_+ D_- (e)_j. \end{aligned}$$

Equations (3.41) and (3.44) (without θ -terms, $D_+ D_-$, $D_+ D_c$ and $D_+ D_+ D_-$ terms due to Airy term) lead to the following upper bound

$$\begin{aligned} \|\mathcal{B}^{Burgers} e\|_{\ell_\Delta^2}^2 &\leq \|e\|_{\ell_\Delta^2}^2 \left\{ 1 + \Delta t \|u\|_{\ell^\infty}^2 + \frac{3 \Delta t^2}{4} \|D_+(u)\|_{\ell^\infty}^2 + \Delta t \|D_+(u)\|_{\ell^\infty} \right. \\ &+ \frac{2 \Delta t^2}{3 \Delta x} \lambda \|D_c(u)\|_{\ell^\infty} \|e\|_{\ell^\infty} + \frac{\tau \Delta t^2}{\Delta x} \|D_+(u)\|_{\ell^\infty} + \frac{\tau \Delta t^2}{\Delta x} \|D_c(u)\|_{\ell^\infty} + \frac{\Delta t^3}{\Delta x^2} \|D_+(u)\|_{\ell^\infty}^2 \Big\} \\ &+ \sum_{j \in \mathbb{Z}} \Delta x (D_+(e)_j)^2 \left[\tau^2 \Delta t^2 + \frac{\Delta x^2 \Delta t}{6} \lambda D_+(e)_j - \tau \Delta t \Delta x - \frac{\Delta t^2 \Delta x \tau}{6} \lambda D_+(e)_j \right] \\ &+ \Delta t^2 \|D_c(e)\|_{\ell_\Delta^2}^2 \left\{ \lambda^2 \|e\|_{\ell^\infty}^2 + \|u\|_{\ell^\infty}^2 - \tau^2 + 2 \lambda \|e\|_{\ell^\infty} \|u\|_{\ell^\infty} + \frac{8 \Delta x^2}{3} \lambda \|D_c(u)\|_{\ell^\infty} \|D_c(e)\|_{\ell^\infty} \right. \\ &\left. + \frac{2 \tau \Delta x}{3} \lambda \|D_c(e)\|_{\ell^\infty} \right\}. \end{aligned}$$

However, we almost recognize $C_b^{Burgers}$ defined by (3.80b) in the previous equation. The only difference results in $\lambda D_+(e)_j$ instead of $D_+(e)_j$:

$$\tau^2 \Delta t^2 + \frac{\Delta x^2 \Delta t}{6} \lambda D_+(e)_j - \tau \Delta t \Delta x - \frac{\Delta t^2 \Delta x \tau}{6} \lambda D_+(e)_j.$$

We choose ε_0 such that

$$\varepsilon_0 \leq (3 \|u\|_{\ell^\infty})^{\frac{2}{1-2\gamma}}, \quad (4.43)$$

to ensure

$$\frac{\Delta x}{6} \lambda D_+(e)_j \leq \frac{\lambda \|e\|_{\ell^\infty}}{3} \leq \frac{\Delta x^{\frac{1}{2}-\gamma}}{3} \leq \frac{\varepsilon_0^{\frac{1}{2}-\gamma}}{3} \leq \|u\|_{\ell^\infty} < \tau.$$

Combined with the hyperbolic CFL-type condition $\frac{\tau \Delta t}{\Delta x} \leq 1 - \beta$, this implies

$$\tau^2 \Delta t^2 + \frac{\Delta x^2 \Delta t}{6} \lambda D_+(e)_j - \tau \Delta t \Delta x - \frac{\Delta t^2 \Delta x \tau}{6} \lambda D_+(e)_j = \left(\frac{\Delta x}{6} \lambda D_+(e)_j - \tau \right) (\Delta x - \tau \Delta t) \Delta t \leq 0$$

In the same way, we almost recognize $C_c^{Burgers}$, defined by (3.80c)

$$\lambda^2 \|e\|_{\ell^\infty}^2 + \|u\|_{\ell^\infty}^2 - \tau^2 + 2\lambda \|e\|_{\ell^\infty} \|u\|_{\ell^\infty} + \frac{8\Delta x^2}{3} \lambda \|D_c(u)\|_{\ell^\infty} \|D_c(e)\|_{\ell^\infty} + \frac{2\tau\Delta x}{3} \lambda \|D_c(e)\|_{\ell^\infty}.$$

For ε_0 satisfying

$$\varepsilon_0^{1-2\gamma} + 2\varepsilon_0^{\frac{1}{2}-\gamma} \|u\|_{\ell^\infty} + \frac{8}{3} \|D_c(u)\|_{\ell^\infty} \varepsilon_0^{\frac{3}{2}-\gamma} + \frac{2\tau}{3} \varepsilon_0^{\frac{1}{2}-\gamma} \leq \alpha^2, \quad (4.44)$$

the condition $\|u\|_{\ell^\infty} + \alpha \leq \tau$ implies

$$\lambda^2 \|e\|_{\ell^\infty}^2 + \|u\|_{\ell^\infty}^2 - \tau^2 + 2\lambda \|e\|_{\ell^\infty} \|u\|_{\ell^\infty} + \frac{8\Delta x^2}{3} \lambda \|D_c(u)\|_{\ell^\infty} \|D_c(e)\|_{\ell^\infty} + \frac{2\tau\Delta x}{3} \lambda \|D_c(e)\|_{\ell^\infty} \leq 0.$$

Proposition 4.32 results from the fact that there exists C_1 such that $\sqrt{1 + C\Delta t} \leq 1 + C_1\Delta t$ with

$$C := \|u\|_{\ell^\infty}^2 + \frac{3\Delta t}{4} \|D_+(u)\|_{\ell^\infty}^2 + \|D_+(u)\|_{\ell^\infty} + \frac{2\Delta t}{3\Delta x} \|D_c(u)\|_{\ell^\infty} \Delta x^{\frac{1}{2}-\gamma} + \frac{\tau\Delta t}{\Delta x} \|D_+(u)\|_{\ell^\infty} \\ + \frac{\tau\Delta t}{\Delta x} \|D_c(u)\|_{\ell^\infty} + \frac{\Delta t^2}{\Delta x^2} \|D_+(u)\|_{\ell^\infty}^2.$$

The upper bound ε_0 must be chosen such that Conditions (4.43) and (4.44) be satisfied. $\square$

The next result is an immediate consequence of the preceding one.

Proposition 4.33. Consider $(u_j)_{j \in \mathbb{Z}} \in \ell^\infty(\mathbb{Z})$ such that $(D_+(u)_j)_{j \in \mathbb{Z}} \in \ell^\infty(\mathbb{Z})$ and $\lambda \geq 0$ respectively $\beta \in (0, 1)$. Fix $\alpha > 0$ and τ such that

$$\|u\|_{\ell^\infty} + \alpha \leq \tau.$$

Then, there exist two sufficiently small positive numbers $\varepsilon_0, \varepsilon_1$ such that the following holds true. Consider two positive reals $\Delta t, \Delta x$ such that

$$\frac{\tau\Delta t}{\Delta x} \leq 1 - \beta, \quad \Delta x \leq \varepsilon_0$$

and $(e_j)_{j \in \mathbb{Z}} \in \ell^2(\mathbb{Z}), (v_j)_{j \in \mathbb{Z}} \in \ell^\infty(\mathbb{Z})$ such that

$$\lambda \|e\|_{\ell^\infty} \leq \Delta x^{\frac{1}{2}-\gamma}, \text{ with } \gamma \in (0, \frac{1}{2}), \quad \|v\|_{\ell^\infty} \leq \varepsilon_1 \text{ and } \|D_+(v)\|_{\ell^\infty} \leq 1.$$

Then, there exists a positive constant C depending on the ℓ^∞ -norms of $(u_j)_{j \in \mathbb{Z}}$ and $(D_+(u)_j)_{j \in \mathbb{Z}}$ such that :

$$\left\| e - \Delta t D_c \left(e \left(u + v + \lambda \frac{e}{2} \right) \right) + \frac{\tau}{2} \Delta x \Delta t D_+ D_-(e) \right\|_{\ell_\Delta^2} \leq (1 + C\Delta t) \|e\|_{\ell_\Delta^2}. \quad (4.45)$$

Proof. Let us consider $\tau > \|u\|_{\ell^\infty}$. Let us suppose that ε_1 is chosen small enough such that

$$\|u + v\|_{\ell^\infty} \leq \varepsilon_1 + \|u\|_{\ell^\infty} < \tau.$$

Then, taking a smaller Δt and Δx if necessary, we may apply Proposition 4.32 with $u + v$ instead of u in order to establish Estimate (4.45). $\square$

The latest lemma concerns the same Burgers-type estimate but for the derivative of $(e_j)_{j \in \mathbb{Z}}$.

Proposition 4.34. Consider $(u_j)_{j \in \mathbb{Z}} \in \ell^\infty(\mathbb{Z})$ such that $(D_+(u)_j)_{j \in \mathbb{Z}} \in \ell^\infty(\mathbb{Z})$ and $\lambda \geq 0$ respectively $\beta \in (0, 1)$. Fix $\alpha > 0$ and τ such that

$$\|u\|_{\ell^\infty} + \alpha \leq \tau.$$

Then, there exist two sufficiently small positive numbers $\varepsilon_0, \varepsilon_1$ such that the following holds true. Consider two positive reals $\Delta t, \Delta x$ such that

$$\frac{\tau \Delta t}{\Delta x} \leq 1 - \beta, \quad \Delta x \leq \varepsilon_0$$

and $(e_j)_{j \in \mathbb{Z}} \in \ell^2(\mathbb{Z}), (v_j)_{j \in \mathbb{Z}} \in \ell^\infty(\mathbb{Z})$ such that

$$\lambda \|e\|_{\ell^\infty}, \lambda \|D_+ e\|_{\ell^\infty} \leq \Delta x^{\frac{1}{2}-\gamma}, \text{ with } \gamma \in (0, \frac{1}{2}), \quad (4.46a)$$

$$\|v\|_{\ell^\infty} < \varepsilon_1 \text{ and } \|D_+ v\|_{\ell^\infty}, \|D_+ D_- (v)\|_{\ell^\infty}, \|D_+ D_- (u)\|_{\ell^\infty} \leq 1. \quad (4.46b)$$

Then, there exist two positive constants C_1, C_2 depending on the ℓ^∞ -norms of $(u_j)_{j \in \mathbb{Z}}, (D_+(u)_j)_{j \in \mathbb{Z}}$ and $(D_c D_+(u)_j)_{j \in \mathbb{Z}}$ such that :

$$\left\| D_+ e - \Delta t D_+ D_c \left(e \left(u + v + \lambda \frac{e}{2} \right) \right) + \frac{\tau}{2} \Delta x \Delta t D_+ D_+ D_- (e) \right\|_{\ell_\Delta^2} \leq C_1 \Delta t \|e\|_{\ell_\Delta^2} + (1 + C_2 \Delta t) \|D_+ e\|_{\ell_\Delta^2}.$$

Proof. Let us observe that

$$\begin{aligned} & D_+(e)_j - \Delta t D_+ D_c \left(e \left(u + v + \lambda \frac{e}{2} \right) \right)_j + \frac{\tau}{2} \Delta x \Delta t D_+ D_+ D_- (e)_j \\ &= D_+(e)_j - \Delta t D_c \left(D_+(e) \left(u + v + \lambda \frac{e}{2} \right) \right)_j - \Delta t D_c \left(\mathcal{S}^1(e) D_+ \left(u + v + \lambda \frac{e}{2} \right) \right)_j + \frac{\tau}{2} \Delta x \Delta t D_+ D_- D_+(e)_j \\ &= D_+(e)_j - \Delta t D_c \left(D_+(e) \left(u + v + \frac{\lambda}{2} \mathcal{S}^1(e) + \lambda \frac{e}{2} \right) \right)_j + \frac{\tau}{2} \Delta x \Delta t D_+ D_- D_+(e)_j \\ &\quad - \Delta t D_c \left(\mathcal{S}^1(e) D_+(u) \right)_j - \Delta t D_c \left(\mathcal{S}^1(e) D_+(v) \right)_j. \end{aligned}$$

Owing to the hypothesis (4.46a)-(4.46b) and Proposition 4.33 with $v + \frac{\lambda}{2} \mathcal{S}^1(e) + \lambda \frac{e}{2}$ instead of v and 0 instead of λ , we may choose ε_0 small enough which ensures the existence of a positive constant C_2 such that :

$$\left\| D_+(e) - \Delta t D_c \left(D_+(e) \left(u + v + \frac{\lambda}{2} \mathcal{S}^1(e) + \lambda \frac{e}{2} \right) \right) + \frac{\tau}{2} \Delta x \Delta t D_+ D_- D_+(e) \right\|_{\ell_\Delta^2} \leq (1 + C_2 \Delta t) \|D_+(e)\|_{\ell_\Delta^2}. \quad (4.47)$$

Moreover, using the derivation formula (3.20), it transpires that

$$\begin{aligned} & \left\| -\Delta t D_c \left(\mathcal{S}^1(e) D_+(u) \right) - \Delta t D_c \left(\mathcal{S}^1(e) D_+(v) \right) \right\|_{\ell_\Delta^2} \\ & \leq \Delta t (\|D_c D_+(u)\|_{\ell^\infty} + \|D_c D_+(v)\|_{\ell^\infty}) \|e\|_{\ell_\Delta^2} + \Delta t (\|D_+(u)\|_{\ell^\infty} + \|D_+(v)\|_{\ell^\infty}) \|D_+(e)\|_{\ell_\Delta^2}. \end{aligned} \quad (4.48)$$

The conclusion follows from Estimates (4.47) and (4.48). $\square$

4.4.2 The case $b = d = 0$

The case $a < 0, b = c = d = 0$

The convergence error satisfies :

$$\begin{cases} e_j^{n+1} + \Delta t D_c(f)_j^{n+1} + a \Delta t D_+ D_- D_c(f)_j^{n+1} = e_j^n - \Delta t D_c(e u_\Delta)_j^n - \Delta t D_c(e f)_j^n \\ \quad - \Delta t D_c(\eta_\Delta f)_j^n + \frac{\tau_1}{2} \Delta t \Delta x D_+ D_- (e)_j^n - \Delta t (\epsilon_1)_j^n, \\ f_j^{n+1} + \Delta t D_c(e)_j^{n+1} = f_j^n - \Delta t D_c(f(u_\Delta + \frac{1}{2}f))_j^n + \frac{\tau_2}{2} \Delta t \Delta x D_+ D_- (f)_j^n - \Delta t (\epsilon_2)_j^n, \end{cases} \quad (4.49)$$

with

$$\begin{cases} \max \{ \tau_1, \tau_2 \} \Delta t \leq \Delta x (1 - \beta_0), \\ \|u_\Delta^n\|_{\ell^\infty} + \alpha_0 \leq \min \{ \tau_1, \tau_2 \}, \quad n \in \llbracket 0, N \rrbracket. \end{cases}$$

We consider the energy functional

$$\mathcal{E}(e, f) := \|e\|_{\ell_\Delta^2}^2 + \|f\|_{\ell_\Delta^2}^2 + (-a) \|D_+ f\|_{\ell_\Delta^2}^2.$$

Proof of Proposition 4.31 in the case $a < 0, b = c = d = 0$. In order to recover a discrete H^1 -type control for $(f_j^n)_{(n,j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}}$ (necessary for the control of $\mathcal{E}(e, f)$), let us apply $\sqrt{-a}D_+$ to the second equation of (4.49). We get that :

$$\begin{aligned} \sqrt{-a}D_+(f)_j^{n+1} + \Delta t \sqrt{-a}D_+D_c(e)_j^{n+1} &= \sqrt{-a}D_+(f)_j^n - \sqrt{-a}\Delta t D_cD_+ \left(f \left(u_\Delta + \frac{1}{2}f \right) \right)_j^n \\ &\quad + \sqrt{-a}\frac{\tau_2}{2}\Delta t \Delta x D_+D_+D_-(f)_j^n - \sqrt{-a}\Delta t D_+(\epsilon_2)_j^n. \end{aligned} \quad (4.50)$$

We consider now the three equations system comprised of System (4.49) with added Equation (4.50). As previously, we square the equations, and add them together. Young's inequality enables us to obtain, with C_0 a constant

$$\begin{aligned} &\|e^{n+1} + \Delta t (I + aD_+D_-) D_c(f)^{n+1}\|_{\ell_\Delta^2}^2 + \|f^{n+1} + \Delta t D_c(e)^{n+1}\|_{\ell_\Delta^2}^2 + \|\sqrt{-a}D_+(f)^{n+1} + \Delta t \sqrt{-a}D_+D_c(e)^{n+1}\|_{\ell_\Delta^2}^2 \\ &\leq (\Delta t + \Delta t^2) C_0 \|\epsilon_1^n\|_{\ell_\Delta^2}^2 + (\Delta t + \Delta t^2) C_0 \|\epsilon_2^n\|_{\ell_\Delta^2}^2 + (-a) (\Delta t + (\Delta t)^2) C_0 \|D_+(\epsilon_2)^n\|_{\ell_\Delta^2}^2 \\ &\quad + \|\text{RHS}^{a < 0}\|_{\ell_\Delta^2}^2, \end{aligned} \quad (4.51)$$

with

$$\begin{aligned} \|\text{RHS}^{a < 0}\|_{\ell_\Delta^2}^2 &:= (1 + C_0\Delta t) \left\| e^n - \Delta t D_c(eu_\Delta)^n - \Delta t D_c(ef)^n + \frac{\tau_1}{2}\Delta t \Delta x D_+D_-(e)^n \right\|_{\ell_\Delta^2}^2 \\ &\quad + (\Delta t + (\Delta t)^2) C_0 \|D_c(\eta_\Delta f)^n\|_{\ell_\Delta^2}^2 \\ &\quad + (1 + C_0\Delta t) \left\| f^n - \Delta t D_c \left(f \left(u_\Delta + \frac{1}{2}f \right) \right)^n + \frac{\tau_2}{2}\Delta t \Delta x D_+D_-(f)^n \right\|_{\ell_\Delta^2}^2 \\ &\quad + (1 + C_0\Delta t) (-a) \left\| D_+(f)^n - \Delta t D_cD_+ \left(f \left(u_\Delta + \frac{1}{2}f \right) \right)^n + \frac{\tau_2}{2}\Delta t \Delta x D_+D_+D_-(f)^n \right\|_{\ell_\Delta^2}^2. \end{aligned}$$

Let us consider the right hand side of (4.51). Owing to Proposition 4.33 and Proposition 4.32 and thanks to hypothesis of Proposition 4.31 ($\|e^n\|_{\ell^\infty}, \|f^n\|_{\ell^\infty}, \|D_+(f)^n\|_{\ell^\infty} \leq \Delta x^{\frac{1}{2}-\gamma}$), we have that

$$\left\| e^n - \Delta t D_c(eu_\Delta)^n - \Delta t D_c(ef)^n + \frac{\tau_1}{2}\Delta t \Delta x D_+D_-(e)^n \right\|_{\ell_\Delta^2}^2 \leq (1 + C_1\Delta t) \|e^n\|_{\ell_\Delta^2}^2, \quad (4.52)$$

respectively that

$$\left\| f^n - \Delta t D_c \left(f \left(u_\Delta + \frac{1}{2}f \right) \right)^n + \frac{\tau_2}{2}\Delta t \Delta x D_+D_-(f)^n \right\|_{\ell_\Delta^2}^2 \leq (1 + C_2\Delta t) \|f^n\|_{\ell_\Delta^2}^2. \quad (4.53)$$

Proposition (4.34) ensures the existence of constants C_3, C_4 such that :

$$\begin{aligned} &(-a) \left\| D_+(f)^n - \Delta t D_+D_c \left(f \left(u_\Delta + \frac{1}{2}f \right) \right)^n + \frac{\tau_2}{2}\Delta t \Delta x D_+D_+D_-(f)^n \right\|_{\ell_\Delta^2}^2 \\ &\leq (1 + C_3\Delta t) (-a) \|D_+(f)^n\|_{\ell_\Delta^2}^2 + C_4(-a)\Delta t \|f^n\|_{\ell_\Delta^2}^2. \end{aligned} \quad (4.54)$$

Finally, we have that, thanks Relation (3.20)

$$\begin{aligned} \|D_c(\eta_\Delta f)^n\|_{\ell_\Delta^2}^2 &\leq 2 \|D_c(\eta_\Delta)^n\|_{\ell_\infty}^2 \|f^n\|_{\ell_\Delta^2}^2 + 2 \|\eta_\Delta^n\|_{\ell_\infty}^2 \|D_c(f)^n\|_{\ell_\Delta^2}^2 \\ &\leq 2 \max \left\{ \|D_c(\eta_\Delta)^n\|_{\ell_\infty}^2, \frac{\|\eta_\Delta^n\|_{\ell_\infty}^2}{-a} \right\} \mathcal{E}(e^n, f^n). \end{aligned} \quad (4.55)$$

Next, we compute the left hand side of (4.51). First of all, we observe that

$$\begin{aligned} &\left\| \sqrt{-a} D_+(f)^{n+1} + \Delta t \sqrt{-a} D_+ D_c(e)^{n+1} \right\|_{\ell_\Delta^2}^2 \\ &= -a \|D_+(f)^{n+1}\|_{\ell_\Delta^2}^2 - a \Delta t^2 \|D_+ D_c(e)^{n+1}\|_{\ell_\Delta^2}^2 + 2a \Delta t \sum_{j \in \mathbb{Z}} \Delta x D_+ D_-(f)_j^{n+1} D_c(e)_j^{n+1}. \end{aligned}$$

Moreover, one has

$$\begin{aligned} &\left\| e^{n+1} + \Delta t (I + a D_+ D_-) D_c(f)^{n+1} \right\|_{\ell_\Delta^2}^2 + \left\| f^{n+1} + \Delta t D_c(e)^{n+1} \right\|_{\ell_\Delta^2}^2 = \|e^{n+1}\|_{\ell_\Delta^2}^2 \\ &+ \sum_{j \in \mathbb{Z}} 2a \Delta t \Delta x e_j^{n+1} D_+ D_- D_c(f)_j^{n+1} + \Delta t^2 \left\| (I + a D_+ D_-) D_c(f)^{n+1} \right\|_{\ell_\Delta^2}^2 + \|f^{n+1}\|_{\ell_\Delta^2}^2 + \Delta t^2 \|D_c(e)^{n+1}\|_{\ell_\Delta^2}^2. \end{aligned}$$

Eventually, we get, for the left hand side,

$$\begin{aligned} &\left\| e^{n+1} + \Delta t (I + a D_+ D_-) D_c f^{n+1} \right\|_{\ell_\Delta^2}^2 + \left\| f^{n+1} + \Delta t D_c e^{n+1} \right\|_{\ell_\Delta^2}^2 + \left\| \sqrt{-a} D_+ f^{n+1} + \Delta t \sqrt{-a} D_+ D_c e^{n+1} \right\|_{\ell_\Delta^2}^2 \\ &= \mathcal{E}(e^{n+1}, f^{n+1}) + \Delta t^2 \left\| (I + a D_+ D_-) D_c(f)^{n+1} \right\|_{\ell_\Delta^2}^2 + \Delta t^2 \|D_c(e)^{n+1}\|_{\ell_\Delta^2}^2 - a \Delta t^2 \|D_+ D_c(e)^{n+1}\|_{\ell_\Delta^2}^2 \\ &\geq \mathcal{E}(e^{n+1}, f^{n+1}). \end{aligned} \quad (4.56)$$

Gathering relations (4.51), (4.52), (4.53), (4.54), (4.55) and (4.56) yields

$$\begin{aligned} \mathcal{E}(e^{n+1}, f^{n+1}) &\leq \left(\Delta t + (\Delta t)^2 \right) C_0 \left(\|\epsilon_1^n\|_{\ell_\Delta^2}^2 + \|\epsilon_2^n\|_{\ell_\Delta^2}^2 + (-a) \|D_+ \epsilon_2^n\|_{\ell_\Delta^2}^2 \right) \\ &+ (1 + C \Delta t) \mathcal{E}(e^n, f^n). \end{aligned}$$

□

4.4.3 The case $b = 0, d > 0$

The case $a = b = c = 0, d > 0$.

In this case, without any difficulties, we are able to prove a more general result. Indeed, we will show that the following general θ -scheme :

$$\begin{cases} \frac{1}{\Delta t} (\eta_j^{n+1} - \eta_j^n) + (1 - \theta) D_c(u)_j^n + \theta D_c(u)_j^{n+1} + D_c(\eta u)_j^n = \frac{\tau_1 \Delta x}{2} D_+ D_- (\eta)_j^n, \\ \frac{1}{\Delta t} (I - d D_+ D_-) (u_j^{n+1} - u_j^n) + (1 - \theta) D_c(\eta)_j^n + \theta D_c(\eta)_j^{n+1} + \frac{1}{2} D_c(u^2)_j^n = 0. \end{cases} \quad (4.57)$$

is adapted to study the classical Boussinesq system, with

$$\begin{cases} \tau_1 \Delta t \leq \Delta x (1 - \beta_0), \\ \|u_\Delta^n\|_{\ell_\infty} + \alpha_0 \leq \tau_1, \quad n \in \llbracket 0, N \rrbracket. \end{cases}$$

The convergence error verifies :

$$\left\{ \begin{array}{l} (E^-)_j^n + \Delta t (1 - \theta) D_c (f)_j^n + \theta \Delta t D_c (f)_j^{n+1} + \Delta t D_c (eu_\Delta)_j^n + \Delta t D_c (\eta_\Delta f)_j^n \\ \quad + \Delta t D_c (ef)_j^n = \frac{\tau_1}{2} \Delta t \Delta x D_+ D_- (e)_j^n - \Delta t (\epsilon_1)_j^n, \\ (I - d D_+ D_-) (F^-)_j^n + \Delta t (1 - \theta) D_c (e)_j^n + \theta \Delta t D_c (e)_j^{n+1} + \Delta t D_c (fu_\Delta)_j^n \\ \quad + \frac{\Delta t}{2} D_c (f^2)_j^n = -\Delta t (\epsilon_2)_j^n. \end{array} \right. \quad (4.58)$$

Recall that in this case :

$$\mathcal{E}(e, f) := \|e\|_{\ell_\Delta^2}^2 + \|f\|_{\ell_\Delta^2}^2 + d \|D_+ f\|_{\ell_\Delta^2}^2.$$

Proof of Proposition 4.31 in the case $a = b = c = 0, d > 0$. Multiply the second equation of (4.58) with $(F^+)_j^n$ we recognize (4.28), (4.31), (4.35), and a part of (4.37), all with $a = 0$. Proceeding as we did in Section 4.3.1, we obtain that there exist two constants $C_{1,1}$ and $C_{2,1}$ depending on d , $\|u_\Delta^n\|_{\ell^\infty}$, $\|Du_\Delta^n\|_{\ell^\infty}$ such that :

$$\begin{aligned} \|f^{n+1}\|_{\ell_\Delta^2}^2 + d \|D_+ (f)^{n+1}\|_{\ell_\Delta^2}^2 - \|f^n\|_{\ell_\Delta^2}^2 - d \|D_+ (f)^n\|_{\ell_\Delta^2}^2 &\leq \Delta t \|\epsilon_2^n\|_{\ell_\Delta^2}^2 + \Delta t C_{1,1} \mathcal{E}(e^n, f^n) \\ &+ \Delta t C_{2,1} \mathcal{E}(e^{n+1}, f^{n+1}) - \Delta t (1 - \theta) \sum_{j \in \mathbb{Z}} \Delta x D_c (e)_j^n (F^+)_j^n - \theta \Delta t \sum_{j \in \mathbb{Z}} \Delta x D_c (e)_j^{n+1} (F^+)_j^n. \end{aligned} \quad (4.59)$$

Notice that Relation (3.28) combined with Cauchy-Schwarz inequality, Young inequality and the upper bound $\|D_c(\cdot)\|_{\ell_\Delta^2} \leq \|D_+(\cdot)\|_{\ell_\Delta^2}$ (for example in Inequality (3.25)) simplifies $\sum_{j \in \mathbb{Z}} \Delta x D_c (e)_j^n (F^+)_j^n$

$$\begin{aligned} \sum_{j \in \mathbb{Z}} \Delta x D_c (e)_j^n (F^+)_j^n &\leq \|e^n\|_{\ell_\Delta^2}^2 + \frac{1}{2} \|D_c (f)^n\|_{\ell_\Delta^2}^2 + \frac{1}{2} \|D_c (f)^{n+1}\|_{\ell_\Delta^2}^2 \\ &\leq \max \left\{ 1, \frac{1}{2d} \right\} \mathcal{E}(e^n, f^n) + \frac{1}{2d} \mathcal{E}(e^{n+1}, f^{n+1}). \end{aligned}$$

A similar inequality holds true for $\sum_{j \in \mathbb{Z}} \Delta x D_c (e)_j^{n+1} (F^+)_j^n$.

Next we rewrite the first equation of (4.58) as

$$\begin{aligned} e_j^{n+1} &= e_j^n - \Delta t D_c (e (u_\Delta + f))_j^n + \frac{\tau_1}{2} \Delta t \Delta x D_+ D_- (e)_j^n - \Delta t (1 - \theta) D_c (f)_j^n - \theta \Delta t D_c (f)_j^{n+1} \\ &\quad - \Delta t D_c (\eta_\Delta f)_j^n - \Delta t (\epsilon_1)_j^n. \end{aligned}$$

Thus, using the Proposition (4.33) and the Relation (3.20) we get that

$$\begin{aligned} \|e^{n+1}\|_{\ell_\Delta^2} &\leq \|e^n - \Delta t D_c (e (u_\Delta + f))_j^n + \frac{\tau_1}{2} \Delta t \Delta x D_+ D_- (e)_j^n\|_{\ell_\Delta^2} \\ &\quad + \Delta t \|(1 - \theta) D_c (f)_j^n + \theta D_c (f)_j^{n+1}\|_{\ell_\Delta^2} + \Delta t \|D_c (\eta_\Delta f)_j^n\|_{\ell_\Delta^2} + \Delta t \|\epsilon_1^n\|_{\ell_\Delta^2} \\ &\leq (1 + C \Delta t) \|e^n\|_{\ell_\Delta^2} + \Delta t \left((1 - \theta) \|D_c (f)_j^n\|_{\ell_\Delta^2} + \theta \|D_c (f)_j^{n+1}\|_{\ell_\Delta^2} \right) \\ &\quad + \Delta t \max \{ \|\eta_\Delta^n\|_{\ell^\infty}, \|D_c (\eta_\Delta)_j^n\|_{\ell^\infty} \} \left(\|f^n\|_{\ell_\Delta^2} + \|D_c (f)_j^n\|_{\ell_\Delta^2} \right) + \Delta t \|\epsilon_1^n\|_{\ell_\Delta^2} \\ &\leq (1 + C \Delta t) \|e^n\|_{\ell_\Delta^2} + \Delta t C_{1,2} \mathcal{E}^{\frac{1}{2}}(e^n, f^n) + C_{2,2} \Delta t \mathcal{E}^{\frac{1}{2}}(e^{n+1}, f^{n+1}) + \Delta t \|\epsilon_1^n\|_{\ell_\Delta^2}. \end{aligned} \quad (4.60)$$

with $C_{1,2}$ and $C_{2,2}$ depending on $\|\eta_\Delta^n\|_{\ell^\infty}$ and $\|D_c (\eta_\Delta)_j^n\|_{\ell^\infty}$. From (4.60), we deduce that

$$\begin{aligned} \|e^{n+1}\|_{\ell_\Delta^2}^2 &\leq (\Delta t + \Delta t^2) C_0 \|\epsilon_1^n\|_{\ell_\Delta^2}^2 + (1 + C_3 \Delta t) \|e^n\|_{\ell_\Delta^2}^2 \\ &\quad + (\Delta t + \Delta t^2) C_{1,3} \mathcal{E}(e^n, f^n) + (\Delta t + \Delta t^2) C_{2,3} \mathcal{E}(e^{n+1}, f^{n+1}). \end{aligned} \quad (4.61)$$

Adding up the estimates yields

$$(1 - C_1 \Delta t) \mathcal{E} \left(e^{n+1}, f^{n+1} \right) \leq (\Delta t + \Delta t^2) C_0 \|\epsilon_1^n\|_{\ell_\Delta^2}^2 + \Delta t \|\epsilon_2^n\|_{\ell_\Delta^2}^2 + (1 + C_2 \Delta t) \mathcal{E} \left(e^n, f^n \right).$$

□

The case $a < 0, b = 0, c = 0, d > 0$

In this case, the convergence error satisfies :

$$\begin{cases} e_j^{n+1} + \Delta t D_c(f)_j^{n+1} + a \Delta t D_+ D_- D_c(f)_j^{n+1} = e_j^n - \Delta t D_c(e u_\Delta)_j^n - \Delta t D_c(e f)_j^n - \Delta t D_c(\eta_\Delta f)_j^n \\ \quad + \frac{\tau_1}{2} \Delta t \Delta x D_+ D_-(e)_j^n - \Delta t (\epsilon_1)_j^n, \\ (I - d D_+ D_-) f_j^{n+1} + \Delta t D_c(e)_j^{n+1} = (I - d D_+ D_-) f_j^n - \Delta t D_c(f(u_\Delta + \frac{1}{2}f))_j^n - \Delta t (\epsilon_2)_j^n, \end{cases}$$

with

$$\begin{cases} \tau_1 \Delta t \leq \Delta x (1 - \beta_0), \\ \|u_\Delta^n\|_{\ell_\infty} + \alpha_0 \leq \tau_1, \quad n \in \llbracket 0, N \rrbracket. \end{cases}$$

For that case, we will work with the energy functional :

$$\mathcal{E}(e, f) := d \|e\|_{\ell_\Delta^2}^2 + (-a) \|(I - d D_+ D_-) f\|_{\ell_\Delta^2}^2,$$

which is better adapted to the system corresponding to the particular values of the parameters in view here. Of course, $\mathcal{E}$ is equivalent to the energy from (4.4).

Proof of Proposition 4.31 in the case $a < 0, b = c = 0, d > 0$. By summing up the square of the ℓ_Δ^2 -norm of the first equation by d , the square of the ℓ_Δ^2 norm of the second one by $(-a)$, we get that (thanks Relations (3.27) and (3.28))

$$\begin{aligned} \mathcal{E} \left(e^{n+1}, f^{n+1} \right) &+ 2(a + d) \Delta t \Delta x \sum_{j \in \mathbb{Z}} e_j^{n+1} D_c(f)_j^{n+1} + d \Delta t^2 \|D_c(f)^{n+1}\|_{\ell_\Delta^2}^2 + 2 \Delta t^2 (-a) d \|D_+ D_c(f)^{n+1}\|_{\ell_\Delta^2}^2 \\ &+ (-a) \Delta t^2 \|D_c(e)^{n+1}\|_{\ell_\Delta^2}^2 + \Delta t^2 a^2 d \|D_+ D_- D_c(f)^{n+1}\|_{\ell_\Delta^2}^2 \\ &\leq (\Delta t + \Delta t^2) C_0 d \|\epsilon_1^n\|_{\ell_\Delta^2}^2 + (\Delta t + \Delta t^2) C_0 (-a) \|\epsilon_2^n\|_{\ell_\Delta^2}^2 \\ &+ (1 + \Delta t C_0) d \left\| e^n - \Delta t D_c(e(u_\Delta + f))_j^n + \frac{\tau_1}{2} \Delta t \Delta x D_+ D_-(e)_j^n \right\|_{\ell_\Delta^2}^2 \\ &+ (1 + \Delta t C_0) (-a) \|(I - d D_+ D_-) f^n\|_{\ell_\Delta^2}^2 + (\Delta t + \Delta t^2) C_0 d \|D_c(\eta_\Delta f)^n\|_{\ell_\Delta^2}^2 \\ &+ (\Delta t + \Delta t^2) C_0 (-a) \left\| D_c \left(f \left(u_\Delta + \frac{1}{2} f \right) \right)_j^n \right\|_{\ell_\Delta^2}^2. \end{aligned}$$

We notice that

$$\begin{aligned} 2 \Delta t \Delta x (d + a) \sum_{j \in \mathbb{Z}} e_j^{n+1} D_c(f)_j^{n+1} &\geq -\Delta t (d - a) \|D_c(f)^{n+1}\|_{\ell_\Delta^2}^2 - \Delta t (d - a) \|e^{n+1}\|_{\ell_\Delta^2}^2 \\ &\geq -\Delta t C_1 \mathcal{E} \left(e^{n+1}, f^{n+1} \right). \end{aligned}$$

Using Proposition 4.33 we get that

$$(1 + \Delta t C_0) d \left\| e^n - \Delta t D_c(e(u_\Delta + f))_j^n + \frac{\tau_1}{2} \Delta t \Delta x D_+ D_-(e)_j^n \right\|_{\ell_\Delta^2}^2 \leq (1 + C_2 \Delta t) d \|e^n\|_{\ell_\Delta^2}^2.$$

Also, we have due to (3.20) :

$$\|D_c(\eta_\Delta f)^n\|_{\ell_\Delta^2}^2 \leq 2 \|D_c(\eta_\Delta)^n\|_{\ell_\infty}^2 \|f^n\|_{\ell_\Delta^2}^2 + 2 \|\eta_\Delta^n\|_{\ell_\infty}^2 \|D_c(f)^n\|_{\ell_\Delta^2}^2 \leq C_3 \mathcal{E} \left(e^n, f^n \right).$$

Proceeding as above, we get that :

$$\left\| D_c \left(f \left(u_\Delta + \frac{1}{2}f \right) \right)^n \right\|_{\ell_\Delta^2}^2 \leq \max \{ \|D_c(u_\Delta)^n\|_{\ell_\infty}^2, \|u_\Delta^n\|_{\ell_\infty}^2, \|D_c(f)^n\|_{\ell_\infty}^2, \|f^n\|_{\ell_\infty}^2 \} C_4 \mathcal{E}(e^n, f^n).$$

Adding up the above estimate gives us :

$$(1 - C_1 \Delta t) \mathcal{E}(e^{n+1}, f^{n+1}) \leq (\Delta t + (\Delta t)^2) C_0 d \|\epsilon_1^n\|_{\ell_\Delta^2}^2 + (\Delta t + (\Delta t)^2) C_0 (-a) \|\epsilon_2^n\|_{\ell_\Delta^2}^2 + (1 + C_5 \Delta t) \mathcal{E}(e^n, f^n).$$

□

The case $a < 0, b = 0, c < 0, d > 0$

In this case, the convergence error satisfies :

$$\begin{cases} e_j^{n+1} + \Delta t D_c(f)_j^{n+1} + a \Delta t D_+ D_- D_c(f)_j^{n+1} = e_j^n - \Delta t D_c(e u_\Delta)_j^n - \Delta t D_c(e f)_j^n \\ \quad - \Delta t D_c(\eta_\Delta f)_j^n + \frac{\tau_1}{2} \Delta t \Delta x D_+ D_- (e)_j^n - \Delta t (\epsilon_1)_j^n, \\ (I - d D_+ D_-) f_j^{n+1} + \Delta t D_c(e)_j^{n+1} + c \Delta t D_+ D_- D_c(e)_j^{n+1} = (I - d D_+ D_-) f_j^n \\ \quad - \Delta t D_c(f(u_\Delta + \frac{1}{2}f))_j^n - \Delta t (\epsilon_2)_j^n, \end{cases} \quad (4.62)$$

with

$$\begin{cases} \tau_1 \Delta t \leq \Delta x (1 - \beta_0), \\ \|u_\Delta^n\|_{\ell_\infty} + \alpha_0 \leq \tau_1, \quad n \in \llbracket 0, N \rrbracket. \end{cases}$$

Again, in order to close the estimates and prove the convergence of the scheme, we will be using the following energy functional :

$$\mathcal{E}(e, f) := (-a) \|e\|_{\ell_\Delta^2}^2 + (-cd) \|D_+(e)\|_{\ell_\Delta^2}^2 + (-a) \|(I - d D_+ D_-) f\|_{\ell_\Delta^2}^2,$$

which is equivalent to the one from (4.4).

Proof of Proposition 4.31 in the case $a < 0, b = 0, c < 0, d > 0$. In order to derive a H^1 control on e^{n+1} , let us apply $\sqrt{-cd} D_+$ in the first equation, to obtain

$$\begin{aligned} & \sqrt{-cd} D_+(e)_j^{n+1} + \Delta t \sqrt{-cd} D_+ D_c(f)_j^{n+1} + \sqrt{-cd} a \Delta t D_+ D_+ D_- D_c(f)_j^{n+1} \\ &= \sqrt{-cd} D_+(e)_j^n - \Delta t \sqrt{-cd} D_+ D_c(e u_\Delta)_j^n - \Delta t \sqrt{-cd} D_+ D_c(e f)_j^n - \sqrt{-cd} \Delta t D_+ D_c(\eta_\Delta f)_j^n \\ &+ \frac{\tau_1}{2} \sqrt{-cd} \Delta t \Delta x D_+ D_+ D_- (e)_j^n - \sqrt{-cd} \Delta t D_+ (\epsilon_1)_j^n. \end{aligned} \quad (4.63)$$

We focus now on the system composed of both equations of (4.62) with Equation (4.63) in addition. We will multiply the first and second equations of (4.62) by $(-a)$ thereafter. As before, we square the three equalities to compute the ℓ_Δ^2 -norm.

Left hand side $\|\text{LHS}^{b=0}\|_{\ell_\Delta^2}^2$. Let us observe that

$$\begin{aligned} \|\text{LHS}^{b=0}\|_{\ell_\Delta^2}^2 &:= (-a) \left\| e^{n+1} + \Delta t D_c(f)^{n+1} + a \Delta t D_+ D_- D_c(f)^{n+1} \right\|_{\ell_\Delta^2}^2 \\ &+ (-cd) \left\| D_+(e)^{n+1} + \Delta t D_+ D_c(f)^{n+1} + a \Delta t D_+ D_+ D_- D_c(f)^{n+1} \right\|_{\ell_\Delta^2}^2 \\ &+ (-a) \left\| (I - d D_+ D_-) f^{n+1} + \Delta t D_c(e)^{n+1} + c \Delta t D_+ D_- D_c(e)^{n+1} \right\|_{\ell_\Delta^2}^2. \end{aligned}$$

- We develop each term.

$$\begin{aligned}
 & (-a) \left\| e^{n+1} + \Delta t D_c(f)^{n+1} + a \Delta t D_+ D_- D_c(f)^{n+1} \right\|_{\ell_\Delta^2}^2 \\
 &= (-a) \left\| e^{n+1} \right\|_{\ell_\Delta^2}^2 + (-a) \Delta t^2 \left\| D_c(f)^{n+1} \right\|_{\ell_\Delta^2}^2 - a^3 \Delta t^2 \left\| D_+ D_- D_c(f)^{n+1} \right\|_{\ell_\Delta^2}^2 \\
 &+ 2(-a) \Delta t \Delta x \sum_{j \in \mathbb{Z}} e_j^{n+1} D_c(f)_j^{n+1} - 2a^2 \Delta t \Delta x \sum_{j \in \mathbb{Z}} e_j^{n+1} D_+ D_- D_c(f)_j^{n+1} + 2a^2 \Delta t^2 \left\| D_+ D_c(f)^{n+1} \right\|_{\ell_\Delta^2}^2.
 \end{aligned} \tag{4.64}$$

- Along with

$$\begin{aligned}
 & (-cd) \left\| D_+(e)^{n+1} + \Delta t D_+ D_c(f)^{n+1} + a \Delta t D_+ D_+ D_- D_c(f)^{n+1} \right\|_{\ell_\Delta^2}^2 \\
 &= (-cd) \left\| D_+(e)^{n+1} \right\|_{\ell_\Delta^2}^2 + (-cd) \Delta t^2 \left\| D_+ D_c(f)^{n+1} \right\|_{\ell_\Delta^2}^2 + (-cd) a^2 \Delta t^2 \left\| D_+ D_+ D_- D_c(f)^{n+1} \right\|_{\ell_\Delta^2}^2 \\
 &+ 2acd \Delta t^2 \left\| D_+ D_- D_c(f)^{n+1} \right\|_{\ell_\Delta^2}^2 - 2cd \Delta t \Delta x \sum_{j \in \mathbb{Z}} D_+(e)_j^{n+1} D_+ D_c(f)_j^{n+1} \\
 &- 2acd \Delta t \Delta x \sum_{j \in \mathbb{Z}} D_+(e)_j^{n+1} D_+ D_+ D_- D_c(f)_j^{n+1},
 \end{aligned} \tag{4.65}$$

- and

$$\begin{aligned}
 & (-a) \left\| (I - dD_+ D_-) f^{n+1} + \Delta t D_c(e)^{n+1} + c \Delta t D_+ D_- D_c(e)^{n+1} \right\|_{\ell_\Delta^2}^2 \\
 &= (-a) \left\| (I - dD_+ D_-) f^{n+1} \right\|_{\ell_\Delta^2}^2 + (-a) c^2 \Delta t^2 \left\| D_+ D_- D_c(e)^{n+1} \right\|_{\ell_\Delta^2}^2 + (-a) \Delta t^2 \left\| D_c(e)^{n+1} \right\|_{\ell_\Delta^2}^2 \\
 &+ 2ac \Delta t^2 \left\| D_+ D_c(e)^{n+1} \right\|_{\ell_\Delta^2}^2 + 2(-a) \Delta t \Delta x \sum_{j \in \mathbb{Z}} (I - dD_+ D_-) f_j^{n+1} D_c(e)_j^{n+1} \\
 &- 2ac \Delta t \Delta x \sum_{j \in \mathbb{Z}} f_j^{n+1} D_+ D_- D_c(e)_j^{n+1} + 2acd \Delta t \Delta x \sum_{j \in \mathbb{Z}} D_+ D_-(f)_j^{n+1} D_+ D_- D_c(e)_j^{n+1}.
 \end{aligned} \tag{4.66}$$

Observe that the last term from (4.65) cancels with the last term of (4.66), according to Relations (3.27) and (3.28). The same is true for $\sum_{j \in \mathbb{Z}} e_j^{n+1} D_c(f)_j^{n+1}$ in (4.64) and $\sum_{j \in \mathbb{Z}} f_j^{n+1} D_c(e)_j^{n+1}$ in (4.66). Therefore, by integrating by part $\sum_{j \in \mathbb{Z}} e_j^{n+1} D_+ D_- D_c(f)_j^{n+1}$ in (4.64) and $\sum_{j \in \mathbb{Z}} f_j^{n+1} D_+ D_- D_c(e)_j^{n+1}$ in (4.66), it yields

$$\begin{aligned}
 \|\text{LHS}^{b=0}\|_{\ell_\Delta^2}^2 &\geq \mathcal{E}(e^{n+1}, f^{n+1}) + 2a^2 \Delta t \Delta x \sum_{j \in \mathbb{Z}} D_c(e)_j^{n+1} D_+ D_-(f)_j^{n+1} \\
 &- 2cd \Delta t \Delta x \sum_{j \in \mathbb{Z}} D_+(e)_j^{n+1} D_+ D_c(f)_j^{n+1} + 2ad \Delta t \Delta x \sum_{j \in \mathbb{Z}} D_+ D_-(f)_j^{n+1} D_c(e)_j^{n+1} \\
 &- 2ac \Delta t \Delta x \sum_{j \in \mathbb{Z}} D_+ D_-(f)_j^{n+1} D_c(e)_j^{n+1}.
 \end{aligned}$$

Young inequality enables us to lower bound this left hand side.

$$\begin{aligned}
 \|\text{LHS}^{b=0}\|_{\ell_\Delta^2}^2 &\geq \mathcal{E}(e^{n+1}, f^{n+1}) - (a^2 - cd + d(-a) + ac) \Delta t \left\{ \|D_+(e)^{n+1}\|_{\ell_\Delta^2}^2 + \|D_+ D_-(f)^{n+1}\|_{\ell_\Delta^2}^2 \right\} \\
 &\geq (1 - C_1 \Delta t) \mathcal{E}(e^{n+1}, f^{n+1}).
 \end{aligned}$$

Right hand side $\|\text{RHS}^{b=0}\|_{\ell_\Delta^2}^2$. Let us now focus on the right hand side of the squared equations.

$$\begin{aligned} \|\text{RHS}^{b=0}\|_{\ell_\Delta^2}^2 := & (-a) \left\| e^n - \Delta t D_c (eu_\Delta)^n - \Delta t D_c (ef)^n + \frac{\tau_1}{2} \Delta t \Delta x D_+ D_- (e)^n - \Delta t D_c (\eta_\Delta f)^n - \Delta t \epsilon_1^n \right\|_{\ell_\Delta^2}^2 \\ & - cd \left\| D_+ (e)^n - \Delta t D_+ D_c (e(u_\Delta + f))^n + \frac{\tau_1}{2} \Delta t \Delta x D_+ D_+ D_- (e)^n - \Delta t D_+ D_c (\eta_\Delta f)^n - \Delta t D_+ (\epsilon_1)^n \right\|_{\ell_\Delta^2}^2 \\ & - a \left\| (I - dD_+ D_-) f^n - \Delta t D_c \left(f \left(u_\Delta + \frac{1}{2} f \right) \right)^n - \Delta t \epsilon_2^n \right\|_{\ell_\Delta^2}^2. \end{aligned}$$

• Firstly, using Cauchy-Schwarz inequality and Proposition 4.33, we get that

$$\begin{aligned} & (-a) \left\| e^n - \Delta t D_c (eu_\Delta)^n - \Delta t D_c (ef)^n + \frac{\tau_1}{2} \Delta t \Delta x D_+ D_- (e)^n - \Delta t D_c (\eta_\Delta f)^n - \Delta t \epsilon_1^n \right\|_{\ell_\Delta^2}^2 \\ & \leq (-a) \left\| e^n - \Delta t D_c (eu_\Delta)^n - \Delta t D_c (ef)^n + \frac{\tau_1}{2} \Delta t \Delta x D_+ D_- (e)^n \right\|_{\ell_\Delta^2}^2 (1 + C_0 \Delta t) \\ & \quad + (-a)(\Delta t + \Delta t^2) C_0 \|D_c (\eta_\Delta f)^n\|_{\ell_\Delta^2}^2 + (-a)(\Delta t + \Delta t^2) C_0 \|\epsilon_1^n\|_{\ell_\Delta^2}^2 \\ & \leq (1 + C \Delta t) (-a) \|e^n\|_{\ell_\Delta^2}^2 + (\Delta t + \Delta t^2) (-a) C_0 \|D_c (\eta_\Delta f)^n\|_{\ell_\Delta^2}^2 + (\Delta t + \Delta t^2) (-a) C_0 \|\epsilon_1^n\|_{\ell_\Delta^2}^2. \end{aligned} \quad (4.67)$$

• Using Proposition 4.34 and Young inequality, we get that

$$\begin{aligned} & -cd \left\| D_+ (e)^n - \Delta t D_+ D_c (e(u_\Delta + f))^n + \frac{\tau_1}{2} \Delta t \Delta x D_+ D_+ D_- (e)^n - \Delta t D_+ D_c (\eta_\Delta f)^n - \Delta t D_+ (\epsilon_1)^n \right\|_{\ell_\Delta^2}^2 \\ & \leq -cd(1 + C_0 \Delta t) \left\| D_+ e^n - \Delta t D_+ D_c (e(u_\Delta + f))^n + \frac{\tau_1}{2} \Delta t \Delta x D_+ D_+ D_- (e)^n \right\|_{\ell_\Delta^2}^2 \\ & \quad - cd(\Delta t + \Delta t^2) C_0 \|D_+ D_c (\eta_\Delta f)^n\|_{\ell_\Delta^2}^2 - cd(\Delta t + \Delta t^2) C_0 \|D_+ (\epsilon_1)^n\|_{\ell_\Delta^2}^2 \\ & \leq -cd \left(\Delta t C_1 \|e^n\|_{\ell_\Delta^2}^2 + (1 + C_2 \Delta t) \|D_+ (e)^n\|_{\ell_\Delta^2}^2 \right) - cd(\Delta t + \Delta t^2) C_0 \|D_+ (\epsilon_1)^n\|_{\ell_\Delta^2}^2 \\ & \quad - cd(\Delta t + \Delta t^2) \max \left\{ \|\eta_\Delta^n\|_{\ell_\infty}^2, \|D_+ (\eta_\Delta)^n\|_{\ell_\infty}^2, \|D_+ D_c (\eta_\Delta)^n\|_{\ell_\infty}^2 \right\} \left(\|f^n\|_{\ell_\Delta^2}^2 + \|D_+ (f)^n\|_{\ell_\Delta^2}^2 + \|D_+ D_- (f)^n\|_{\ell_\Delta^2}^2 \right) \\ & \leq -cd(\Delta t + \Delta t^2) C_0 \|D_+ (\epsilon_1)^n\|_{\ell_\Delta^2}^2 + (-cd) \|D_+ e^n\|_{\ell_\Delta^2}^2 + \Delta t C_3 \mathcal{E}(e^n, f^n). \end{aligned} \quad (4.68)$$

• Using once again Young inequality, we get that

$$\begin{aligned} & -a \left\| (I - dD_+ D_-) f^n - \Delta t D_c \left(f \left(u_\Delta + \frac{1}{2} f \right) \right)^n - \Delta t \epsilon_2^n \right\|_{\ell_\Delta^2}^2 \\ & \leq -a(1 + C_0 \Delta t) \|(I - dD_+ D_-) f^n\|_{\ell_\Delta^2}^2 - a(\Delta t + \Delta t^2) C_0 \left\| D_c \left(f \left(u_\Delta + \frac{1}{2} f \right) \right)^n \right\|_{\ell_\Delta^2}^2 - a(\Delta t + \Delta t^2) C_0 \|\epsilon_2^n\|_{\ell_\Delta^2}^2 \\ & \leq -a(\Delta t + \Delta t^2) C_0 \|\epsilon_2^n\|_{\ell_\Delta^2}^2 + C \Delta t \max \{1, \|u_\Delta^n\|_{\ell_\infty}, \|D_+ (u_\Delta)^n\|_{\ell_\infty}\} \mathcal{E}(e^n, f^n) + (-a) \|(I - dD_+ D_-) f^n\|_{\ell_\Delta^2}^2. \end{aligned} \quad (4.69)$$

From (4.64)- (4.68) we get that

$$\begin{aligned} \mathcal{E}(e^{n+1}, f^{n+1}) (1 - C_1 \Delta t) & \leq (\Delta t + \Delta t^2) C_0 ((-a) \|\epsilon_1^n\|_{\ell_\Delta^2}^2 - a \|\epsilon_2^n\|_{\ell_\Delta^2}^2 + (-c)d \|D_+ \epsilon_1^n\|_{\ell_\Delta^2}^2) \\ & \quad + (1 + \Delta t C_2) \mathcal{E}(e^n, f^n). \end{aligned}$$

□

4.4.4 The case $d = 0, b > 0$

The case $a = c = d = 0, b > 0$.

In this case, as for the classical Boussinesq system (the case $a = b = c = 0, d > 0$, page 137) we derive estimates for a more general scheme : the following θ scheme where the advection term is discretized according to a convex combination of n and $n + 1$

$$\begin{cases} \frac{1}{\Delta t} (I - bD_+D_-) (\eta_j^{n+1} - \eta_j^n) + (1 - \theta)D_c(u)_j^n + \theta D_c(u)_j^{n+1} + D_c(\eta u)_j^n = 0, \\ \frac{1}{\Delta t} (u_j^{n+1} - u_j^n) + (1 - \theta)D_c(\eta)_j^n + \theta D_c(\eta)_j^{n+1} + \frac{1}{2}D_c(u^2)_j^n = \frac{\tau_2 \Delta x}{2} D_+D_- (\eta)_j^n, \end{cases}$$

with

$$\begin{cases} \tau_2 \Delta t \leq \Delta x (1 - \beta_0), \\ \|u_\Delta^n\|_{\ell^\infty} + \alpha_0 \leq \tau_2, \quad n \in \llbracket 0, N \rrbracket. \end{cases}$$

The convergence error verifies :

$$\begin{cases} (I - bD_+D_-) (E^-)_j^n + \Delta t (1 - \theta) D_c(f)_j^n + \theta \Delta t D_c(f)_j^{n+1} + \Delta t D_c(eu_\Delta)_j^n \\ \quad + \Delta t D_c(\eta_\Delta f)_j^n + \Delta t D_c(ef)_j^n = -\Delta t (\epsilon_1)_j^n, \\ (F^-)_j^n + \Delta t (1 - \theta) D_c(e)_j^n + \theta \Delta t D_c(e)_j^{n+1} + \Delta t D_c(fu_\Delta)_j^n + \frac{\Delta t}{2} D_c(f^2)_j^n \\ \quad = \frac{\tau_2}{2} \Delta t \Delta x D_+D_- (f)_j^n - \Delta t (\epsilon_2)_j^n. \end{cases} \quad (4.70)$$

We work with

$$\mathcal{E}(e, f) := \|e\|_{\ell_\Delta^2}^2 + b \|D_+(e)\|_{\ell_\Delta^2}^2 + \|f\|_{\ell_\Delta^2}^2, \quad (4.71)$$

and the proof of energy estimate (4.41) is as follows.

Proof of Proposition 4.31 in the case $a = c = d = 0, b > 0$. We multiply the first equation by $(E^+)_j^n$. The left hand side corresponds to (4.27) with $c = 0$. We obtain

$$\begin{aligned} & \left(\|e^{n+1}\|_{\ell^2}^2 + b \|D_+(e)^{n+1}\|_{\ell^2}^2 \right) - \left(\|e^n\|_{\ell^2}^2 + b \|D_+(e)^n\|_{\ell^2}^2 \right) \\ &= -\Delta t \Delta x \sum_{j \in \mathbb{Z}} \left[(1 - \theta) D_c(f)_j^n + \theta D_c(f)_j^{n+1} + D_c(eu_\Delta)_j^n + D_c(\eta_\Delta f)_j^n + D_c(ef)_j^n + (\epsilon_1)_j^n \right] (E^+)_j^n. \end{aligned}$$

Integration by parts (3.27) gives

$$\begin{aligned} & \left(\|e^{n+1}\|_{\ell^2}^2 + b \|D_+(e)^{n+1}\|_{\ell^2}^2 \right) - \left(\|e^n\|_{\ell^2}^2 + b \|D_+(e)^n\|_{\ell^2}^2 \right) \\ &= -\Delta t \Delta x \sum_{j \in \mathbb{Z}} (\epsilon_1)_j^n (E^+)_j^n + \Delta t \Delta x \sum_{j \in \mathbb{Z}} \left[(1 - \theta) f_j^n + \theta f_j^{n+1} + e_j^n (u_\Delta)_j^n + (\eta_\Delta)_j^n f_j^n + e_j^n f_j^n \right] D_c(E^+)_j^n \\ &\leq \Delta t \|\epsilon_1^n\|_{\ell_\Delta^2}^2 + \Delta t C_1 \mathcal{E}(e^{n+1}, f^{n+1}) + \Delta t C_2 \mathcal{E}(e^n, f^n), \end{aligned} \quad (4.72)$$

with C_1 and C_2 two constants proportional to

$$1 + \frac{1}{b} \max \{1, \|u_\Delta^n\|_{\ell^\infty}, \|\eta_\Delta^n\|_{\ell^\infty}\}.$$

Using the second equation of (4.70), the triangle inequality along with Proposition 4.32 we get that :

$$\begin{aligned} \|f^{n+1}\|_{\ell^2} &\leq \left\| f^n - \Delta t D_c(fu_\Delta)^n - \frac{\Delta t}{2} D_c(f^2)^n + \frac{\tau_2}{2} \Delta t \Delta x D_+D_- (f)^n \right\|_{\ell^2} \\ &\quad + \Delta t \|\epsilon_2^n\|_{\ell_\Delta^2} + \Delta t (1 - \theta) \|D_c(e)^n\|_{\ell_\Delta^2} + \Delta t \theta \|D_c(e)^{n+1}\|_{\ell_\Delta^2} \\ &\leq \Delta t \|\epsilon_2^n\|_{\ell_\Delta^2} + (1 + C \Delta t) \|f^n\|_{\ell_\Delta^2} + \Delta t (1 - \theta) \|D_c(e)^n\|_{\ell_\Delta^2} + \Delta t \theta \|D_c(e)^{n+1}\|_{\ell_\Delta^2}. \end{aligned} \quad (4.73)$$

Thus, by adding up estimate (4.72) with the square of the estimate (4.73) we get that

$$(1 - C_3 \Delta t) \mathcal{E}(e^{n+1}, f^{n+1}) \leq \Delta t \|\epsilon_1^n\|_{\ell_\Delta^2}^2 + (\Delta t + \Delta t^2) C_0 \|\epsilon_2^n\|_{\ell_\Delta^2}^2 + (1 + C_4 \Delta t) \mathcal{E}(e^n, f^n).$$

□

The case $a < 0, b > 0, c = d = 0$.

The convergence error satisfies :

$$\begin{cases} (I - bD_+D_-) e_j^{n+1} + \Delta t D_c(f)_j^{n+1} + a\Delta t D_+D_-D_c(f)_j^{n+1} = (I - bD_+D_-) e_j^n \\ \quad - \Delta t D_c(eu_\Delta)_j^n - \Delta t D_c(ef)_j^n - \Delta t D_c(\eta_\Delta f)_j^n - \Delta t (\epsilon_1)_j^n, \\ f_j^{n+1} + \Delta t D_c(e)_j^{n+1} = f_j^n - \Delta t D_c\left(f\left(u_\Delta + \frac{1}{2}f\right)\right)_j^n + \frac{\tau_2}{2}\Delta t \Delta x D_+D_- (f)_j^n - \Delta t (\epsilon_2)_j^n, \end{cases} \quad (4.74)$$

with

$$\begin{cases} \tau_2 \Delta t \leq \Delta x (1 - \beta_0), \\ \|u_\Delta^n\|_{\ell^\infty} + \alpha_0 \leq \tau_2, \quad n \in \llbracket 0, N \rrbracket. \end{cases}$$

Consider the energy functional we will work with

$$\mathcal{E}(e, f) := \|e\|_{\ell_\Delta^2}^2 + b \|D_+e\|_{\ell_\Delta^2}^2 + \|f\|_{\ell_\Delta^2}^2 + (-a) \|D_+f\|_{\ell_\Delta^2}^2.$$

Proof of Proposition 4.31 in the case $a < 0, b > 0, c = d = 0$. Let us multiply the first equation of (4.74) with $2e_j^{n+1}$ to obtain :

$$\begin{aligned} & 2 \|e^{n+1}\|_{\ell_\Delta^2}^2 + 2b \|D_+(e)^{n+1}\|_{\ell_\Delta^2}^2 + 2\Delta t \Delta x \sum_{j \in \mathbb{Z}} D_c(f)_j^{n+1} e_j^{n+1} + 2a\Delta t \Delta x \sum_{j \in \mathbb{Z}} D_+D_-D_c(f)_j^{n+1} e_j^{n+1} \\ &= 2\Delta x \sum_{j \in \mathbb{Z}} (I - bD_+D_-) e_j^n e_j^{n+1} - 2\Delta t \Delta x \sum_{j \in \mathbb{Z}} \left[D_c(eu_\Delta)_j^n + D_c(ef)_j^n + D_c(\eta_\Delta f)_j^n + (\epsilon_1)_j^n \right] e_j^{n+1} \\ &\leq \|e^{n+1}\|_{\ell_\Delta^2}^2 + b \|D_+e^{n+1}\|_{\ell_\Delta^2}^2 + \|e^n\|_{\ell_\Delta^2}^2 + b \|D_+e^n\|_{\ell_\Delta^2}^2 \\ &\quad + 2\Delta t \|D_c(eu_\Delta)_j^n + D_c(ef)_j^n + D_c(\eta_\Delta f)_j^n\|_{\ell_\Delta^2} \|e^{n+1}\|_{\ell_\Delta^2} + \Delta t \|\epsilon_1^n\|_{\ell_\Delta^2}^2 + \Delta t \|e^{n+1}\|_{\ell_\Delta^2}^2. \end{aligned} \quad (4.75)$$

The last inequality is due to Cauchy-Schwarz and Young inequalities.

Notice that

$$\|D_c(eu_\Delta)_j^n + D_c(ef)_j^n + D_c(\eta_\Delta f)_j^n\|_{\ell_\Delta^2} \leq C_1 \mathcal{E}^{\frac{1}{2}}(e^n, f^n),$$

with C_1 proportional to

$$\max \{ \|u_\Delta^n\|_{\ell^\infty}, \|D_c(u_\Delta)_j^n\|_{\ell^\infty}, \|\eta_\Delta^n\|_{\ell^\infty}, \|D_c(\eta_\Delta)_j^n\|_{\ell^\infty} \} \max \left\{ \frac{1}{b}, \frac{1}{-a}, 1 \right\}.$$

Thus, Equation (4.75) becomes

$$\begin{aligned} & \|e^{n+1}\|_{\ell_\Delta^2}^2 + b \|D_+(e)^{n+1}\|_{\ell_\Delta^2}^2 + 2\Delta t \Delta x \sum_{j \in \mathbb{Z}} D_c(f)_j^{n+1} e_j^{n+1} + 2a\Delta t \Delta x \sum_{j \in \mathbb{Z}} D_+D_-D_c(f)_j^{n+1} e_j^{n+1} \\ &\leq \|e^n\|_{\ell_\Delta^2}^2 + b \|D_+e^n\|_{\ell_\Delta^2}^2 + C_1 \Delta t \mathcal{E}^{\frac{1}{2}}(e^n, f^n) \|e^{n+1}\|_{\ell_\Delta^2} + \Delta t \|\epsilon_1^n\|_{\ell_\Delta^2}^2 + \Delta t \|e^{n+1}\|_{\ell_\Delta^2}^2. \end{aligned} \quad (4.76)$$

Next, taking the square of the ℓ_Δ^2 -norm of the second equation of (4.74) and using Proposition 4.33, we get that

$$\begin{aligned} & \|f^{n+1}\|_{\ell_\Delta^2}^2 + 2\Delta t \Delta x \sum_{j \in \mathbb{Z}} f_j^{n+1} D_c(e)_j^{n+1} + \Delta t^2 \|D_c(e)^{n+1}\|_{\ell_\Delta^2}^2 \\ &\leq (\Delta t + \Delta t^2) C_0 \|\epsilon_2^n\|_{\ell_\Delta^2}^2 + (1 + C_0 \Delta t) \left\| f^n - \Delta t D_c\left(f\left(u_\Delta + \frac{1}{2}f\right)\right)_j^n + \frac{\tau_2}{2}\Delta t \Delta x D_+D_- (f)_j^n \right\|_{\ell_\Delta^2}^2 \\ &\leq (\Delta t + \Delta t^2) C_0 \|\epsilon_2^n\|_{\ell_\Delta^2}^2 + (1 + C_2 \Delta t) \|f^n\|_{\ell_\Delta^2}^2. \end{aligned} \quad (4.77)$$

When we will sum up Equations (4.76) and (4.77), both terms

$$\sum_{j \in \mathbb{Z}} D_c(f)_j^{n+1} e_j^{n+1} \quad \text{and} \quad \sum_{j \in \mathbb{Z}} f_j^{n+1} D_c(e)_j^{n+1}$$

will cancel each other, thanks to Relation (3.28). We have to cancel $\sum_{j \in \mathbb{Z}} D_+ D_- D_c(f)_j^{n+1} e_j^{n+1}$ in (4.76) too. This is the aim of the following computation.

Applying $\sqrt{-a}D_+$ into the second equation of (4.74) and taking the square of the ℓ_Δ^2 -norm of the resulting equation yields (with Proposition 4.34)

$$\begin{aligned} & (-a) \left\| D_+(f)^{n+1} \right\|_{\ell_\Delta^2}^2 + 2(-a) \Delta t \Delta x \sum_{j \in \mathbb{Z}} D_+(f)_j^{n+1} D_+ D_c(e)_j^{n+1} + (-a) \Delta t^2 \left\| D_+ D_c(e)^{n+1} \right\|_{\ell_\Delta^2}^2 \\ & \leq (\Delta t + \Delta t^2) C_0 (-a) \left\| D_+(\epsilon_2)^n \right\|_{\ell_\Delta^2}^2 \\ & + (1 + C_0 \Delta t) (-a) \left\| D_+(f)^n - \Delta t D_+ D_c \left(f \left(u_\Delta + \frac{1}{2} f \right) \right)^n + \frac{\tau_2}{2} \Delta t \Delta x D_+ D_- D_c(f)^n \right\|_{\ell_\Delta^2}^2 \\ & \leq (\Delta t + \Delta t^2) C_0 (-a) \left\| D_+(\epsilon_2)^n \right\|_{\ell_\Delta^2}^2 + (1 + C_3 \Delta t) (-a) \left\| D_+(f)^n \right\|_{\ell_\Delta^2}^2 + \Delta t C_4 (-a) \left\| f^n \right\|_{\ell_\Delta^2}^2. \end{aligned} \quad (4.78)$$

Adding up the estimates (4.76), (4.77) and (4.78) leads to

$$\begin{aligned} (1 - C_5 \Delta t) \mathcal{E}(e^{n+1}, f^{n+1}) & \leq \Delta t \left\| \epsilon_1^n \right\|_{\ell_\Delta^2}^2 + (\Delta t + \Delta t^2) C_0 \left(\left\| \epsilon_2^n \right\|_{\ell_\Delta^2}^2 + (-a) \left\| D_+ \epsilon_2^n \right\|_{\ell_\Delta^2}^2 \right) \\ & + (1 + C_6 \Delta t) \mathcal{E}(e^n, f^n). \end{aligned}$$

□

The case $a < 0, b > 0, c < 0, d = 0$.

Finally, in this case, the convergence error satisfies :

$$\left\{ \begin{array}{l} (I - bD_+ D_-) e_j^{n+1} + \Delta t D_c(f)_j^{n+1} + a \Delta t D_+ D_- D_c(f)_j^{n+1} = (I - bD_+ D_-) e_j^n \\ \quad - \Delta t D_c(eu_\Delta)_j^n - \Delta t D_c(ef)_j^n - \Delta t D_c(\eta_\Delta f)_j^n - \Delta t (\epsilon_1)_j^n, \\ f_j^{n+1} + \Delta t D_c(e)_j^{n+1} + c \Delta t D_+ D_- D_c(e)_j^{n+1} = f_j^n - \Delta t D_c \left(f \left(u_\Delta + \frac{1}{2} f \right) \right)_j^n \\ \quad + \frac{\tau_2}{2} \Delta t \Delta x D_+ D_- (f)_j^n - \Delta t (\epsilon_2)_j^n, \end{array} \right. \quad (4.79)$$

with

$$\left\{ \begin{array}{l} \tau_2 \Delta t \leq \Delta x (1 - \beta_0), \\ \|u_\Delta^n\|_{\ell^\infty} + \alpha_0 \leq \tau_2, \quad n \in \llbracket 0, N \rrbracket. \end{array} \right.$$

In this case, we will use the energy functional :

$$\mathcal{E}(e, f) := (-c) \left\| (I - bD_+ D_-) e \right\|_{\ell_\Delta^2}^2 + \|f\|_{\ell_\Delta^2}^2 + (-a) b \left\| D_+(f) \right\|_{\ell_\Delta^2}^2.$$

Proof of Proposition 4.31 in the case $a < 0, b > 0, c < 0, d = 0$. Taking the square of the first equations of (4.79) and multiplying the result with $(-c)$ yields

$$\begin{aligned} \|\text{LHS}_1^{d=0}\|_{\ell_\Delta^2}^2 & := (-c) \left\| (I - bD_+ D_-) e^{n+1} \right\|_{\ell_\Delta^2}^2 + (-c) \Delta t^2 \left\| D_c(f)^{n+1} \right\|_{\ell_\Delta^2}^2 \\ & + (-c) a^2 \Delta t^2 \left\| D_+ D_- D_c(f)^{n+1} \right\|_{\ell_\Delta^2}^2 + 2ac \Delta t^2 \left\| D_+ D_c(f)^{n+1} \right\|_{\ell_\Delta^2}^2 \\ & - 2c \Delta t \Delta x \sum_{j \in \mathbb{Z}} (I - bD_+ D_-) e_j^{n+1} D_c(f)_j^{n+1} - 2ac \Delta t \Delta x \sum_{j \in \mathbb{Z}} e_j^{n+1} D_+ D_- D_c(f)_j^{n+1} \\ & + 2acb \Delta t \Delta x \sum_{j \in \mathbb{Z}} D_+ D_- (e)_j^{n+1} D_+ D_- D_c(f)_j^{n+1}. \end{aligned} \quad (4.80)$$

Taking the square of the ℓ_Δ^2 -norm of the second equation of (4.79), we get that

$$\begin{aligned} \|\text{LHS}_2^{d=0}\|_{\ell_\Delta^2}^2 &:= \|f^{n+1}\|_{\ell_\Delta^2}^2 + \Delta t^2 \|D_c(e)^{n+1}\|_{\ell_\Delta^2}^2 + (c\Delta t)^2 \|D_+ D_- D_c(e)^{n+1}\|_{\ell_\Delta^2}^2 - 2c\Delta t^2 \|D_+ D_c(e)^{n+1}\|_{\ell_\Delta^2}^2 \\ &\quad + 2\Delta t\Delta x \sum_{j \in \mathbb{Z}} f_j^{n+1} D_c(e)_j^{n+1} + 2c\Delta t\Delta x \sum_{j \in \mathbb{Z}} f_j^{n+1} D_+ D_- D_c(e)_j^{n+1}. \end{aligned} \quad (4.81)$$

Let us observe that the terms

$$2\Delta t\Delta x \sum_{j \in \mathbb{Z}} f_j^{n+1} D_c(e)_j^{n+1} \quad \text{and} \quad 2c\Delta t\Delta x \sum_{j \in \mathbb{Z}} f_j^{n+1} D_+ D_- D_c(e)_j^{n+1}, \quad \text{in (4.81)}$$

and

$$-2ac\Delta t\Delta x \sum_{j \in \mathbb{Z}} e_j^{n+1} D_+ D_- D_c(f)_j^{n+1} \quad \text{and} \quad -2c\Delta t\Delta x \sum_{j \in \mathbb{Z}} (I - bD_+ D_-) e_j^{n+1} D_c(f)_j^{n+1}, \quad \text{in (4.80)}$$

can be lower bounded by the energy $-\Delta t C \mathcal{E}(e^{n+1}, f^{n+1})$ (by using integrations by parts (3.27) and Cauchy-Schwarz inequality) such that they do not raise any issues. In order to get rid of the term

$$2acb\Delta t\Delta x \sum_{j \in \mathbb{Z}} D_+ D_- (e)_j^{n+1} D_+ D_- D_c(f)_j^{n+1} \quad \text{in (4.80)} \quad (4.82)$$

we will apply $\sqrt{(-a)b}D_+$ into the second equation of (4.79) and consider the square of the ℓ_Δ^2 -norm of the result. First, let us compute :

$$\begin{aligned} \|\text{LHS}_3^{d=0}\|_{\ell_\Delta^2}^2 &:= (-a)b \left\| D_+(f)^{n+1} + \Delta t D_+ D_c(e)^{n+1} + c\Delta t D_+ D_+ D_- D_c(e)^{n+1} \right\|_{\ell_\Delta^2}^2 \\ &= (-a)b \left\| D_+(f)^{n+1} \right\|_{\ell_\Delta^2}^2 + (-a)b\Delta t^2 \left\| D_+ D_c(e)^{n+1} \right\|_{\ell_\Delta^2}^2 + (-a)b(c\Delta t)^2 \left\| D_+ D_+ D_- D_c(e)^{n+1} \right\|_{\ell_\Delta^2}^2 \\ &\quad + 2(-a)b\Delta t\Delta x \sum_{j \in \mathbb{Z}} D_+(f)_j^{n+1} D_+ D_c(e)_j^{n+1} + 2(-a)bc\Delta t\Delta x \sum_{j \in \mathbb{Z}} D_+(f)_j^{n+1} D_+ D_+ D_- D_c(e)_j^{n+1} \\ &\quad + 2abc\Delta t^2 \left\| D_+ D_- D_c(e)^{n+1} \right\|_{\ell_\Delta^2}^2. \end{aligned} \quad (4.83)$$

Of course the problematic term from (4.82) will cancel with

$$-2abc\Delta t\Delta x \sum_{j \in \mathbb{Z}} D_+(f)_j^{n+1} D_+ D_+ D_- D_c(e)_j^{n+1}$$

appearing in (4.83). The additional term $2(-a)b\Delta t\Delta x \sum_{j \in \mathbb{Z}} D_+(f)_j^{n+1} D_+ D_c(e)_j^{n+1}$ in (4.83) can be once again lower bounded by the energy $-\Delta t C \mathcal{E}(e^{n+1}, f^{n+1})$ (by using integrations by parts (3.27) and Cauchy-Schwarz inequality).

Let us now interest to the right hand side. For the first equation of (4.79), we obtain by Young inequality

$$\|\text{RHS}_1^{d=0}\|_{\ell_\Delta^2}^2 \leq (-c)(\Delta t + \Delta t^2) C_0 \|\epsilon_1^n\|_{\ell_\Delta^2}^2 + (1 + C_0\Delta t)(-c) \|(I - bD_+ D_-) e^n\|_{\ell_\Delta^2}^2 + C_1\Delta t \mathcal{E}(e^n, f^n).$$

In the previous inequality, we have upper bounded $\Delta t \| -D_c(eu_\Delta)^n - D_c(ef)^n - D_c(\eta_\Delta f)^n \|_{\ell_\Delta^2}^2$ by

$$\max \{ \|u_\Delta^n\|_{\ell^\infty}, \|D_c(u_\Delta)^n\|_{\ell^\infty}, \|\eta_\Delta^n\|_{\ell^\infty}, \|D_c(\eta_\Delta)^n\|_{\ell^\infty}, \} \max \left\{ \frac{1}{(-a)b}, \frac{1}{2b(-c)} \right\} \mathcal{E}(e^n, f^n).$$

The second equation of (4.79) gives

$$\begin{aligned} \|\text{RHS}_2^{d=0}\|_{\ell_\Delta^2}^2 &\leq (\Delta t + \Delta t^2) C_0 \|\epsilon_2^n\|_{\ell_\Delta^2}^2 + (1 + C_0 \Delta t) \left\| f^n - \Delta t D_c \left(f \left(u_\Delta + \frac{1}{2} f \right) \right)^n + \frac{\tau_2}{2} \Delta t \Delta x D_+ D_- (f)^n \right\|_{\ell_\Delta^2}^2 \\ &\leq (\Delta t + \Delta t^2) C_0 \|\epsilon_2^n\|_{\ell_\Delta^2}^2 + (1 + C_1 \Delta t) \|f^n\|_{\ell_\Delta^2}^2. \end{aligned}$$

Moreover, using Proposition (4.34) we get

$$\|\text{RHS}_3^{d=0}\|_{\ell_\Delta^2}^2 \leq (-a) b (\Delta t + \Delta t^2) C_0 \|D_+ \epsilon_2^n\|_{\ell_\Delta^2}^2 + C_2 \Delta t (-a) b \|f^n\|_{\ell_\Delta^2}^2 + (1 + C_3 \Delta t) (-a) b \|D_+ f^n\|_{\ell_\Delta^2}^2.$$

Putting together Estimates (4.80), (4.81) (4.83) and the upper bounds of $\text{RHS}_i^{d=0}$, for $i \in \{1, 2, 3\}$, we get that

$$\begin{aligned} (1 - C_4 \Delta t) \mathcal{E}(e^{n+1}, f^{n+1}) &\leq (\Delta t + \Delta t^2) C_0 (-c) \|\epsilon_1^n\|_{\ell_\Delta^2}^2 + (\Delta t + \Delta t^2) C_0 \left(\|\epsilon_2^n\|_{\ell_\Delta^2}^2 + (-ab) \|D_+ \epsilon_2^n\|_{\ell_\Delta^2}^2 \right) \\ &\quad + (1 + C_5 \Delta t) \mathcal{E}(e^n, f^n). \end{aligned}$$

□

4.5 Experimental results

To illustrate our numerical results, we compare in this section the exact solution and the numerical one, for several values of a , b , c and d . The simulation is done by fixing $[0, L]$ the space domain with $L = 40$. Moreover, we use periodic boundary conditions. Those conditions are not absorbing boundary conditions, which would mimic perfectly the behavior on $\mathbb{Z}$, but we fix the final time T small enough and we take the initial conditions localized enough in order to minimize boundary effects.

The comparisons between the exact solution and the numerical one (Figures 4.2-4.7) are done with the space cell size $\Delta x = \frac{1}{26}$ and a time step $\Delta t = 0.001$. The convergence results (Tables 4.1-4.4) are computed with a number of cells J satisfying $J \in \{640, 1280, 2560, 5120, 10240\}$. In these tables, we choose the final time $T = 2$ and the time step is taken to $\Delta t^n = \frac{\Delta x}{\max_{t \in [0, t^n], x \in [0, L]} |u(t, x)|}$ to verify the CFL-type condition. In the case where $bd = 0$, we choose the Rusanov coefficients τ_i verifying $\tau_1^n \geq \|u_\Delta^n\|_{\ell^\infty}$ and $\tau_2^n \geq \|u_\Delta^n\|_{\ell^\infty}$, when they are needed.

4.5.1 Linear case

First of all, we test the linear case

$$\begin{cases} (I - b \partial_x^2) \partial_t \eta + (I + a \partial_x^2) \partial_x u = 0, \\ (I - d \partial_x^2) \partial_t u + (I + c \partial_x^2) \partial_x \eta = 0, \\ \eta|_{t=0} = \eta_0, u|_{t=0} = u_0. \end{cases}$$

For that case, the discrete energy is conserved when b and d are non zero and for the Crank-Nicolson scheme (as the one we chose previously). We illustrate this conservation in Figure 4.1. We choose $a = -\frac{7}{30}$, $b = \frac{7}{15}$, $c = -\frac{2}{5}$ and $d = \frac{1}{2}$ and initialize the scheme with

$$\begin{cases} \eta_0(x) = \frac{3}{8} \text{sech}^2 \left(\frac{1}{2} \sqrt{\frac{5}{7}} \left(x - \frac{L}{2} \right) \right), \\ u_0(x) = \frac{1}{2\sqrt{2}} \text{sech}^2 \left(\frac{1}{2} \sqrt{\frac{5}{7}} \left(x - \frac{L}{2} \right) \right). \end{cases}$$

Section 4.5. Experimental results

For $\Delta x = \frac{1}{2^5}$, $\Delta t = 0.001$ and nor τ_1 neither τ_2 we obtain the conservation of the discret energy, more or less 10^{-11} , as illustrated in Figure 4.1.

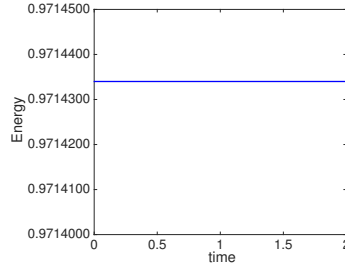


FIGURE 4.1 – Energy conservation in the linear case

4.5.2 Case $b > 0, d > 0$

First we consider the BBM-BBM system ($a = 0, b = \frac{1}{6}, c = 0, d = \frac{1}{6}$) for which two different exact solutions are known. The first of them is obtained in [14] and corresponds to the exact traveling-wave solution of [30] :

$$(A) \begin{cases} \eta(t, x) = \frac{15}{4} \left\{ -2 + \cosh \left(3\sqrt{\frac{2}{5}} \left(x - \frac{L}{2} - \frac{5}{2}t \right) \right) \right\} \operatorname{sech}^4 \left(\frac{3}{\sqrt{10}} \left(x - \frac{L}{2} - \frac{5}{2}t \right) \right), \\ u(t, x) = \frac{15}{2} \operatorname{sech}^2 \left(\frac{3}{\sqrt{10}} \left(x - \frac{L}{2} - \frac{5}{2}t \right) \right). \end{cases}$$

The results are represented in Figure 4.2.

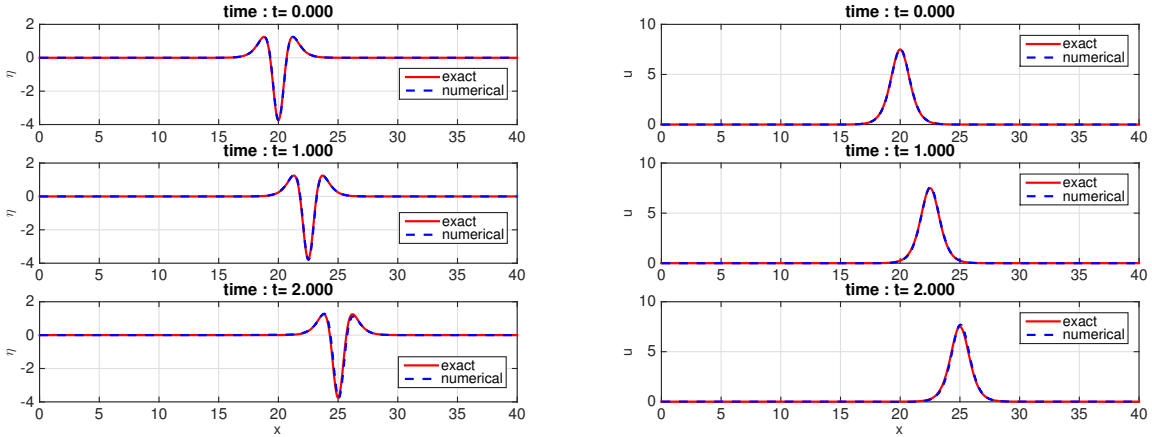
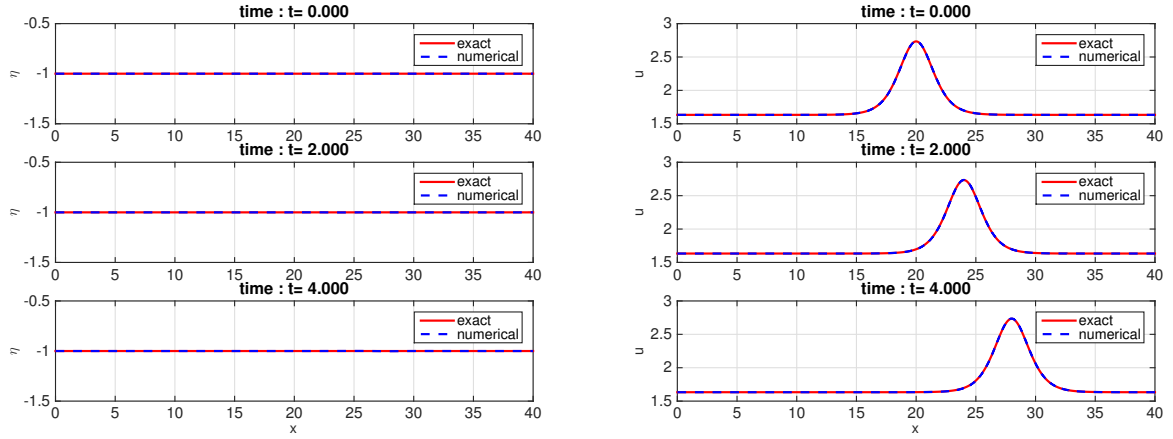


FIGURE 4.2 – Case (A) where $a = 0, b = \frac{1}{6}, c = 0, d = \frac{1}{6}$: results for η (left) and u (right)

Next, we consider the exact solution obtained in [30], whose expression is

$$(B) \begin{cases} \eta(t, x) = -1, \\ u(t, x) = C_s \left(1 - \frac{\rho}{6} \right) + \frac{C_s \rho}{2} \operatorname{sech}^2 \left(\frac{\sqrt{\rho}}{2} \left(x - \frac{L}{2} - C_s t \right) \right). \end{cases}$$

We choose $C_s = 2$ and $\rho = 1.1$ in order to construct the numerical solution. The corresponding result can be found in Figure 4.3.


 FIGURE 4.3 – Case (B) where $a = 0, b = \frac{1}{6}, c = 0, d = \frac{1}{6}$: results for η (left) and u (right)

In order to illustrate our theoretical rate of convergence, we perform the previous examples with increasingly smaller space meshes. The experimental rates are gathered in Table 4.1.

Δx	Case (A)		Case (B)	
	energy error	exp. rate	energy error	exp. rate
$6.25000 \cdot 10^{-2}$	$4.48993 \cdot 10^0$		$8.51815 \cdot 10^{-2}$	
$3.12500 \cdot 10^{-2}$	$2.05132 \cdot 10^0$	1.13270	$4.14750 \cdot 10^{-2}$	1.03830
$1.56250 \cdot 10^{-2}$	$9.80969 \cdot 10^{-1}$	1.06450	$2.04332 \cdot 10^{-2}$	1.02133
$7.81250 \cdot 10^{-3}$	$4.79738 \cdot 10^{-1}$	1.03181	$1.01409 \cdot 10^{-2}$	1.01073
$3.90625 \cdot 10^{-3}$	$2.37234 \cdot 10^{-1}$	1.01580	$5.05189 \cdot 10^{-3}$	1.00529

 TABLE 4.1 – Experimental rates of convergence for $b, d \neq 0$

Another example regards the case $a = -\frac{7}{30}, b = \frac{7}{15}, c = -\frac{2}{5}$ and $d = \frac{1}{2}$ which is discussed in [30]. The exact solution is

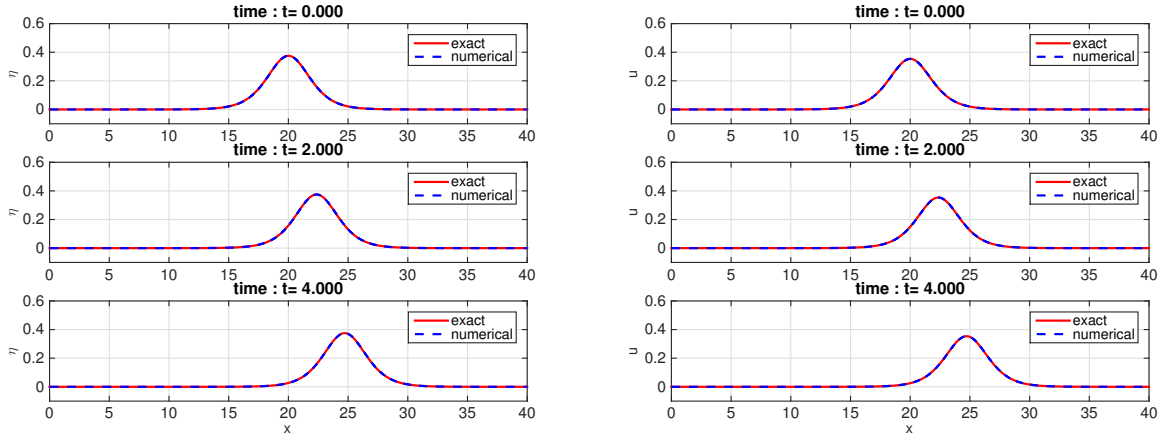
$$(C) \begin{cases} \eta(t, x) = \frac{3}{8} \operatorname{sech}^2 \left(\frac{1}{2} \sqrt{\frac{5}{7}} \left(x - \frac{L}{2} - 5 \frac{\sqrt{2}}{6} t \right) \right), \\ u(t, x) = \frac{1}{2\sqrt{2}} \operatorname{sech}^2 \left(\frac{1}{2} \sqrt{\frac{5}{7}} \left(x - \frac{L}{2} - 5 \frac{\sqrt{2}}{6} t \right) \right). \end{cases}$$

The results are gathered in Figure 4.4 and Table 4.2.

The last example of this paragraph is $a = 0, b = \frac{1}{3}, c = -\frac{1}{3}$ and $d = \frac{1}{3}$, studied in [30]. The exact solution writes

$$(D) \begin{cases} \eta(t, x) = -1, \\ u(t, x) = \left(1 - \frac{\rho}{3}\right) C_s + C_s \rho \operatorname{sech}^2 \left(\frac{\sqrt{\rho}}{2} \left(x - \frac{L}{2} - C_s t \right) \right), \end{cases}$$

with $C_s = 3$ and $\rho = 2$ to obtain Figure 4.5 and Table 4.3.


 FIGURE 4.4 – Case (C) where $a = -\frac{7}{30}$, $b = \frac{7}{15}$, $c = -\frac{2}{5}$, $d = \frac{1}{2}$: results for η (left) and u (right)

Δx	Case (C)	
	energy error	exp. rate
$6.25000 \cdot 10^{-2}$	$2.27860 \cdot 10^{-2}$	1.01692 1.00439 0.97887
$3.12500 \cdot 10^{-2}$	$1.126019 \cdot 10^{-2}$	
$1.56250 \cdot 10^{-2}$	$5.612993 \cdot 10^{-3}$	
$7.81250 \cdot 10^{-3}$	$2.847910 \cdot 10^{-3}$	

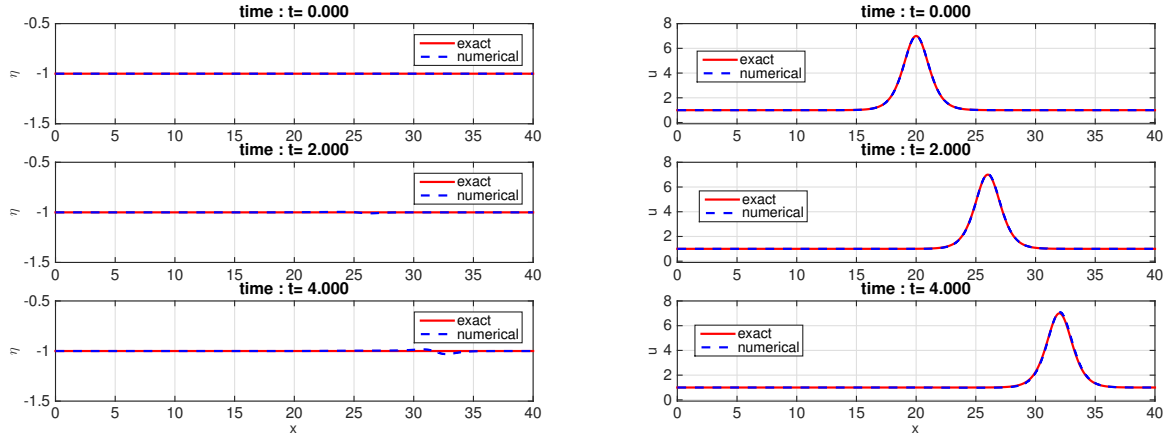
 TABLE 4.2 – Experimental rates of convergence for $a, b, c, d \neq 0$

Δx	Case (D)	
	energy error	exp. rate
$6.25000 \cdot 10^{-2}$	$6.39353 \cdot 10^{-1}$	1.04826 1.02399 1.01195 1.00596
$3.12500 \cdot 10^{-2}$	$3.09159 \cdot 10^{-1}$	
$1.56250 \cdot 10^{-2}$	$1.52031 \cdot 10^{-1}$	
$7.81250 \cdot 10^{-3}$	$7.53884 \cdot 10^{-2}$	
$3.90625 \cdot 10^{-3}$	$3.75388 \cdot 10^{-2}$	

 TABLE 4.3 – Experimental rates of convergence for $b, c, d \neq 0$

Remark 4.35. In these four examples, there is nor viscosity τ_1 neither τ_2 . The terms b and d enable to control and stabilize the scheme, even for the nonlinear Rusanov term, as shown in Figures 4.2-4.5.

With the previous choices of Δt and Δx , we recover the first order convergence, but notice that the schemes would be second order accurate in space and first order in time. When $\Delta t = \Delta x^2$, the second order is well found.

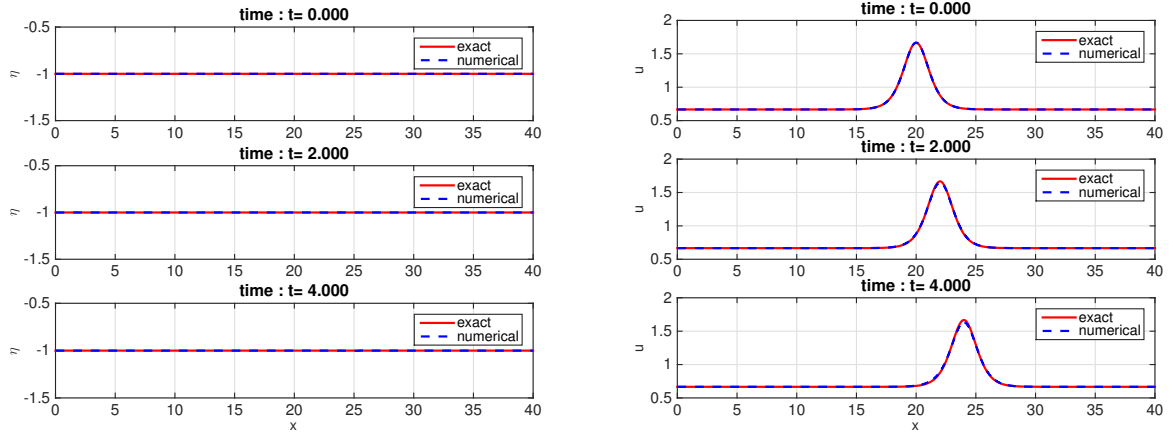

 FIGURE 4.5 – Case (D) where $a = 0, b = \frac{1}{3}, c = -\frac{1}{3}, d = \frac{1}{3}$: results for η (left) and u (right)

4.5.3 Case $bd = 0$

We decide to cancel b by testing two examples. First, we take $a = b = c = 0, d = \frac{1}{6}$. The exact solution is detailed in [30] :

$$(E) \begin{cases} \eta(t, x) = -1, \\ u(t, x) = \left(1 - \frac{\rho}{6}\right) C_s + \frac{C_s \rho}{2} \operatorname{sech}^2 \left(\frac{\sqrt{\rho}}{2} \left(x - \frac{L}{2} - C_s t \right) \right). \end{cases}$$

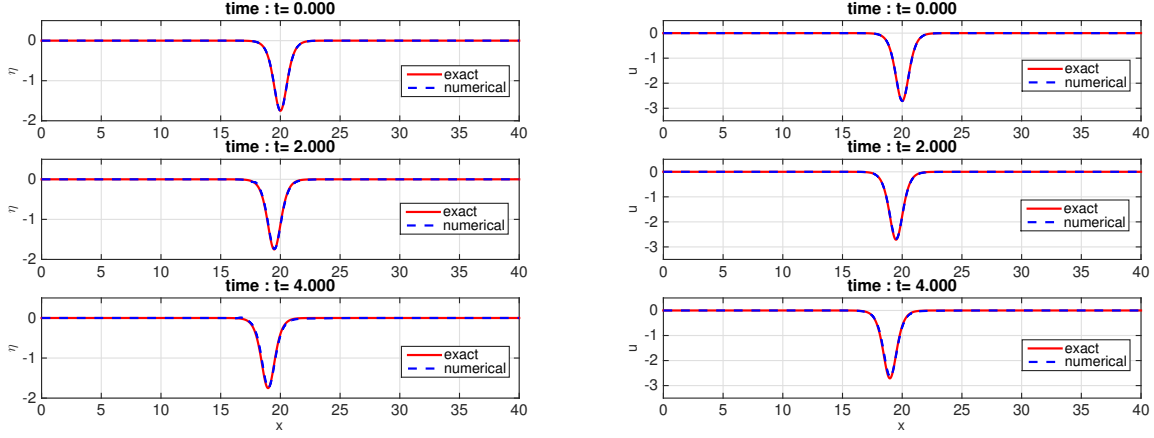
For Figure 4.6, we have fixed $C_s = 1$ and $\rho = 2$.


 FIGURE 4.6 – Case (E) where $a = b = c = 0, d = \frac{1}{6}$: results for η (left) and u (right)

In the second example, we choose the case $a = -\frac{1}{6}, b = c = 0$ and $d = \frac{1}{2}$, for which we know the exact solution developed in [30]

$$(F) \begin{cases} \eta(t, x) = -\frac{7}{4} \operatorname{sech}^2 \left(\frac{\sqrt{7}}{2} \left(x - \frac{L}{2} + \frac{1}{\sqrt{15}} t \right) \right), \\ u(t, x) = -\frac{7}{2} \sqrt{\frac{3}{5}} \operatorname{sech}^2 \left(\frac{\sqrt{7}}{2} \left(x - \frac{L}{2} + \frac{1}{\sqrt{15}} t \right) \right). \end{cases}$$

We obtain Figure 4.7.


 FIGURE 4.7 – Case (F) where $a = -\frac{1}{6}$, $b = c = 0$, $d = \frac{1}{2}$: results for η (left) and u (right)

The experimental rates of convergence in the two cases $b = 0$ are gathered in Table 4.4. In all

Δx	Case (E)		Case (F)	
	energy error	exp. rate	energy error	exp. rate
$6.25000 \cdot 10^{-2}$	$5.94214 \cdot 10^{-2}$		$3.62176 \cdot 10^{-1}$	
$3.12500 \cdot 10^{-2}$	$3.01052 \cdot 10^{-2}$	0.98097	$1.92823 \cdot 10^{-1}$	0.93164
$1.56250 \cdot 10^{-2}$	$1.51573 \cdot 10^{-2}$	0.99000	$1.00366 \cdot 10^{-1}$	0.94780
$7.81250 \cdot 10^{-3}$	$7.60581 \cdot 10^{-3}$	0.99483	$5.16267 \cdot 10^{-2}$	0.95784
$3.90625 \cdot 10^{-3}$	$3.80985 \cdot 10^{-3}$	0.99737	$2.63453 \cdot 10^{-2}$	0.96715

 TABLE 4.4 – Experimental rates of convergence when $b = 0$ and $d > 0$

the previous experimental results, the first order convergence is recovered, which matches to the theoretical results.

4.6 Afterword : traveling waves collision

Recently, in [15], the authors simulate the collision of two traveling waves moving in opposite directions in $[-L, L]$ in the BBM-BBM case ($a = 0, b = \frac{1}{6}, c = 0, d = \frac{1}{6}$). Motivated, by their results, we simulated the same phenomena but for different values of the a, b, c, d parameters.

First, we used our numerical results and performed the same experiment described in [15]. The initial condition is fixed to

$$\begin{cases} \eta(t, x) = \eta_+(t, x) + \eta_-(t, x), \\ u(t, x) = u_+(t, x) + u_-(t, x), \end{cases}$$

with

$$\begin{cases} \eta_{\pm}(t, x) = \frac{15}{2} \operatorname{sech}^2 \left(\frac{3}{\sqrt{10}} \left(x - x_{\pm} \pm \frac{5}{2} t \right) \right) - \frac{45}{4} \operatorname{sech}^4 \left(\frac{3}{\sqrt{10}} \left(x - x_{\pm} \pm \frac{5}{2} t \right) \right), \\ u_{\pm}(t, x) = \mp \frac{15}{2} \operatorname{sech}^2 \left(\frac{3}{\sqrt{10}} \left(x - x_{\pm} \pm \frac{5}{2} t \right) \right), \end{cases}$$

where $x_{\pm} = \pm \frac{L}{2}$.

The space domain is fixed at $[-14, 14]$ and initially, the traveling waves are centered in $x_+ = 7$ and $x_- = -7$. We choose the same space size and time step as in [15], namely $\Delta x = 0.02$ and $\Delta t = 0.00001$. The simulation suggest that blow-up occurs while the explosion time appears to be around $t = 4.5$ as shown in Figure 4.8, result which is very close to the one obtained in [15].

We believe that it is an interesting feature that performing the same kind of experiments for different values of the parameters suggests different qualitative behavior. Indeed, for $(a = -\frac{7}{30}, b = \frac{7}{15}, c = -\frac{2}{5}, d = \frac{1}{2})$, Figure 4.9 suggests that there is no blow-up resulting from the collision of two traveling waves. The experiment is performed for the same value of Δt and Δx as in the previous case. This time, the initial value is

$$\begin{cases} \eta(t, x) = \eta_+(t, x) + \eta_-(t, x), \\ u(t, x) = u_+(t, x) + u_-(t, x), \end{cases}$$

with

$$\begin{cases} \eta_{\pm}(t, x) = \frac{3}{8} \operatorname{sech}^2 \left(\frac{1}{2} \sqrt{\frac{5}{7}} \left(x - x_{\pm} \mp \frac{5\sqrt{2}}{6} t \right) \right), \\ u_{\pm}(t, x) = \pm \frac{1}{2\sqrt{2}} \operatorname{sech}^2 \left(\frac{1}{2} \sqrt{\frac{5}{7}} \left(x - x_{\pm} \mp \frac{5\sqrt{2}}{6} t \right) \right), \end{cases}$$

where $x_{\pm} = \pm \frac{L}{2}$.

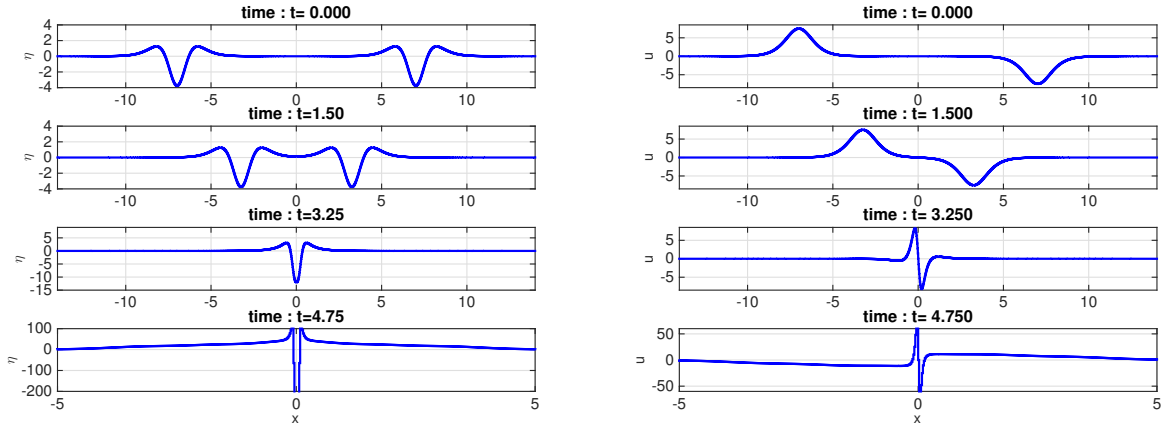


FIGURE 4.8 – Explosion for $(a = 0, b = \frac{1}{6}, c = 0, d = \frac{1}{6})$: results for η (left) and u (right)

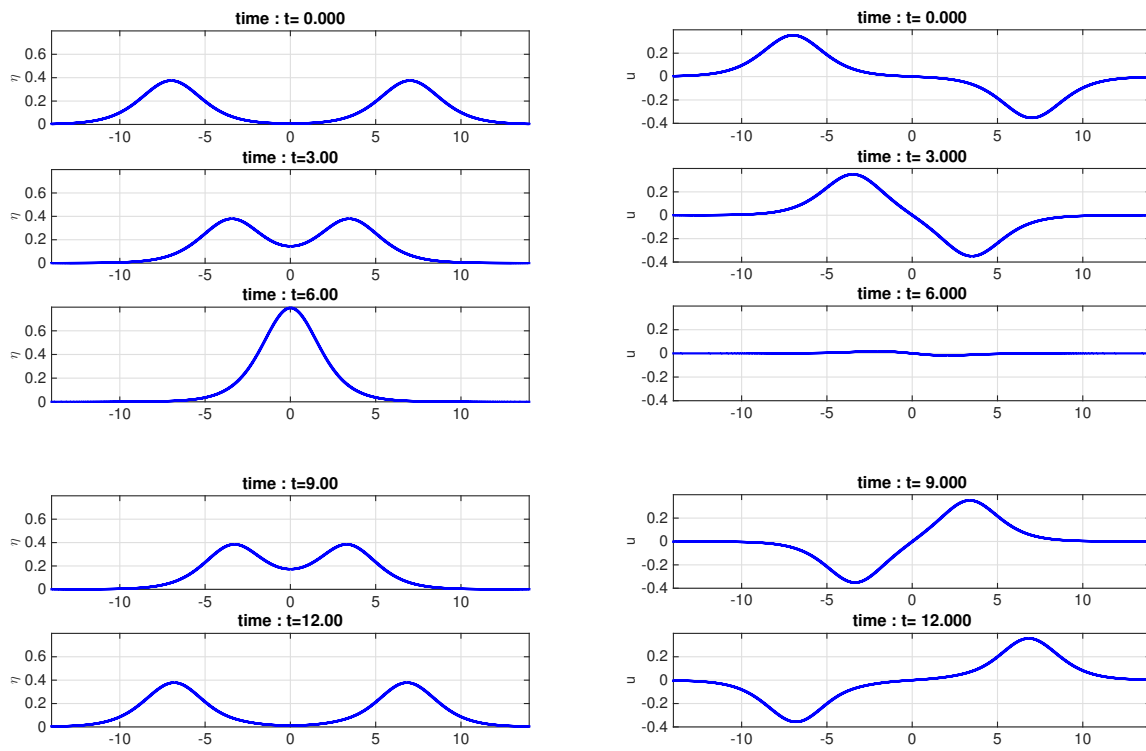


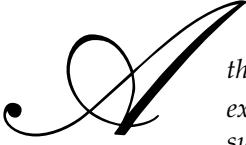
FIGURE 4.9 – No collision for $(a = -\frac{7}{30}, b = \frac{7}{15}, c = -\frac{2}{5}, d = \frac{1}{6})$: results for η (left) and u (right)

QUATRIÈME PARTIE

Étude de l'équation de Korteweg-de Vries-Kuramoto-Sivashinsky

CHAPITRE 5

Small-amplitude traveling waves for KdV-KS with fourth order diffusion

 specific dispersive-diffusive Korteweg-de Vries-Kuramoto-Sivashinsky equation is studied, the one without the second order diffusion : $\partial_t v + \partial_x \left(\frac{v^2}{2} \right) = \mu \partial_x^3 v - \zeta \partial_x^4 v$. We exhibit the existence of small amplitude traveling wave solution. Arguments from dynamical systems such that central manifold theorem, bifurcation theory and Lyapunov function are used. As an aside, a digression is made on the KdV-KS equation without fourth diffusive term, the KdV-Burgers equation : studying traveling waves for this equation could be limited to the study of traveling waves for the well-known Fisher-KPP equation.

Summary

5.1	Introduction (en français)	159
5.2	Speed of propagation and equilibrium points	163
5.3	The extended system and the small amplitude hypothesis	164
5.4	The center manifold reduction	166
5.4.1	Study of the linear part	167
5.4.2	Projection on the center manifold	168
5.4.3	Foliation in α	169
5.4.4	Blowup of the system : a singular change of coordinates	171
5.5	Singular limit system	173
5.6	Perturbed system : a homoclinic bifurcation	174
5.6.1	The Melnikov integral	175
5.6.2	Study of the homoclinic bifurcation	178
5.7	Global «weak» Lyapunov function : elimination of pathogenic cases	179
5.7.1	A global «weak» Lyapunov function	179
5.7.2	The trapping region	181
5.8	Afterword : Traveling waves for KdV-Burgers equation	184
5.8.1	From KdV-Burgers equation to Fisher-KPP equation	185

5.1 Introduction (en français)

Nous nous intéressons dans ce chapitre à montrer l'existence d'onde progressive (ou *traveling wave*) pour l'équation de KdV⁻-KS dégénérée sans diffusion d'ordre deux (i.e l'équation de KdV⁻ avec diffusion d'ordre quatre)

$$\partial_t v + \partial_x \left(\frac{v^2}{2} \right) = \mu \partial_x^3 v - \varsigma \partial_x^4 v. \quad (\text{KdV}^- - 4)$$

Dans la suite, nous considérons une non linéarité plus générale

$$\partial_t v + \partial_x [f(v)] = \mu \partial_x^3 v - \varsigma \partial_x^4 v, \quad (5.1)$$

avec $\mu > 0$ et $\varsigma > 0$ deux constantes et f une fonction convexe au moins de classe $\mathcal{C}^2$ telle que

$$f(0) = f'(0) = 0 \quad \text{et} \quad f''(0) > 0. \quad (5.2)$$

Remarque 5.1. • D'après les hypothèses (5.2), il existe une fonction h au moins de classe $\mathcal{C}^0$ telle que

$$f(x) = x^2 h(x).$$

L'onde progressive que nous considérons a le profil suivant

$$v(t, x) := u \left(\frac{x - ct}{\varsigma^{\frac{1}{3}}} \right) \quad (5.3)$$

avec la vitesse c et le comportement asymptotique

$$\begin{cases} \lim_{\xi \rightarrow -\infty} u(\xi) = u_- = 0 \\ \lim_{\xi \rightarrow +\infty} u(\xi) = u_+ < 0 \end{cases} \quad \text{et} \quad \lim_{\xi \rightarrow \pm\infty} u^{(i)}(\xi) = 0, \quad \text{pour } i \in \{1, 2, 3\}.$$

Remarque 5.2. En réalité, les hypothèses $f(0) = f'(0) = 0$ et $u_- = 0$ ne sont pas restrictives. En effet, étudier les ondes progressives joignant u_- à u_+ à vitesse $c > 0$ pour l'équation

$$\partial_t v + \partial_x [g(v)] = \mu \partial_x^3 v - \varsigma \partial_x^4 v,$$

avec g une fonction $\mathcal{C}^2$ convexe est équivalent à étudier les ondes progressives joignant 0 à $u_+ - u_- < 0$ à vitesse $c - g'(u_-)$ pour l'équation

$$\partial_t v + \partial_x [f(v)] = \mu \partial_x^3 v - \varsigma \partial_x^4 v$$

avec

$$f(\xi) := g(\xi + u_-) - g(u_-) - g'(u_-)\xi.$$

L'hypothèse $f(0) = f'(0) = 0$ est alors bien satisfaite.

Le résultat principal de ce chapitre est le théorème suivant.

Théorème 5.3. Onde progressive pour (5.1)

Supposons que dans l'équation (5.1), f soit convexe telle que $f(0) = f'(0) = 0$ et $f''(0) > 0$. Il existe $\eta > 0$ tel que pour tout $u_+ \in]-\eta, 0[$, il existe une onde progressive (de petite amplitude) joignant 0 à u_+ et se déplaçant à vitesse $c = \frac{f(u_+)}{u_+}$, solution de (5.1).

Section 5.1. Introduction (en français)

Remarque 5.4. Une telle solution est illustrée par exemple en Figure 5.1. Les différents paramètres sont fixés comme suit : $\mu = 0.01$ et $\zeta = \mu^{\frac{3}{2}}$. La condition initiale est égale à la fonction indicatrice : $u_0(x) = \frac{1}{10} \mathbb{1}_{\{x < \frac{L}{8}\}}$. Le domaine $[0, L]$ avec $L = 20$ est considéré avec des conditions aux bords de Dirichlet

$$u(t, 0) \equiv 0.1, \quad \forall t \geq 0 \quad \text{et} \quad u(t, L) \equiv 0, \quad \forall t \geq 0.$$

et le temps final T est fixé à $T = 80$. À partir de $t \approx 40$, l'onde progressive semble atteinte (sous-figure 5.1(a)). En sous-figure 5.1(b), le comportement de l'onde progressive est illustré pour $40 < t < 80$ et sans surprise, le profil semble se traduire dans le sens des x croissants sans changer d'amplitude. La condition initiale ne satisfait pas $u_- = 0$ mais comme expliqué en remarque 5.2, cette hypothèse est juste une simplification de calcul.

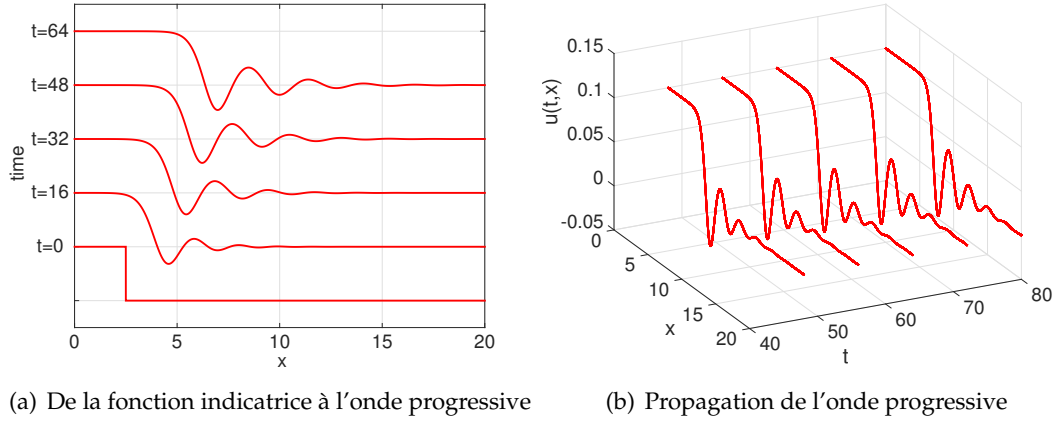


FIGURE 5.1 – Onde progressive pour (KdV⁻ - 4) ($f(x) = \frac{x^2}{2}$) - À gauche : pour $0 < t < 64$, à droite : pour $40 < t < 80$



Variétés stable, instable et centrale. Comme expliqué en introduction générale, étudier l'existence d'onde progressive revient à étudier l'existence d'orbite hétérocline entre deux points d'équilibre. Dans l'espace des phases, l'EDO étudiée s'écrit

$$\frac{d}{d\xi} \mathbf{X}(\xi) = F(\mathbf{X}(\xi)), \quad \mathbf{X} \in \mathbb{R}^n, \quad (5.4)$$

la fonction F dépendant, entre autre, de la vitesse c et de la non linéarité f .

Afin de savoir s'il existe une orbite hétérocline joignant deux points d'équilibre $\mathbf{X}_-$ et $\mathbf{X}_+$, il est important de connaître (au moins qualitativement) le comportement des orbites au voisinage de ces points. Une notion qui sera primordiale dans cette étude est la notion de variétés stable, instable et centrale, dont nous rappelons ici les énoncés (qui peuvent se retrouver dans [52] ou [100])

Remarque 5.5. Soit le système différentiel (5.4) avec F un champ de vecteur régulier. Supposons que $\mathbf{X}_0$ soit un point d'équilibre du système. On notera $\mathbb{E}_{\mathbf{X}_0}^s$, $\mathbb{E}_{\mathbf{X}_0}^u$ et $\mathbb{E}_{\mathbf{X}_0}^c$ respectivement les espaces propres stable, instable et central. C'est-à-dire, les espaces propres associés aux valeurs propres de partie réelle respectivement strictement négative, strictement positive ou nulle de la jacobienne de F en $\mathbf{X}_0$.

Théorème 5.6 (théorème de la variété stable/instable). Soit le système différentiel (5.4). Supposons que $\mathbf{X}_0$ soit un point d'équilibre du système et que F soit de classe C^k , $k \geq 2$ dans un voisinage de $\mathbf{X}_0$. Alors, il existe un voisinage $\mathcal{V}(\mathbf{X}_0) \subset \mathbb{R}^n$ de $\mathbf{X}_0$ et deux applications $\Phi_s : \mathbb{E}_{\mathbf{X}_0}^s \rightarrow \mathbb{E}_{\mathbf{X}_0}^u \oplus \mathbb{E}_{\mathbf{X}_0}^c$ et $\Phi_u : \mathbb{E}_{\mathbf{X}_0}^u \rightarrow \mathbb{E}_{\mathbf{X}_0}^s \oplus \mathbb{E}_{\mathbf{X}_0}^c$, toutes deux de classe C^k , tels que :

- $\Phi_i(\mathbf{X}_0) = \mathbf{0}$, $D\Phi_i(\mathbf{X}_0) = \mathbf{0}$, pour $i \in \{s, u\}$,
 - l'ensemble $\mathcal{W}_{\mathbf{X}_0}^i := \{\mathbf{X}_i + \Phi_i(\mathbf{X}_i), \mathbf{X}_i \in \mathbb{E}_{\mathbf{X}_0}^i\} \cap \mathcal{V}(\mathbf{X}_0)$, pour $i \in \{s, u\}$ est localement invariant par le flot associé à (5.4).
 - Si $\mathbf{X}(0) \in \mathcal{W}_{\mathbf{X}_0}^s$ alors $\mathbf{X}(\xi) \xrightarrow{\xi \rightarrow +\infty} \mathbf{X}_0$. Si $\mathbf{X}(0) \in \mathcal{W}_{\mathbf{X}_0}^u$ alors $\mathbf{X}(\xi) \xrightarrow{\xi \rightarrow -\infty} \mathbf{X}_0$.
 - Si $DF(\mathbf{X}_0)$ n'a pas de valeur propre de partie réelle nulle, alors
- $$\mathcal{W}_{\mathbf{X}_0}^s = \{\mathbf{X}(0) \in \mathcal{V}(\mathbf{X}_0), \mathbf{X}(\xi) \xrightarrow{\xi \rightarrow +\infty} \mathbf{X}_0\} \quad \text{et} \quad \mathcal{W}_{\mathbf{X}_0}^u = \{\mathbf{X}(0) \in \mathcal{V}(\mathbf{X}_0), \mathbf{X}(\xi) \xrightarrow{\xi \rightarrow -\infty} \mathbf{X}_0\}.$$

Théorème 5.7 (théorème de la variété centrale). *Supposons les mêmes hypothèses que pour le théorème 5.6. Supposons de plus que $DF(\mathbf{X}_0)$ a des valeurs propres de partie réelle nulle. Alors, il existe un voisinage $\mathcal{V}(\mathbf{X}_0) \subset \mathbb{R}^n$ de $\mathbf{X}_0$ et une application $\Phi_c : \mathbb{E}_{\mathbf{X}_0}^c \rightarrow \mathbb{E}_{\mathbf{X}_0}^s \oplus \mathbb{E}_{\mathbf{X}_0}^u$ de classe C^k , tels que :*

- $\Phi_c(\mathbf{X}_0) = \mathbf{0}$, $D\Phi_c(\mathbf{X}_0) = \mathbf{0}$,
- l'ensemble $\mathcal{W}_{\mathbf{X}_0}^c := \{\mathbf{X}_c + \Phi_c(\mathbf{X}_c), \mathbf{X}_c \in \mathbb{E}_{\mathbf{X}_0}^c\} \cap \mathcal{V}(\mathbf{X}_0)$ est localement invariant par le flot associé à (5.4).
- Si $\mathbf{X}(\xi) \in \mathcal{V}(\mathbf{X}_0)$, $\forall \xi \in \mathbb{R}$ alors $\mathbf{X}(\xi) \in \mathcal{W}_{\mathbf{X}_0}^c$. Autrement dit, $\mathcal{W}_{\mathbf{X}_0}^c$ contient les solutions bornées et petites dans un voisinage de $\mathbf{X}_0$ du système (5.4).

Définition 5.8. L'ensemble $\mathcal{W}_{\mathbf{X}_0}^s$ est appelé la variété (locale) stable, $\mathcal{W}_{\mathbf{X}_0}^u$ la variété (locale) instable et $\mathcal{W}_{\mathbf{X}_0}^c$ la variété centrale.

Remarque 5.9. • Puisque les variétés sont définies par $\mathcal{W}_{\mathbf{X}_0}^i := \{\mathbf{X}_i + \Phi_i(\mathbf{X}_i), \mathbf{X}_i \in \mathbb{E}_{\mathbf{X}_0}^i\} \cap \mathcal{V}(\mathbf{X}_0)$, pour $i \in \{s, u, c\}$ on dit que $\mathcal{W}_{\mathbf{X}_0}^i$ est tangente à l'espace propre $\mathbb{E}_{\mathbf{X}_0}^i$.
• Les variétés stable et instable sont uniques par contre, il n'y a pas toujours unicité de la variété centrale.

Nous utiliserons principalement le troisième point du théorème de la variété stable/instable : si la trajectoire part d'un point initial situé sur la variété stable, elle converge dans le futur vers le point d'équilibre. Si elle part d'un point initial situé sur la variété instable, c'est qu'elle provient dans le passé du point d'équilibre. Pour ce qui est du théorème de la variété centrale, nous utiliserons principalement le dernier point : les orbites bornées du système global (5.4) sont des images des orbites sur la variété centrale. Connaître la dynamique de la variété centrale permet donc de connaître la dynamique bornée du système global (5.4).



Méthode suivie. Nous détaillons ici la méthode suivie pour montrer l'existence d'une orbite hétérocline de petite amplitude $\epsilon := |u_- - u_+| \ll 1$. Nous rappelons que dans l'espace des phases notre système du premier ordre s'écrit

$$\frac{d}{d\xi} \mathbf{X}(\xi) = F(\mathbf{X}(\xi)), \quad \mathbf{X} \in \mathbb{R}^n$$

et que l'orbite joint les états d'équilibre $\mathbf{X}_-$ et $\mathbf{X}_+$ à vitesse c . Nous projetons ce système sur la variété centrale afin de réduire sa dimension. Il faut toutefois nous assurer que les points d'équilibre $\mathbf{X}_-$ et $\mathbf{X}_+$ appartiennent bien à cette variété centrale.

Dans un premier temps, il est nécessaire de montrer qu'il existe des solutions qui «décollent» du point d'équilibre $\mathbf{X}_-$ (ce qui n'est *a priori* pas évident puisqu'il s'agit justement d'un point d'équilibre). Nous utilisons pour cela la variété instable : les trajectoires situées dessus proviennent dans le passé de $\mathbf{X}_-$. Il nous faut donc prendre une trajectoire telle que

Section 5.1. Introduction (en français)

$X(0)$ appartienne à cette variété instable et, par le théorème de la variété instable, nous avons $\lim_{\xi \rightarrow -\infty} X(\xi) = X_-$.

Il faut ensuite montrer que la trajectoire dans le futur issue de ce $X(0)$ est bien globale (pour que sa limite quand $\xi \rightarrow +\infty$ ait bien un sens). L'existence de cette trajectoire est assurée par le théorème de Cauchy-Lipschitz. Pour montrer sa globalité, nous montrons qu'elle reste confinée dans un domaine «piège» et utilisons le théorème de sortie de tout compact (théorème 5.52).

Il nous reste à construire ce domaine invariant. En nous rappelant que nous nous sommes restreints à l'étude des ondes de petites amplitudes $\epsilon := |u_- - u_+| \ll 1$, nous étudions le système limite correspondant à $\epsilon = 0$. Ce système possède une orbite homocline et nous étudions son devenir lorsque le système est perturbé par ϵ . Le compact invariant qui va piéger la solution est dans ce cas obtenu par la combinaison de l'étude de cette bifurcation d'homocline et par une fonction de Lyapunov globale.

La dernière étape consiste à montrer la convergence vers l'état d'équilibre X_+ . Nous résumons ci-dessous ces étapes.

Étapes clés : preuve de l'existence d'une orbite hétérocline

- Détermination du système autonome de degrés 1 dans l'espace des phases
 - Réduction du système par projection sur la variété centrale
 - Preuve qu'il existe une solution X qui «décolle» de X_- , par le théorème de la variété instable
 - Quitte à étudier les ondes progressives de petite amplitude, étude de l'homocline du système limite singulier (lorsque $u_- = u_+$)
 - Étude de la bifurcation de l'homocline quand l'amplitude $u_- - u_+$ varie
 - Preuve que la solution X est globale par étude des lignes de niveau d'une fonction de Lyapunov globale et par le théorème de sortie de tout compact
 - Convergence de la solution vers l'état d'équilibre X_+
-



Organisation du chapitre. La section 5.2 a pour but d'étudier les points d'équilibre du système et de déterminer ceux qui sont attractifs et ceux qui sont répulsifs. Nous utilisons en section 5.3 la méthode du système augmenté qui consiste à considérer la vitesse de propagation comme une variable du système afin de créer une variété centrale. La projection sur la variété centrale correspond à la section 5.4. Nous montrons en section 5.5 que le système limite correspondant à $u_+ = u_-$ possède une homocline qui correspond à celle de KdV avec temps rétrograde et montrons qu'elle se transforme en une orbite hétérocline pour le système perturbé, cf sections 5.6 et 5.7.

Nous finissons par une petite digression sur l'équation (KdV⁻ - B) en section 5.8. En remarquant que le système autonome de (KdV⁻ - B) dans l'espace des phases correspond à celui de l'équation de Fisher-KPP (sous-section 5.8.1), nous montrons les résultats d'existence d'ondes progressives pour (KdV⁻ - B) déjà connus mais en nous ramenant à l'équation de Fisher-KPP. Là encore, nous remettons à l'échelle les ondes progressives afin que les coefficients de l'EDO dans l'espace des phases soient de l'ordre de l'unité en cherchant une onde progressive sous la forme $v(t, x) = u\left(\frac{x-ct}{\sqrt{\mu}}\right)$.



5.2 Speed of propagation and equilibrium points

The ansatz u necessary satisfies the following EDO

$$-cu'(\xi) + [f(u(\xi))]' = \frac{\mu}{\varsigma^{\frac{2}{3}}}u^{(3)}(\xi) - u^{(4)}(\xi), \quad \xi \in \mathbb{R}.$$

An integration of the previous EDO combined with the asymptotic behavior at $-\infty$ give

$$-cu(\xi) + f(u(\xi)) = \frac{\mu}{\varsigma^{\frac{2}{3}}}u^{(2)}(\xi) - u^{(3)}(\xi), \quad \xi \in \mathbb{R}, \quad (5.5)$$

which gives the following relation for c and f , when $\xi \rightarrow +\infty$

$$c = \frac{f(u_+)}{u_+}. \quad (5.6)$$

Remark 5.10. We recall that c is well-defined since $u_+ < 0$. Notice that in our case ($u_+ < u_-$ and f convex such that $f(0) = f'(0) = 0$ and $f''(0) > 0$), the speed c is negative. Indeed, the convexity of f implies the growth of the function

$$g := x \mapsto \begin{cases} \frac{f(x)}{x}, & \text{if } x \neq 0, \\ f'(0), & \text{if } x = 0. \end{cases}$$

Hypotheses $f'(0) = 0$ and $u_+ < 0$ conclude on the sign of c .

As explained in Section 5.1, the first step is to transform the EDO satisfied by the ansatz u by a first-order autonomous system in the phase space

$$\frac{d}{d\xi} \begin{pmatrix} u \\ u' \\ u'' \end{pmatrix} = \begin{pmatrix} u' \\ u'' \\ \frac{\mu}{\varsigma^{\frac{2}{3}}}u'' + cu - f(u) \end{pmatrix}. \quad (5.7)$$

Remark 5.11. The rescaling by $\varsigma^{\frac{1}{3}}$ in the ansatz (5.3) enables to determine the weights of dispersion with respect to diffusion. Indeed, the resulting term $\frac{\mu}{\varsigma^{\frac{2}{3}}}$ in the autonomous system must be of zero-order so $\mu^3 = \mathcal{O}(\varsigma^2)$. Thereafter, we denote $\Lambda := \frac{\mu}{\varsigma^{\frac{2}{3}}} > 0$ and one has $\Lambda = \mathcal{O}(1)$.

Let us first have a look on the equilibrium points of that autonomous system.

Proposition 5.12. System (5.7) has two equilibrium points $(0, 0, 0)^t$ and $(u_+, 0, 0)^t$.

• The point $(0, 0, 0)^t$ is a saddle point with $\mathcal{W}_{(0,0,0)}^u$ (respectively $\mathcal{W}_{(0,0,0)}^s$) its unstable manifold (resp. stable manifold). Moreover, one has

$$\dim(\mathcal{W}_{(0,0,0)}^u) = 2 \quad \text{and} \quad \dim(\mathcal{W}_{(0,0,0)}^s) = 1.$$

• Whereas the point $(u_+, 0, 0)^t$ is a saddle point with

$$\dim(\mathcal{W}_{(u_+,0,0)}^u) = 1 \quad \text{and} \quad \dim(\mathcal{W}_{(u_+,0,0)}^s) = 2.$$

Proof. • For the point $(0, 0, 0)^t$ the jacobian matrix of the linearized system is $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ c & 0 & \Lambda \end{pmatrix}$ and the characteristic polynomial writes $\chi(x) := -x^3 + \Lambda x^2 + c$. We denote x_i for $i \in \{1, 2, 3\}$, the (real

or complex) roots. Vieta's formulas give relations between these roots and the coefficients of χ :

$$\begin{cases} x_1 + x_2 + x_3 = \Lambda > 0, & (5.8a) \\ x_1x_2 + x_2x_3 + x_1x_3 = 0, & (5.8b) \\ x_1x_2x_3 = c < 0. & (5.8c) \end{cases}$$

Relation (5.8c) implies $x_1, x_2, x_3 \neq 0$.

Let us first suppose that there exist one real root, say x_1 , and two complex conjugate roots, say x_2 and x_3 . Since $x_2x_3 = \Re(x_2)^2 + \Im(x_2)^2 > 0$, Relation (5.8c) implies $x_1 < 0$. Relation (5.8b) gives

$$2x_1\Re(x_2) + \Re(x_2)^2 + \Im(x_2)^2 = 0.$$

Hence, the real part of x_2 (and x_3) must be positive.

In the case where all the roots are real, there are necessary one non positive root, say $x_1 < 0$ and the other roots have the same sign, say x_2 and x_3 , according to (5.8c). Thanks to relation (5.8b), we conclude that

$$x_1(x_2 + x_3) = -x_2x_3 < 0.$$

The sum $x_2 + x_3$ is then positive and we conclude that $x_2 > 0$ and $x_3 > 0$.

In a nutshell, in all cases, there are two roots with positive real part and one root with negative real part, Proposition 5.12 is a straightforward consequence.

- For the point $(u_+, 0, 0)^t$, the jacobian matrix of the linearized system is $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ c - f'(u_+) & 0 & \Lambda \end{pmatrix}$.

We recall that the tangent to a differentiable convex function is always below the function, it yields

$$f(x) \geq f'(u_+)(x - u_+) + f(u_+), \quad \forall x \in \mathbb{R}.$$

Fix $x = 0$ in the previous inequality (and $f(0) = 0$) gives $(f''(0) > 0$ thus the inequality is strict)

$$f(u_+) < f'(u_+)u_+$$

then (since $u_+ < 0$)

$$c - f'(u_+) = \frac{f(u_+)}{u_+} - f'(u_+) > 0.$$

The previous argumentation made for the point $(0, 0, 0)^t$ is thus valid for $(u_+, 0, 0)^t$ provided that we change $c < 0$ into $c - f'(u_+) > 0$. The sign of the roots are so reversed. $\square$

5.3 The extended system and the small amplitude hypothesis

Our aim in this chapter is to prove the existence of an orbit, solution of (5.7), connecting $(0, 0, 0)^t$ to $(u_+, 0, 0)^t$. It is nothing else than proving the intersection between the unstable manifold of $(0, 0, 0)^t$ and the stable manifold of $(u_+, 0, 0)^t$. Thus, we are looking for the intersection of two 2-dimensional surfaces in a 3-dimensional space. Moreover, these 2-dimensional surfaces depend on c and u_+ , then the intersection seems difficult to prove by topological arguments. First of all, we overcome this ordeal by bringing back the system to a well-known situation : the situation near $(0, 0, 0)^t$. This is why we consider only small amplitude traveling waves which correspond to limiting our study to the orbits near $(0, 0, 0)^t$.

The second idea to overcome this issue is to consider the following extended (or augmented) system

$$\frac{d}{d\tilde{\zeta}} \begin{pmatrix} u \\ u' \\ u'' \\ c \end{pmatrix} = \begin{pmatrix} u' \\ u'' \\ \Lambda u'' + cu - f(u) \\ 0 \end{pmatrix} \quad (5.9)$$

instead of System (5.7). The «extended system» method is an usual hint in fork theory, for example in [52] or in the study of homoclinic bifurcations in [71]. Studying the existence of the ansatz u is then equivalent to studying the existence of an heteroclinic orbit connecting $(0, 0, 0, c)^t$ to $(u_+, 0, 0, c)^t$, solution of the extended system (5.9).

Thereafter, we will focus our attention only on small amplitude traveling waves, which means $|u_+ - u_-| \ll 1$. Our choice of $u_- = 0$ implies $|u_+| \ll 1$. This condition feeds back into the speed of the traveling wave. Indeed, Relation (5.6) together with the small amplitude hypothesis imply

$$|c| = \frac{f''(0)}{2}|u_+| + \mathcal{O}(u_+^2) \ll 1.$$

For the case of (KdV⁻-4) equation, the non linearity is $f(u) = \frac{u^2}{2}$ then $c = \frac{u_+}{2}$ and we immediately see that a small amplitude $|u_+| \ll 1$ implies a small speed of propagation $|c| \ll 1$. The choice of a small amplitude traveling wave (and thus a small traveling speed) enables us to deal only with local orbit near $(0, 0, 0, 0)^t$. This observation will be essential and spearheads the reduction on the center manifold (cf Section 5.4).

Hereafter, we note

$$\bar{\epsilon} := |c| = -c > 0,$$

and introduce a small parameter $\epsilon > 0$. The extended system will be studied in the phase space $\mathbb{R}^3 \times (-\eta, 0)$ where $\eta > 0$ and $(-\eta, 0)$ is a neighborhood of 0 in $\mathbb{R}^-$ and rewrites

$$\frac{d}{d\tilde{\zeta}} \begin{pmatrix} u \\ u' \\ u'' \\ -\epsilon \end{pmatrix} = \begin{pmatrix} u' \\ u'' \\ \Lambda u'' - \epsilon u - f(u) \\ 0 \end{pmatrix} \quad (5.10)$$

Proposition 5.13. *There exists $(-\eta_1, \eta_1) \subset \mathbb{R}$ a neighborhood of 0 with $\eta_1 > 0$, such that, for $\sigma \in (-\eta_1, \eta_1)$, there exists a unique $\mathcal{X}(\sigma) \in \mathbb{R}$ such that $\frac{f(\mathcal{X}(\sigma))}{\mathcal{X}(\sigma)} = \sigma$. Moreover, it holds*

- $\mathcal{X}(0) = 0$, and $\lim_{\sigma \rightarrow 0} \frac{\mathcal{X}(\sigma)}{\sigma} = \frac{1}{h(0)}$, with h introduced in Remark 5.1,
- if $\sigma < 0$, $\mathcal{X}(\sigma) < 0$,
- if c defined by (5.6) belongs to $(-\eta_1, \eta_1)$, $\mathcal{X}(c) = u_+$.

Proof. • We reuse the function g introduced in Remark 5.10

$$g := \mathbb{R}^* \rightarrow \mathbb{R} \\ x \mapsto \frac{f(x)}{x}$$

This function is extended by continuity to $\mathbb{R}$ by defining $g(0) = f'(0) = 0$ and $g'(0) = \frac{f''(0)}{2} \neq 0$ and is so of class $\mathcal{C}^1$. Since f is a $\mathcal{C}^2$ function, strictly convex in a neighborhood of 0. Thus, g is an increasing function locally around 0, say in $(-\eta_1, \eta_1)$, with $\eta_1 > 0$.

By the bijection theorem applied on $(-\eta_1, \eta_1)$, there exists a unique $\mathcal{X}(\sigma) \in \mathbb{R}$ such that $g(\mathcal{X}(\sigma)) = \frac{f(\mathcal{X}(\sigma))}{\mathcal{X}(\sigma)} = \sigma$ for $\sigma \in (-\eta_1, \eta_1)$. Since $g(0) = 0$, then by unicity of $\mathcal{X}(0)$, one has $\mathcal{X}(0) = 0$.

• Moreover, the function h introduced in Remark 5.1 verifies $\mathcal{X}(\sigma)h(\mathcal{X}(\sigma)) = \sigma$. It gives, for all $\sigma \neq 0$, $h(\mathcal{X}(\sigma)) \neq 0$ and

$$\frac{\mathcal{X}(\sigma)}{\sigma} = \frac{1}{h(\mathcal{X}(\sigma))}.$$

The previous relation enables to compute $\lim_{\sigma \rightarrow 0} \frac{\mathcal{X}(\sigma)}{\sigma} = \frac{1}{h(0)}$ (it is well-defined since $h(0) = \frac{f''(0)}{2} \neq 0$).

Section 5.4. The center manifold reduction

- Furthermore, $\sigma < 0$ means $g(\mathcal{X}(\sigma)) < g(0)$ and the increasing of g in $(-\eta_1, \eta_1)$ leads to $\mathcal{X}(\sigma) < 0$.
- To conclude, if $c \in (-\eta_1, \eta_1)$, Equation (5.6) of c and the unicity of $\mathcal{X}(c)$ imply $\mathcal{X}(c) = u_+$. $\square$

Remark 5.14. For the (KdV⁻-4) equation, the non linearity is $f(x) = \frac{x^2}{2}$ and the existence of the function $\mathcal{X}$ is trivial since $\mathcal{X}(\sigma) = 2\sigma$.

Henceforth, we fix η small enough to have $(-\eta, 0) \subset (-\eta_1, \eta_1)$.

Proposition 5.15. The extended system (5.10) has two types of equilibrium points : $(0, 0, 0, \sigma)^t$ and $(\mathcal{X}(\sigma), 0, 0, \sigma)^t$ with $\sigma \in (-\eta, 0)$. Furthermore, we can report the dimensions of the unstable, stable and center manifolds for each points

$$\dim(\mathcal{W}_{(0,0,0,\sigma)}^u) = 2, \quad \dim(\mathcal{W}_{(0,0,0,\sigma)}^s) = 1 \quad \text{and} \quad \dim(\mathcal{W}_{(0,0,0,\sigma)}^c) = 1,$$

and

$$\dim(\mathcal{W}_{(\mathcal{X}(\sigma),0,0,\sigma)}^u) = 1, \quad \dim(\mathcal{W}_{(\mathcal{X}(\sigma),0,0,\sigma)}^s) = 2 \quad \text{and} \quad \dim(\mathcal{W}_{(\mathcal{X}(\sigma),0,0,\sigma)}^c) = 1.$$

Proof. • The jacobian matrix of $(0, 0, 0, \sigma)^t$ is $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sigma & 0 & \Lambda & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, with the following characteristic

polynomial $\chi(x) := -x[-x^3 + \Lambda x^2 + \sigma]$. In the brackets, we recognize the characteristic polynomial of the point $(0, 0, 0)^t$ for System (5.7). We have already proved that there are two eigenvalues with a positive real part and one negative real part eigenvalue.

• The jacobian matrix of $(\mathcal{X}(\sigma), 0, 0, \sigma)^t$ is $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sigma - f'(\mathcal{X}(\sigma)) & 0 & \Lambda & \mathcal{X}(\sigma) \\ 0 & 0 & 0 & 0 \end{pmatrix}$ with the characteristic polynomial $\chi(x) := -x[-x^3 + \Lambda x^2 + \sigma - f'(\mathcal{X}(\sigma))]$. The only difference is the sign of $\sigma - f'(\mathcal{X}(\sigma))$. Indeed, the definition of $\mathcal{X}(\sigma)$ leads to $\sigma = \frac{f(\mathcal{X}(\sigma))}{\mathcal{X}(\sigma)}$ and the convexity of f implies $\frac{f(x)}{x} - f'(x) > 0$ as soon as $x < 0$ (this argument has already been used in the proof of Proposition 5.12). Since $\sigma \in (-\eta, 0)$, one has

$$\sigma - f'(\mathcal{X}(\sigma)) > 0.$$

The signs of the eigenvalues are thus reversed with respect to the $(0, 0, 0, \sigma)^t$ -case. $\square$

5.4 The center manifold reduction

We will use a method detailed for example in [52] to compute the projection of our system on the center manifold. We first decompose the extended system (5.10) into a linear part (around $(0, 0, 0, 0)^t$) and a non linear one

$$\frac{d}{d\xi} \begin{pmatrix} u \\ u' \\ u'' \\ -\epsilon \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \Lambda & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ u' \\ u'' \\ -\epsilon \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ -\epsilon u - f(u) \\ 0 \end{pmatrix}.$$

The linear part corresponds to the matrix

$$\mathcal{L} := \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \Lambda & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The rest (*i.e* the non linear part) corresponds to the function

$$\mathcal{R} \begin{pmatrix} u \\ u' \\ u'' \\ -\epsilon \end{pmatrix} := \begin{pmatrix} 0 \\ 0 \\ -\epsilon u - f(u) \\ 0 \end{pmatrix}.$$

5.4.1 Study of the linear part

Before investigating the center manifold, we have to determine the center (and unstable) eigenspaces of the linear part $\mathcal{L}$.

Proposition 5.16. *The matrix $\mathcal{L}$ has two eigenvalues 0 and Λ whose characteristics are as follows.*

- The eigenvalue Λ is a simple eigenvalue.
- The eigenvalue 0 is an algebraically third eigenvalue but a geometrically double one : there is only two linearly independents eigenvectors and we have to consider the generalized eigenspace

$$\mathbb{E}_{(0,0,0,0)}^u := \text{Span} \left\{ \begin{pmatrix} \frac{1}{\Lambda^2} \\ \frac{1}{\Lambda} \\ 1 \\ 0 \end{pmatrix} \right\} \quad \text{and} \quad \mathbb{E}_{(0,0,0,0)}^{c,generalized} := \text{Span} \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}$$

Proof. • The two eigenvectors associated to 0 are $\zeta_1 := (0, 0, 0, 1)^t$ and $\zeta_2 := (1, 0, 0, 0)^t$. In order to complete the generalized center eigenspace we choose a vector ζ_3 such that $\mathcal{L}\zeta_3 = \zeta_2$. We can fix $\zeta_3 := (0, 1, 0, 0)^t$.

- The eigenvector associated to the eigenvalue Λ is $\zeta_4 := (\frac{1}{\Lambda^2}, \frac{1}{\Lambda}, 1, 0)^t$.

□

Remark 5.17. *The center manifold reduction for the linear part is nothing but a decomposition in Jordan blocks of the matrix $\mathcal{L}$.*

The basis $(\zeta_1, \zeta_2, \zeta_3, \zeta_4)^t$ satisfies

$$\mathcal{L}\zeta_1 = 0, \quad \mathcal{L}\zeta_2 = 0, \quad \mathcal{L}\zeta_3 = \zeta_2, \quad \mathcal{L}\zeta_4 = \Lambda\zeta_4.$$

The dual basis is $(\zeta_1^*, \zeta_2^*, \zeta_3^*, \zeta_4^*)^t$ with

$$\zeta_1^* := \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \zeta_2^* := \begin{pmatrix} 1 \\ 0 \\ -\frac{1}{\Lambda^2} \\ 0 \end{pmatrix}, \quad \zeta_3^* := \begin{pmatrix} 0 \\ 1 \\ -\frac{1}{\Lambda} \\ 0 \end{pmatrix}, \quad \zeta_4^* := \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

By definition, they verify

$$\mathcal{L}^T \zeta_1^* = 0, \quad \mathcal{L}^T \zeta_2^* = \zeta_3^*, \quad \mathcal{L}^T \zeta_3^* = 0, \quad \mathcal{L}^T \zeta_4^* = \Lambda \zeta_4^*$$

and $\langle \zeta_i, \zeta_j^* \rangle = \delta_{ij}$, $\forall (i, j) \in \{1, 2, 3, 4\}$, where $\mathcal{L}^T$ is the transpose matrix of $\mathcal{L}$.

Section 5.4. The center manifold reduction

Proposition 5.18. Let $\mathbf{X}$ be a point in the phase plan $\mathbb{R}^3 \times (-\eta, 0)$. If $(u, u', u'', -\epsilon)$ are its coordinates in the canonical basis and $(\alpha, \beta, \gamma, \delta)^t$ its coordinates in the new basis $(\zeta_1, \zeta_2, \zeta_3, \zeta_4)^t$ then

$$\begin{cases} \alpha = -\epsilon, \\ \beta = u - \frac{u''}{\Lambda^2}, \\ \gamma = u' - \frac{u''}{\Lambda}, \\ \delta = u'' \end{cases} \quad \text{and} \quad \begin{cases} u = \beta + \frac{\delta}{\Lambda^2}, \\ u' = \gamma + \frac{\delta}{\Lambda}, \\ u'' = \delta, \\ -\epsilon = \alpha. \end{cases} \quad (5.11)$$

5.4.2 Projection on the center manifold

According to the center manifold theorem (Theorem 5.7), there exists a function Φ_c satisfying

$$\Phi_c(\mathbf{0}) = \mathbf{0} \quad \text{and} \quad D\Phi_c(\mathbf{0}) = \mathbf{0} \quad (5.12)$$

such that the center manifold verifies

$$\mathcal{W}_{(0,0,0,0)}^c := \left\{ \mathbf{X}_c + \Phi_c(\mathbf{X}_c), \quad \mathbf{X}_c \in \mathbb{E}_{(0,0,0,0)}^{c, \text{generalized}} \text{ and } \Phi_c(\mathbf{X}_c) \in \mathbb{E}_{(0,0,0,0)}^u \right\}.$$

Remark 5.19. Thanks to Properties (5.12) of Φ_c , a Taylor expansion gives the following relation

$$\Phi_c(\mathbf{X}_c) = \left\langle \mathbf{X}_c, \int_0^1 D^2\Phi_c(s\mathbf{X}_c)(1-s)ds \mathbf{X}_c \right\rangle.$$

Hence, there exist two functions $\mathcal{B}_c$ and Ψ_c defined on $\mathbb{E}_{(0,0,0,0)}^{c, \text{generalized}}$ such that

$$\Phi_c(\mathbf{X}_c) =: \langle \mathbf{X}_c, \mathcal{B}_c(\mathbf{X}_c) \mathbf{X}_c \rangle \quad (5.13)$$

and, for all $\epsilon \in \mathbb{R}$

$$\Phi_c(\epsilon \mathbf{X}_c) = \langle \epsilon \mathbf{X}_c, \mathcal{B}_c(\epsilon \mathbf{X}_c) \epsilon \mathbf{X}_c \rangle = \epsilon^2 \langle \mathbf{X}_c, \mathcal{B}_c(\epsilon \mathbf{X}_c) \mathbf{X}_c \rangle =: \epsilon^2 \Psi_c(\mathbf{X}_c; \epsilon). \quad (5.14)$$

Since $\mathbb{E}_{(0,0,0,0)}^{c, \text{generalized}} = \text{Span} \{ \zeta_1, \zeta_2, \zeta_3 \}$, this suggests that the only non zero component of Φ_c is the ζ_4 direction.

Remark 5.20. Thus there exists φ_c defined on $\mathbb{R}^3$, such that, for $\mathbf{X}_c = \alpha \zeta_1 + \beta \zeta_2 + \gamma \zeta_3 \in \mathbb{E}_{(0,0,0,0)}^{c, \text{generalized}}$,

$$\Phi_c(\mathbf{X}_c) = \varphi_c(-\alpha, \beta, \gamma) \zeta_4.$$

By abuse of notation, we equally use Φ_c or φ_c . The same is true for $\Psi_c(\mathbf{X}_c; \epsilon)$ or $\Psi_c((-\alpha, \beta, \gamma); \epsilon)$ and $\mathcal{B}_c(\mathbf{X}_c)$ or $\mathcal{B}_c(-\alpha, \beta, \gamma)$.

If $\mathbf{X} \in \mathcal{W}_{(0,0,0,0)}^c$ with $\mathbf{X} = \alpha \zeta_1 + \beta \zeta_2 + \gamma \zeta_3 + \delta \zeta_4$, one has the following decomposition

$$\mathbf{X}_c = \alpha \zeta_1 + \beta \zeta_2 + \gamma \zeta_3$$

and

$$\delta = \Phi_c(-\alpha, \beta, \gamma).$$

The function Φ_c cancels in $(0,0,0,0)^t$ by definition but also in some other points, as detailed in the following property.

Proposition 5.21. Let $\sigma \in (-\eta, 0)$ and $\mathcal{X}(\sigma)$ be the unique real such that $\frac{f(\mathcal{X}(\sigma))}{\mathcal{X}(\sigma)} = \sigma$ (as defined in Proposition 5.13), then the function Φ_c satisfies

$$\begin{cases} \Phi_c(-\sigma, 0, 0) = 0, & \forall \sigma \in (-\eta, 0), \\ \Phi_c(-\sigma, \mathcal{X}(\sigma), 0) = 0, & \forall \sigma \in (-\eta, 0). \end{cases} \quad (5.15a)$$

$$(5.15b)$$

Proof. The equilibrium points of the extended system are $(0, 0, 0, \sigma)^t$ and $(\mathcal{X}(\sigma), 0, 0, \sigma)^t$ for $\sigma \in (-\eta, 0)$ in the canonical basis of $\mathbb{R}^4$ (according to Proposition 5.15). Thanks to the change of coordinates (5.11), these equilibrium points write, in the $(\zeta_1, \zeta_2, \zeta_3, \zeta_4)$ basis, $(\sigma, 0, 0, 0)^t$ and $(\sigma, \mathcal{X}(\sigma), 0, 0)^t$.

These equilibrium points correspond to stationary trajectories around $(0, 0, 0, 0)^t$ and so belong to the central manifold $\mathcal{W}_{(0,0,0,0)}^c$, thanks to the central manifold theorem, Theorem 5.7. Thus, the relation $\delta = \Phi_c(-\alpha, \beta, \gamma)$ is valid for the equilibrium points and it gives (5.15a) and (5.15b). $\square$

Once projected on the central manifold $\mathcal{W}_{(0,0,0,0)}^c$, the augmented system (5.10) becomes

$$\begin{cases} \frac{d}{d\zeta} \epsilon = 0, \\ \frac{d}{d\zeta} \beta = \gamma + \frac{\epsilon\beta}{\Lambda^2} + \frac{\epsilon\Phi_c(\epsilon, \beta, \gamma)}{\Lambda^4} + \frac{f\left(\beta + \frac{\Phi_c(\epsilon, \beta, \gamma)}{\Lambda^2}\right)}{\Lambda^2}, \\ \frac{d}{d\zeta} \gamma = \frac{\epsilon\beta}{\Lambda} + \frac{\epsilon\Phi_c(\epsilon, \beta, \gamma)}{\Lambda^3} + \frac{f\left(\beta + \frac{\Phi_c(\epsilon, \beta, \gamma)}{\Lambda^2}\right)}{\Lambda}. \end{cases} \quad (5.16)$$

The δ coordinate satisfies the EDO

$$\frac{d}{d\zeta} \delta = \Lambda\delta - \epsilon\beta - \frac{\epsilon\delta}{\Lambda^2} - f\left(\beta + \frac{\delta}{\Lambda^2}\right).$$

Remark 5.22. The previous system can be rewritten in the form of a linear part and a non linear one

$$\frac{d}{d\zeta} \begin{pmatrix} \epsilon \\ \beta \\ \gamma \\ \delta \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \Lambda \end{pmatrix} \begin{pmatrix} \epsilon \\ \beta \\ \gamma \\ \delta \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{\epsilon\beta}{\Lambda^2} + \frac{\epsilon\Phi_c(\epsilon, \beta, \gamma)}{\Lambda^4} + \frac{f\left(\beta + \frac{\Phi_c(\epsilon, \beta, \gamma)}{\Lambda^2}\right)}{\Lambda^2} \\ \frac{\epsilon\beta}{\Lambda} + \frac{\epsilon\Phi_c(\epsilon, \beta, \gamma)}{\Lambda^3} + \frac{f\left(\beta + \frac{\Phi_c(\epsilon, \beta, \gamma)}{\Lambda^2}\right)}{\Lambda} \\ -\epsilon\beta - \frac{\epsilon\delta}{\Lambda^2} - f\left(\beta + \frac{\delta}{\Lambda^2}\right) \end{pmatrix}.$$

The linear matrix corresponds to the decomposition in Jordan blocks of $\mathcal{L}$ as already mentioned in Remark 5.17.

Remark 5.23. Since the equilibrium points which we are interested in, namely $(0, 0, 0, c)^t$ and $(u_+, 0, 0, c)^t = (\mathcal{X}(c), 0, 0, c)^t$, belongs to the central manifold, our goal is to exhibit an heteroclinic orbit from $(c, 0, 0)^t$ to $(c, u_+, 0)^t = (c, \mathcal{X}(c), 0)^t$ in the (α, β, γ) -frame.

5.4.3 Foliation in α

The projection on the central manifold enables us to reduce the system from a four-dimensional system (5.10) to a three-dimensional one (5.16). Moreover, we have

$$\frac{d}{d\zeta} \alpha = -\frac{d}{d\zeta} \epsilon = 0.$$

Section 5.4. The center manifold reduction

Hence, we can consider only leaves of a foliation successively : we fix $\alpha = -\bar{\epsilon}(= c)$, with $\bar{\epsilon} \in (0, \eta)$. The autonomous system is reduced to a two-dimensional one

$$\begin{cases} \frac{d}{d\bar{\zeta}}\beta = \gamma + \frac{\bar{\epsilon}\beta}{\Lambda^2} + \frac{\bar{\epsilon}\Phi_c(\bar{\epsilon}, \beta, \gamma)}{\Lambda^4} + \frac{f\left(\beta + \frac{\Phi_c(\bar{\epsilon}, \beta, \gamma)}{\Lambda^2}\right)}{\Lambda^2}, \\ \frac{d}{d\bar{\zeta}}\gamma = \frac{\bar{\epsilon}\beta}{\Lambda} + \frac{\bar{\epsilon}\Phi_c(\bar{\epsilon}, \beta, \gamma)}{\Lambda^3} + \frac{f\left(\beta + \frac{\Phi_c(\bar{\epsilon}, \beta, \gamma)}{\Lambda^2}\right)}{\Lambda}. \end{cases} \quad (5.17)$$

Proposition 5.24. *There exists a neighborhood of 0, called $(-\eta_2, \eta_2)$, with $\eta_2 > 0$, such that, for $\bar{\epsilon} \in (-\eta_2, \eta_2)$, the only two equilibrium points of (5.17) are $(0, 0)^t$ and $(\mathcal{X}(-\bar{\epsilon}), 0)^t$, with $\mathcal{X}$ the function introduced in Proposition 5.13*

Proof. The equilibrium points of (5.17) verify

$$\begin{cases} \gamma = 0, \\ \bar{\epsilon} \left[\beta + \frac{\Phi_c(\bar{\epsilon}, \beta, \gamma)}{\Lambda^2} \right] + f\left(\beta + \frac{\Phi_c(\bar{\epsilon}, \beta, \gamma)}{\Lambda^2}\right) = 0. \end{cases}$$

Since we have supposed $f(0) = 0$ in Hypothesis (5.2), equilibrium points correspond to $(\beta, 0)^t$ such that

$$\beta + \frac{\Phi_c(\bar{\epsilon}, \beta, 0)}{\Lambda^2} = 0 \quad \text{or} \quad -\bar{\epsilon} = \frac{f\left(\beta + \frac{\Phi_c(\bar{\epsilon}, \beta, 0)}{\Lambda^2}\right)}{\beta + \frac{\Phi_c(\bar{\epsilon}, \beta, 0)}{\Lambda^2}},$$

that means

$$\beta + \frac{\Phi_c(\bar{\epsilon}, \beta, 0)}{\Lambda^2} = 0 \quad \text{or} \quad \beta + \frac{\Phi_c(\bar{\epsilon}, \beta, 0)}{\Lambda^2} = \mathcal{X}(-\bar{\epsilon}). \quad (5.18)$$

We denote $\mathcal{E}_{\bar{\epsilon}}$ the set of equilibrium points in the (β, γ) -frame. Based on the above consideration, one has

$$\mathcal{E}_{\bar{\epsilon}} := \left\{ (\beta, 0)^t \in \mathbb{R}^2, \text{ such that } \beta + \frac{\Phi_c(\bar{\epsilon}, \beta, 0)}{\Lambda^2} = 0 \quad \text{or} \quad \beta + \frac{\Phi_c(\bar{\epsilon}, \beta, 0)}{\Lambda^2} = \mathcal{X}(-\bar{\epsilon}) \right\}.$$

First and foremost, $(0, 0)^t$ and $(\mathcal{X}(-\bar{\epsilon}), 0)^t$ belongs to $\mathcal{E}_{\bar{\epsilon}}$ thanks to Property 5.21. Let us demonstrate that $\mathcal{E}_{\bar{\epsilon}}$ is reduced to these only two points. We define two functions on $\mathbb{R}^2$

$$\begin{aligned} \mathcal{F}_1(x, y) &:= y + \frac{\Phi_c(x, y, 0)}{\Lambda^2}, \\ \mathcal{F}_2(x, y) &:= y + \frac{\Phi_c(x, y, 0)}{\Lambda^2} - \mathcal{X}(-x). \end{aligned}$$

Both functions are continuous, because of the continuity of $\mathcal{X}$ and Φ_c . They satisfy $\mathcal{F}_1(0, 0) = \mathcal{F}_2(0, 0) = 0$ since $\mathcal{X}(0) = 0$, cf Proposition 5.13. Moreover, the condition $D\Phi_c(\mathbf{0}) = \mathbf{0}$ in (5.12) implies $\partial_y \mathcal{F}_1(0, 0) = \partial_y \mathcal{F}_2(0, 0) = 1$. The implicit function theorem gives the existence of a neighborhood of 0, called $(-\eta_2, \eta_2)$ with $\eta_2 > 0$ and a unique function $\mathfrak{P}_i$, for $i \in \{1, 2\}$ such that : for all $\epsilon \in (-\eta_2, \eta_2)$,

$$\beta + \frac{\Phi_c(\epsilon, \beta, 0)}{\Lambda^2} = 0 \Leftrightarrow \beta = \mathfrak{P}_1(\epsilon) \quad \text{and} \quad \beta + \frac{\Phi_c(\epsilon, \beta, 0)}{\Lambda^2} = \mathcal{X}(-\epsilon) \Leftrightarrow \beta = \mathfrak{P}_2(\epsilon).$$

Then, the equilibrium set verifies

$$\{(0, 0)^t, (\mathcal{X}(-\bar{\epsilon}), 0)^t\} \subseteq \mathcal{E}_{\bar{\epsilon}} \subseteq \{(\mathfrak{P}_1(\bar{\epsilon}), 0)^t, (\mathfrak{P}_2(\bar{\epsilon}), 0)^t\}.$$

For a fixed $\epsilon = \bar{\epsilon}$, the unicity of $\mathfrak{P}_1$ and $\mathfrak{P}_2$ implies $\mathfrak{P}_1(\bar{\epsilon}) = 0$ and $\mathfrak{P}_2(\bar{\epsilon}) = \mathcal{X}(-\bar{\epsilon})$, and Proposition 5.24 is proved. $\square$

Thereafter, we diminish η if necessary in order to have $(-\eta, \eta) \subset (-\eta_1, \eta_1) \cap (-\eta_2, \eta_2)$.

Remark 5.25. In the (β, γ) -frame, our goal is to connect $(0, 0)^t$ to $(\mathcal{X}(c), 0)^t = (u_+, 0)^t$ by an heteroclinic orbit.

5.4.4 Blowup of the system : a singular change of coordinates

The equilibrium points $(0, 0)^t$ and $(\mathcal{X}(-\bar{\epsilon}), 0)^t$ are actually very closed because $\bar{\epsilon}$ corresponds to a small parameter. In the asymptotic case $\bar{\epsilon} \rightarrow 0$, they are even equal. As explained in Section 5.1, to encapsulate the solution in a trapping region and prove that it is global (step 6 of key guidelines, page 162), we choose to study the asymptotic regime and its related homoclinic orbit. We have to blow up the system in order to distinguish as much as possible both equilibrium points and to cancel their dependency into $\bar{\epsilon}$.

We perform the following singular change of coordinates

$$\begin{cases} X(\xi) := \frac{1}{\bar{\epsilon}} \beta \left(\frac{\xi}{\sqrt{\bar{\epsilon}}} \right), \\ Y(\xi) := \frac{1}{\bar{\epsilon}^{\frac{3}{2}}} \gamma \left(\frac{\xi}{\sqrt{\bar{\epsilon}}} \right), \end{cases} \quad (5.19)$$

which leads to the two-dimensional system

$$\begin{cases} \frac{d}{d\xi} X = Y + \frac{\sqrt{\bar{\epsilon}}}{\Lambda^2} \left[X + \frac{\bar{\epsilon} \Psi_c((1, X, \sqrt{\bar{\epsilon}} Y); \bar{\epsilon})}{\Lambda^2} \right. \\ \quad \left. + \left(X + \frac{\bar{\epsilon} \Psi_c((1, X, \sqrt{\bar{\epsilon}} Y); \bar{\epsilon})}{\Lambda^2} \right)^2 h \left(\bar{\epsilon} X + \frac{\bar{\epsilon}^2 \Psi_c((1, X, \sqrt{\bar{\epsilon}} Y); \bar{\epsilon})}{\Lambda^2} \right) \right], \\ \frac{d}{d\xi} Y = \frac{X}{\Lambda} + \frac{\bar{\epsilon} \Psi_c((1, X, \sqrt{\bar{\epsilon}} Y); \bar{\epsilon})}{\Lambda^3} + \frac{1}{\Lambda} \left(X + \frac{\bar{\epsilon} \Psi_c((1, X, \sqrt{\bar{\epsilon}} Y); \bar{\epsilon})}{\Lambda^2} \right)^2 h \left(\bar{\epsilon} X + \frac{\bar{\epsilon}^2 \Psi_c((1, X, \sqrt{\bar{\epsilon}} Y); \bar{\epsilon})}{\Lambda^2} \right). \end{cases} \quad (5.20)$$

We recall that Ψ_c is defined in Remark 5.19 and h in Remark 5.1

Proposition 5.26. For $\bar{\epsilon} \in (0, \eta)$, with $\bar{\epsilon} \neq 0$, there exists only two equilibrium points for (5.20) : $(0, 0)^t$ and $(\frac{\mathcal{X}(-\bar{\epsilon})}{\bar{\epsilon}}, 0)^t$. Moreover, $(0, 0)^t$ is a saddle point (respectively $(\frac{\mathcal{X}(-\bar{\epsilon})}{\bar{\epsilon}}, 0)^t$ a stable spiral) for $\bar{\epsilon} \in (0, \eta)$.

Proof. • The equilibrium points verify

$$\begin{cases} Y = 0, \\ X + \frac{\bar{\epsilon} \Psi_c((1, X, 0); \bar{\epsilon})}{\Lambda^2} + \left(X + \frac{\bar{\epsilon} \Psi_c((1, X, 0); \bar{\epsilon})}{\Lambda^2} \right)^2 h \left(\bar{\epsilon} X + \frac{\bar{\epsilon}^2 \Psi_c((1, X, 0); \bar{\epsilon})}{\Lambda^2} \right) = 0. \end{cases}$$

The definition of h in Remark 5.1 leads to

$$\begin{cases} Y = 0, \\ X + \frac{\bar{\epsilon} \Psi_c((1, X, 0); \bar{\epsilon})}{\Lambda^2} = 0 \quad \text{or} \quad \frac{f(\bar{\epsilon} X + \frac{\bar{\epsilon}^2 \Psi_c((1, X, 0); \bar{\epsilon})}{\Lambda^2})}{\bar{\epsilon} X + \frac{\bar{\epsilon}^2 \Psi_c((1, X, 0); \bar{\epsilon})}{\Lambda^2}} = -\bar{\epsilon}. \end{cases}$$

that means, by definition of $\mathcal{X}$ in Proposition 5.13

$$\begin{cases} Y = 0, \\ X + \frac{\bar{\epsilon} \Psi_c((1, X, 0); \bar{\epsilon})}{\Lambda^2} = 0 \quad \text{or} \quad X + \frac{\bar{\epsilon} \Psi_c((1, X, 0); \bar{\epsilon})}{\Lambda^2} = \frac{\mathcal{X}(-\bar{\epsilon})}{\bar{\epsilon}}. \end{cases}$$

Section 5.4. The center manifold reduction

By defining the two functions

$$\begin{aligned}\mathcal{F}_1(x, y) &:= y + \frac{x\Psi_c((1, y, 0); x)}{\Lambda^2}, \\ \mathcal{F}_2(x, y) &:= y + \frac{x\Psi_c((1, y, 0); x)}{\Lambda^2} + \frac{\mathcal{X}(-x)}{-x}.\end{aligned}$$

One has $\mathcal{F}_1(0, 0) = \mathcal{F}_2(0, -\frac{1}{h(0)}) = 0$ because $\lim_{x \rightarrow 0} \frac{\mathcal{X}(-x)}{-x} = \frac{1}{h(0)}$ thanks to Proposition 5.13. Moreover, one has too $\partial_y \mathcal{F}_1(0, 0) = \partial_y \mathcal{F}_2(0, -\frac{1}{h(0)}) = 1$. We conclude the proof thanks to the implicit function theorem, exactly as the same guidelines used for Proposition 5.24.

- To demonstrate that $(0, 0)^t$ is always a saddle point, we compute the jacobian matrix at $(0, 0)^t$: $\begin{pmatrix} \frac{\sqrt{\bar{\epsilon}}}{\Lambda^2} + \mathcal{O}(\bar{\epsilon}^{\frac{3}{2}}) & 1 + \mathcal{O}(\bar{\epsilon}^2) \\ \frac{1}{\Lambda} + \mathcal{O}(\bar{\epsilon}) & \mathcal{O}(\bar{\epsilon}^{\frac{3}{2}}) \end{pmatrix}$. The eigenvalues are $\frac{\sqrt{\bar{\epsilon}} \pm \sqrt{4\Lambda^3 + \mathcal{O}(\bar{\epsilon})}}{2\Lambda^2}$ and have an opposite sign. The origin is thus a saddle point for the perturbed system.

- Since $\Psi_c((1, \frac{\mathcal{X}(-\bar{\epsilon})}{\bar{\epsilon}}, 0); \bar{\epsilon}) = 0$ and $h(\frac{\mathcal{X}(-\bar{\epsilon})}{\bar{\epsilon}}) = \frac{-\bar{\epsilon}}{\mathcal{X}(-\bar{\epsilon})}$ (thanks to Remark 5.19, Propositions 5.21 and 5.13), the jacobian matrix at $(\frac{\mathcal{X}(-\bar{\epsilon})}{\bar{\epsilon}}, 0)^t$ is $\begin{pmatrix} -\frac{\sqrt{\bar{\epsilon}}}{\Lambda^2} + \mathcal{O}(\bar{\epsilon}^{\frac{3}{2}}) & 1 + \mathcal{O}(\bar{\epsilon}^2) \\ -\frac{1}{\Lambda} + \mathcal{O}(\bar{\epsilon}) & \mathcal{O}(\bar{\epsilon}^{\frac{3}{2}}) \end{pmatrix}$. The eigenvalues are $\frac{-\sqrt{\bar{\epsilon}} \pm i\sqrt{4\Lambda^3 + \mathcal{O}(\bar{\epsilon})}}{2\Lambda^2}$, thus implies that this equilibrium point is a stable spiral. $\square$

Remark 5.27. The square root $\sqrt{4\Lambda^3 + \mathcal{O}(\bar{\epsilon})}$ in the latest line of the previous proof is always well-defined since, Λ is non negative, its order of magnitude is one and $\bar{\epsilon} \ll 1$.

Remark 5.28. In the (X, Y) -frame, our goal is to connect $(0, 0)^t$ to $(\frac{u_+}{-c}, 0)^t = (\frac{\mathcal{X}(c)}{-c}, 0)^t$ by a heteroclinic orbit.

Remark 5.29. Numerically, with the same parameters, the same initial condition and the same Dirichlet boundary conditions that those chosen for Figure 5.1, the desired heteroclinic orbit in the phase space is illustrated in Figure 5.2. Since $f(x) = \frac{x^2}{2}$ in this example, $\mathcal{X}(c) = 2c$ and we want to connect $(0, 0)^t$ to $(-2, 0)^t$.

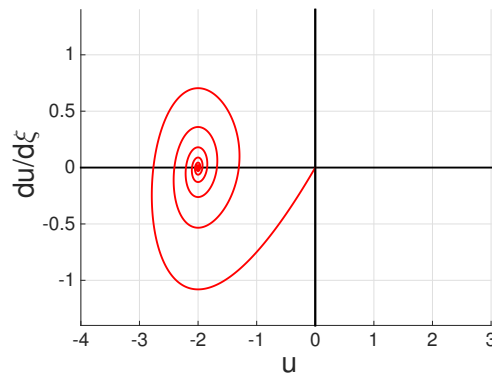


FIGURE 5.2 – Heteroclinic orbit : traveling wave in the phase space in (X, Y) -frame

5.5 Singular limit system

We consider first the limit case where $\bar{\epsilon} = 0$ in (5.20). The singular limit system rewrites

$$\begin{cases} \frac{d}{d\xi} X = Y, \\ \frac{d}{d\xi} Y = \frac{X}{\Lambda} + \frac{X^2 h(0)}{\Lambda}, \end{cases} \quad \xi \in \mathbb{R}. \quad (5.21)$$

This limit system has fruitful properties, for instance, starting with its hamiltonian structure.

Proposition 5.30. *The limit system (5.21) is an hamiltonian system whose Hamiltonian is given by*

$$\begin{aligned} \mathcal{H} &:= \mathbb{R}^2 \rightarrow \mathbb{R} \\ (X, Y) &\mapsto \frac{1}{2\Lambda} \left(X + \frac{1}{h(0)} \right)^2 + \frac{Y^2}{2} - \frac{h(0)}{3\Lambda} \left(X + \frac{1}{h(0)} \right)^3 \end{aligned} \quad (5.22)$$

This function is constant along trajectories of limit system (5.21).

Proof. For $\xi \in \mathbb{R}$, one has

$$\begin{aligned} \frac{d}{d\xi} \mathcal{H}(X(\xi), Y(\xi)) &= \frac{1}{\Lambda} \left(X(\xi) + \frac{1}{h(0)} \right) X'(\xi) + Y(\xi) Y'(\xi) - \frac{h(0)}{\Lambda} \left(X(\xi) + \frac{1}{h(0)} \right)^2 X'(\xi) \\ &= -\frac{1}{\Lambda} X(\xi) X'(\xi) - \frac{h(0)}{\Lambda} X^2(\xi) X'(\xi) + Y(\xi) Y'(\xi) \\ &= 0. \end{aligned}$$

We recognize also the hamiltonian structure of limit system (5.21)

$$\begin{cases} \frac{d}{d\xi} X = \partial_Y \mathcal{H}, \\ \frac{d}{d\xi} Y = -\partial_X \mathcal{H}. \end{cases}$$

□

A particular level curve of $\mathcal{H}$ is the one goes through the origin which corresponds to the level $\frac{1}{6\Lambda h(0)^2}$. Actually, this level curve corresponds to an homoclinic orbit as described in the following property.

Proposition 5.31. *There exists a homoclinic orbit $(X^{hom}, Y^{hom})^t$ at the saddle point $(0, 0)^t$ for the limit system (5.21). This homoclinic orbit crosses the X -axis in $(-\frac{3}{2h(0)}, 0)^t$ and corresponds to the solitary wave with a unit speed for the weighted KdV equation*

$$\partial_t v^{hom} + 2h(0) \partial_x \left(\frac{(v^{hom})^2}{2} \right) + \Lambda \partial_x^3 v^{hom} = 0.$$

The other equilibrium point of (5.21) is the center $(-\frac{1}{h(0)}, 0)^t$.

Proof. • First of all, jacobian matrix at $(0, 0)^t$ writes $\begin{pmatrix} 0 & 1 \\ \frac{1}{\Lambda} & 0 \end{pmatrix}$ with characteristic polynomial $\chi(x) = x^2 - \frac{1}{\Lambda}$. The origin is thus a saddle point with eigenvalues $\frac{1}{\sqrt{\Lambda}} > 0$ and $-\frac{1}{\sqrt{\Lambda}} < 0$.

Section 5.6. Perturbed system : a homoclinic bifurcation

• Moreover, let v^{hom} be the solitary wave at speed one of the weighted KdV equation, then its ansatz $v^{hom}(t, x) = u^{hom}(x - t)$ satisfies

$$-u^{hom} + h(0) (u^{hom})^2 + \Lambda (u^{hom})'' = 0.$$

We fix

$$\begin{cases} X^{hom}(\xi) := -u^{hom}(-\xi), \\ Y^{hom}(\xi) := (u^{hom})'(-\xi), \end{cases}$$

then, $(X^{hom}, Y^{hom})^t$ verifies the limit system (5.21). Then there exists an homoclinic orbit since $u^{hom} \xrightarrow[\xi \rightarrow \pm\infty]{} 0$.

- The homoclinic orbit coincides with the isoline of level $\frac{1}{6\Lambda h(0)^2}$ of $\mathcal{H}$. It passes through the X-axis in $(0, 0)^t$ or $(-\frac{3}{2h(0)}, 0)^t$.
- The other equilibrium point is $(-\frac{1}{h(0)}, 0)^t$. The jacobian matrix at this point is $\begin{pmatrix} 0 & 1 \\ -\frac{1}{\Lambda} & 0 \end{pmatrix}$. Since the eigenvalues are pure imaginary numbers : $\pm i \frac{1}{\sqrt{\Lambda}}$, the equilibrium point $(-\frac{1}{h(0)}, 0)^t$ is a center. $\square$

We have represented in Figure 5.3 those particular orbit of level $\frac{1}{6\Lambda h(0)^2}$ and the level curves of the Hamiltonian $\mathcal{H}$.

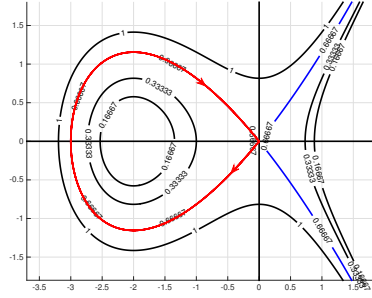


FIGURE 5.3 – Level curves of the Hamiltonian $\mathcal{H}$ in the (KdV⁻-4) case ($h \equiv \frac{1}{2}$) and for $\Lambda = 1$ – The blue curve corresponds to the level $\frac{2}{3}$ and the red one to the homoclinic orbit

5.6 Perturbed system : a homoclinic bifurcation

Thereafter, the main issue is to determined what happens to this homoclinic orbit when we consider the total system (5.20) with parameter $\bar{\epsilon} \neq 0$. There is no reason why the unstable manifold $\mathcal{W}_{(0,0)}^u$ should still intersect the stable manifold $\mathcal{W}_{(0,0)}^s$ and therefore, this homoclinic orbit splits in two. The question then is to identify the mutual positions of $\mathcal{W}_{(0,0)}^u$ and $\mathcal{W}_{(0,0)}^s$. We follow the Melnikov method, introduced by Melnikov in 1963.

First and foremost, from the homoclinic orbit when $\bar{\epsilon} = 0$, cf Figure 5.4(a), there are two possible scenarios : $\mathcal{W}_{(0,0)}^u$ is on the inside of $\mathcal{W}_{(0,0)}^s$, scenario 1 cf Figure 5.4(b), otherwise $\mathcal{W}_{(0,0)}^u$ is on the outside of $\mathcal{W}_{(0,0)}^s$, scenario 2 cf Figure 5.4(c).

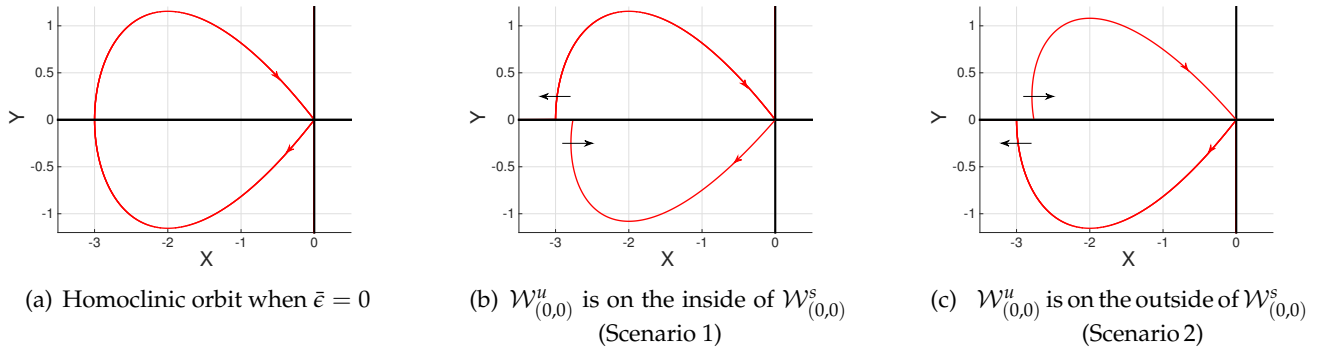


FIGURE 5.4 – Two possible scenarios from the homoclinic orbit

In order to connect $(0,0)^t$ to $(\frac{\mathcal{X}(c)}{-c}, 0)^t$ (i.e. $(-2,0)^t$ in Figure 5.4) in the future, the unstable manifold of $(0,0)^t$ must not pass through the stable manifold of $(0,0)^t$ (otherwise, the homoclinic orbit persists). Thus, the stable manifold plays the role of a fence that blocks solutions of the system (5.20) from crossing it. The mutual positions of $\mathcal{W}_{(0,0)}^u$ and $\mathcal{W}_{(0,0)}^s$ are then determining in the existence of the traveling wave. For instance, the first scenario, for which $\mathcal{W}_{(0,0)}^u$ is on the inside of $\mathcal{W}_{(0,0)}^s$, cf Figure 5.4(b), seems to give the best way to reach $(\frac{\mathcal{X}(c)}{-c}, 0)^t$ (i.e. $(-2,0)^t$ in Figure 5.4) from $(0,0)^t$ without crossing the top fence $\mathcal{W}_{(0,0)}^s$.

Remark 5.32. However, notice that studying the mutual positions $\mathcal{W}_{(0,0)}^u$ and $\mathcal{W}_{(0,0)}^s$ alone is not sufficient to determine the existence of the traveling wave. There may be pathogenic cases as those presented in Figure 5.5. Although the scenario 1 seems to be the ideal case, the solution may not converge to $(\frac{\mathcal{X}(c)}{-c}, 0)^t$ (i.e. $(-2,0)^t$ in Figure 5.4) as illustrated in Subfigure 5.5(a). We can imagine an other situation where the solution converges to $(\frac{\mathcal{X}(c)}{-c}, 0)^t$ (i.e. $(-2,0)^t$ in Figure 5.4) while scenario 2 occurs, see Subfigure 5.5(b).

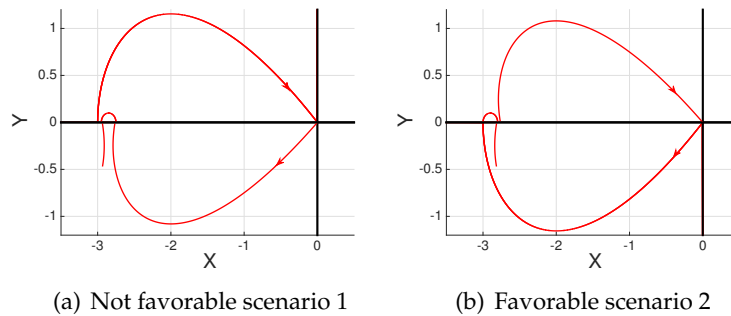


FIGURE 5.5 – Pathogenic cases

To exclude these pathogenic cases, we will use a global Lyapunov function, cf Section 5.7.

5.6.1 The Melnikov integral

In this subsection, we recall the main ideas of the Melnikov method, detailed for example in [71] and [51].

Consider a two-dimensional system with parameter

$$\frac{d}{d\zeta} \mathbf{X} = F(\mathbf{X}; \vartheta), \quad \mathbf{X} = (X_1, X_2)^t \in \mathbb{R}^2, \quad \vartheta \in \mathbb{R}, \quad (5.23)$$

Section 5.6. Perturbed system : a homoclinic bifurcation

where $F := (F_1, F_2)^t$ is C^∞ smooth. Suppose moreover, that, when $\vartheta = 0$, this system has a hyperbolic saddle point X_0^{saddle} (i.e a saddle point with no zero eigenvalue of the jacobian matrix at X_0^{saddle}) and a homoclinic orbit Γ_0 . Let $\xi \in \mathbb{R} \mapsto X_0^{hom}(\xi)$ be a parametrization of this homoclinic orbit. Eventually, suppose that, for ϑ small enough, equilibrium persists and there exists (an other) hyperbolic saddle point X_ϑ^{saddle} .

Fix $\xi_0 = 0$ and consider Σ the cross section going through $X_0^{hom}(0)$. This local section is directed by $\mathbf{n}(X_0^{hom}(0))$, the unit normal to Γ_0 at $X_0^{hom}(0)$. Since, this section Σ is orthogonal to the tangent of the homoclinic orbit at $X_0^{hom}(0)$, one has

$$\mathbf{n}(X_0^{hom}(0)) = \frac{1}{\|F(X_0^{hom}(0); 0)\|} \begin{pmatrix} -F_2(X_0^{hom}(0); 0) \\ F_1(X_0^{hom}(0); 0) \end{pmatrix} \quad (5.24)$$

Remark 5.33. Notice that the orientation of the normal $\mathbf{n}(X_0^{hom}(0))$ is chosen for the basis $(\mathbf{t}(X_0^{hom}(0)), \mathbf{n}(X_0^{hom}(0))^t)$ to be «positively» oriented, when $\mathbf{t}(X_0^{hom}(0))$ stands for the tangent of the homoclinic orbit at $X_0^{hom}(0)$.

Let us now perturb the homoclinic orbit with a small parameter ϑ . Consider X_ϑ^s (respectively X_ϑ^u) the intersection between the stable manifold $\mathcal{W}_{X_\vartheta^{saddle}}^s$ and the section Σ (respectively between the unstable manifold $\mathcal{W}_{X_\vartheta^{saddle}}^u$ and the section Σ).

Remark 5.34. Even with $\vartheta \neq 0$, both stable and unstable manifolds pass through the local cross section Σ , thus both points X_ϑ^s and X_ϑ^u are well-defined. Indeed, Σ is orthogonal to Γ_0 then the intersection between $\mathcal{W}_{X_0^{saddle}}^s$ and Σ is a transversal intersection. The same is true for the intersection between $\mathcal{W}_{X_0^{saddle}}^u$ and Σ . Recall that a transverse intersection is robust : a transversal intersection is stable after a small perturbation, cf [41] for example. Then, when X_0^{saddle} is substituted by X_ϑ^{saddle} , the stable/unstable manifolds $\mathcal{W}_{X_0^{saddle}}^s, \mathcal{W}_{X_0^{saddle}}^u$ are substituted by $\mathcal{W}_{X_\vartheta^{saddle}}^s, \mathcal{W}_{X_\vartheta^{saddle}}^u$ and the transversal intersections are preserved.

Remark 5.35. The definition of X_ϑ^s and X_ϑ^u are illustrated in Figure 5.6. We have chosen a reparametrization of Γ_0 such that $X_0^{hom}(\xi_0 = 0)$ corresponds to the intersection between the homoclinic orbit Γ_0 and the x -axis. The normal vector corresponds here at the horizontal exterior normal vector (in blue in Figure 5.6).

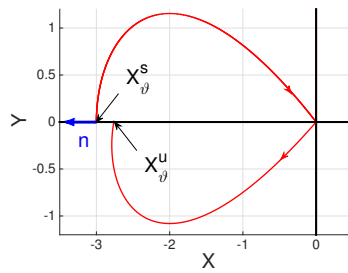


FIGURE 5.6 – Definition of points X_ϑ^u and X_ϑ^s – The normal vector is represented in blue.

We introduce then a function which enables us to measure the displacement (i.e the distance) of the invariant manifolds $\mathcal{W}_{X_\vartheta^{saddle}}^s$ and $\mathcal{W}_{X_\vartheta^{saddle}}^u$, a detailed presentation could be found in [92].

Definition 5.36. Let $(-\eta_3, \eta_3)$, with $\eta_3 > 0$ be a neighborhood of 0. We define $\mathcal{M}$, the split function (also called the separation function) by

$$\begin{aligned} \mathcal{M} &:= (-\eta_3, \eta_3) \rightarrow \mathbb{R} \\ \vartheta &\mapsto \left\langle X_\vartheta^s X_\vartheta^u, \mathbf{n}(X_0^{hom}(0)) \right\rangle, \end{aligned}$$

when $\mathbf{X}_\vartheta^s \mathbf{X}_\vartheta^u$ corresponds to the vector from $\mathbf{X}_\vartheta^s$ to $\mathbf{X}_\vartheta^u$ and $\mathbf{n}(\mathbf{X}_0^{hom}(0))$ the normal vector defined by (5.24).

Remark 5.37. Note that Feng and Hu in [47] define their split function by $\Delta_\vartheta(\xi) := -\langle \mathbf{X}_\vartheta^u \mathbf{X}_\vartheta^s, \mathbf{n}(\mathbf{X}_0^{hom}(\xi)) \rangle$. And thus, for $\xi_0 = 0$, one has $\Delta_\vartheta(0) := \mathcal{M}(\vartheta)$.

It would be very difficult to compute exactly this split function, but the idea of Melnikov was to compute the first terms of a Taylor expansion,

$$\mathcal{M}(\vartheta) = \mathcal{M}(0) + \partial_\vartheta \mathcal{M}(0) \vartheta + \mathcal{O}(\vartheta^2).$$

The existence of the homoclinic orbit at $\vartheta = 0$ implies $\mathcal{M}(0) = 0$, the second term $\partial_\vartheta \mathcal{M}(0)$ is called the Melnikov integral and we can compute it exactly, as stated in the following proposition, cf [92] or [71].

Proposition 5.38. Consider a two-dimensional system

$$\begin{cases} \frac{d}{d\xi} X = F_1(\mathbf{X}; \vartheta), \\ \frac{d}{d\xi} Y = F_2(\mathbf{X}; \vartheta), \end{cases} \quad \text{with } \mathbf{X} = (X, Y) \in \mathbb{R}^2, \vartheta \in \mathbb{R},$$

having at $\vartheta = 0$ a saddle equilibrium point $\mathbf{X}_0^{saddle}$ and a homoclinic orbit Γ_0 . Let $\xi \in \mathbb{R} \mapsto \mathbf{X}_0^{hom}(\xi)$ be a parametrization of this homoclinic orbit. Then, the Melnikov integral is equal to

$$\partial_\vartheta \mathcal{M}(0) = \int_{-\infty}^{+\infty} \left\{ \exp \left[- \int_0^\xi \text{div} F(\mathbf{X}_0^{hom}(\zeta); 0) d\zeta \right] \right\} F(\mathbf{X}_0^{hom}(\xi); 0) \wedge \partial_\vartheta F(\mathbf{X}_0^{hom}(\xi); 0) d\xi,$$

where

$$\text{div} F = (\partial_X F_1 + \partial_Y F_2) \quad \text{and} \quad F \wedge \partial_\vartheta F = (F_1 \partial_\vartheta F_2 - F_2 \partial_\vartheta F_1).$$

We can conclude on the mutual positions of the invariant manifolds as follows.

Corollary 5.39. • If $\partial_\vartheta \mathcal{M}(0) \vartheta > 0$, the unstable manifold is on the outside of the stable manifold (scenario 2 of Figure 5.4(c)).

• If $\partial_\vartheta \mathcal{M}(0) \vartheta < 0$, the unstable manifold is on the inside of the stable manifold (scenario 1 of Figure 5.4(b)).

Remark 5.40. In the conclusions of our Corollary 5.39, we recognize the conclusion of Theorem 9 of [47].

Remark 5.41. There exists a way to define Melnikov integral for any dimensional system. First, note that $\left\{ \exp \left[- \int_0^\xi \text{div} F(\mathbf{X}_0^{hom}(\zeta); 0) d\zeta \right] \right\} F(\mathbf{X}_0^{hom}(\xi); 0) \wedge \partial_\vartheta F(\mathbf{X}_0^{hom}(\xi); 0)$ could be rewritten in terms of a scalar product

$$\left\{ \exp \left[- \int_0^\xi \text{div} F(\mathbf{X}_0^{hom}(\zeta); 0) d\zeta \right] \right\} \| F(\mathbf{X}_0^{hom}(0); 0) \| \langle \mathbf{n}(\mathbf{X}_0^{hom}(\xi)), \partial_\vartheta F(\mathbf{X}_0^{hom}(\xi); 0) \rangle.$$

Secondly, notice that $\left\{ \exp \left[- \int_0^\xi \text{div} F(\mathbf{X}_0^{hom}(\zeta); 0) d\zeta \right] \right\} \| F(\mathbf{X}_0^{hom}(0); 0) \| \mathbf{n}(\mathbf{X}_0^{hom}(\xi))$ is the unique (within a scalar multiple) bounded solution (called $z(\xi)$ thereafter) of the adjoint equation

$$\frac{d}{d\xi} \mathbf{Z} = -\mathbb{A}^t(\xi) \mathbf{Z}, \quad \mathbf{Z} \in \mathbb{R}^2,$$

with $\mathbb{A}(\xi) := D_{\mathbf{X}} F(\mathbf{X}_0^{hom}(\xi); 0)$ the jacobian matrix of F (with respect to $\mathbf{X}$) at $\mathbf{X} = \mathbf{X}_0^{hom}(\xi)$ and $\vartheta = 0$. We can extend the definition of Melnikov integral to any n -dimensional system by defining

$$\partial_\vartheta \mathcal{M}(0) = \int_{-\infty}^{+\infty} \langle z(\xi), \partial_\vartheta F(\mathbf{X}_0^{hom}(\xi); 0) \rangle d\xi.$$

More details can be found in [71].

Section 5.6. Perturbed system : a homoclinic bifurcation

In the particular case of a 2-dimensional Hamiltonian system, the computation of the Melnikov integral is significantly simplified thanks to the following proposition, proved in [71].

Proposition 5.42. *Consider a two-dimensional system endowed with an Hamiltonian structure at $\vartheta = 0$*

$$\frac{d}{d\xi} \mathbf{X} = J(\nabla \mathcal{H})(\mathbf{X}) + \vartheta g(\mathbf{X}), \quad \mathbf{X} = (X, Y)^t \in \mathbb{R}^2, \vartheta \in \mathbb{R}$$

with $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $\nabla \mathcal{H} = (\partial_X \mathcal{H}, \partial_Y \mathcal{H})^t$, having at $\vartheta = 0$ a saddle equilibrium point $\mathbf{X}_0^{\text{saddle}}$ and a homoclinic orbit Γ_0 . As before, let $\xi \in \mathbb{R} \mapsto \mathbf{X}_0^{\text{hom}}(\xi)$ be a parametrization of this homoclinic orbit but in the hamiltonian case, we can also described the homoclinic orbit in terms of level sets of $\mathcal{H} : \Gamma_0 = \{\mathbf{X} \in \mathbb{R}^2, \text{ such that } \mathcal{H}(\mathbf{X}) = \mathcal{H}(\mathbf{X}_0^{\text{saddle}})\}$. Assume the interior of Γ_0 is the domain $\Omega := \{\mathbf{X} \in \mathbb{R}^2, \text{ such that } \mathcal{H}(\mathbf{X}) \leq \mathcal{H}(\mathbf{X}_0^{\text{saddle}})\}$. Then the Melnikov integral along Γ_0 is equal to

$$\partial_{\vartheta} \mathcal{M}(0) = \int_{\Omega} (\text{div } g)(X, Y) dX dY.$$

5.6.2 Study of the homoclinic bifurcation

We substitute $\sqrt{\varepsilon}$ by ϑ in (5.20) in order to reduce system (5.20) to the classical form (5.23) studied in the previous subsection

$$\begin{cases} \frac{d}{d\xi} X = F_1(X, Y; \vartheta), \\ \frac{d}{d\xi} Y = F_2(X, Y; \vartheta), \end{cases} \quad (5.25)$$

with,

$$\begin{cases} F_1(X, Y; \vartheta) = Y + \frac{\vartheta}{\Lambda^2} \left[X + \frac{\vartheta^2 \Psi_c((1, X, \vartheta Y); \vartheta^2)}{\Lambda^2} + \left(X + \frac{\vartheta^2 \Psi_c((1, X, \vartheta Y); \vartheta^2)}{\Lambda^2} \right)^2 h \left(\vartheta^2 X + \frac{\vartheta^4 \Psi_c((1, X, \vartheta Y); \vartheta^2)}{\Lambda^2} \right) \right], \\ F_2(X, Y; \vartheta) = \frac{X}{\Lambda} + \frac{\vartheta^2 \Psi_c((1, X, \vartheta Y); \vartheta^2)}{\Lambda^3} + \frac{\left(X + \frac{\vartheta^2 \Psi_c((1, X, \vartheta Y); \vartheta^2)}{\Lambda^2} \right)^2 h \left(\vartheta^2 X + \frac{\vartheta^4 \Psi_c((1, X, \vartheta Y); \vartheta^2)}{\Lambda^2} \right)}{\Lambda}. \end{cases}$$

If necessary, we diminish η such that $(-\eta, \eta) \subset (-\eta_1, \eta_1) \cap (-\eta_2, \eta_2) \cap (-\eta_3, \eta_3)$.

To apply the Melnikov method, we have to verify that there exists a saddle point $\mathbf{X}_{\vartheta}^{\text{saddle}}$ for $\vartheta \neq 0$. However, the origin $(0, 0)^t$ is always a saddle point, thanks to Proposition 5.26, then

$$\mathbf{X}_{\vartheta}^{\text{saddle}} = (0, 0)^t, \quad \forall \vartheta \in (0, \eta).$$

We will apply Proposition 5.38 in our case. Since the limit system has an Hamiltonian structure, cf Proposition 5.30, one has

$$\partial_X F_1(\mathbf{X}^{\text{hom}}(\xi), 0) + \partial_Y F_2(\mathbf{X}^{\text{hom}}(\xi), 0) = 0.$$

Melnikov integral is nothing but

$$\partial_{\vartheta} \mathcal{M}(0) = \int_{-\infty}^{+\infty} (F_1 \partial_{\vartheta} F_2 - F_2 \partial_{\vartheta} F_1) \left(\mathbf{X}_0^{\text{hom}}(\xi), \mathbf{Y}_0^{\text{hom}}(\xi); 0 \right) d\xi, \quad (5.26)$$

Thus, it yields

$$\begin{cases} \partial_{\vartheta} F_1(X, Y; 0) = \frac{X}{\Lambda^2} + \frac{h(0)X^2}{\Lambda^2}, \\ \partial_{\vartheta} F_2(X, Y; 0) = 0. \end{cases}$$

Menikov integral becomes in our case

$$\partial_{\vartheta} \mathcal{M}(0) = \int_{-\infty}^{+\infty} -\frac{1}{\Lambda^3} \left(X_0^{hom}(\xi) + h(0) \left(X_0^{hom}(\xi) \right)^2 \right)^2 d\xi < 0.$$

Since $\vartheta = \sqrt{\varepsilon} > 0$, Corollary 5.39 implies that the scenario 1 of Figure 5.4 is occurred.

Remark 5.43. As explain in Proposition 5.30, our limit system (5.21) is hamiltonian. However, the perturbed system (5.20) is not exactly as those studied in Proposition 5.42, since in our case, the perturbation g still depends on ϑ . However, at first order in ϑ , one has

$$\begin{cases} \frac{d}{d\xi} X = Y + \frac{\vartheta}{\Lambda^2} [X + h(0)X^2], \\ \frac{d}{d\xi} Y = \frac{X}{\Lambda} + \frac{X^2 h(0)}{\Lambda}, \end{cases}$$

and we can apply Proposition 5.42 with

$$\begin{aligned} g : \quad \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ (X, Y)^t &\mapsto \left(\frac{1}{\Lambda^2} [X + h(0)X^2], 0 \right)^t. \end{aligned}$$

Green's theorem implies

$$\int_{\Omega} (\operatorname{div} g)(X, Y) dXdY = \int_{-\infty}^{+\infty} \left\langle g \left(X_0^{hom}(\xi), Y_0^{hom}(\xi) \right), \mathbf{n} \left(X_0^{hom}(\xi), Y_0^{hom}(\xi) \right) \right\rangle d\xi,$$

where $\mathbf{n} \left(X_0^{hom}(\xi), Y_0^{hom}(\xi) \right)$ is the exterior normal to Ω at point $\left(X_0^{hom}(\xi), Y_0^{hom}(\xi) \right)^t$. We recognize the same relation as before

$$\partial_{\vartheta} \mathcal{M}(0) = - \int_{-\infty}^{+\infty} \frac{1}{\Lambda^3} \left[X_0^{hom}(\xi) + h(0) \left(X_0^{hom}(\xi) \right)^2 \right]^2 d\xi$$

5.7 Global «weak» Lyapunov function : elimination of pathogenic cases

As mentioned in Remark 5.32, scenario 1 alone does not ensure the convergence to the equilibrium state $\left(\frac{\mathcal{X}(-\bar{\varepsilon})}{\bar{\varepsilon}}, 0 \right)^t$, with $\bar{\varepsilon} = -c$. We have to exclude pathogenic cases illustrated in Figure 5.5. Pathogenic case of Subfigure 5.5(b) is automatically excluded since the unstable manifold is in the inside of the stable manifold. It means that we have to justify that the pathogenic case of Subfigure 5.5(a) fails : this is the aim of this section.

5.7.1 A global «weak» Lyapunov function

Let us return to the initial coordinates of the system (5.7) in phase plan $(u, u', u'')^t$.

Proposition 5.44. Let $c \in \mathbb{R}$, define the following function

$$\begin{aligned} \mathcal{L}_c &:= \mathbb{R}^3 \rightarrow \mathbb{R} \\ (x_1, x_2, x_3)^t &\mapsto c \frac{x_1^2}{2} - \mathcal{F}(x_1) + \Lambda \frac{x_2^2}{2} - x_3 x_2 \end{aligned}$$

with $\mathcal{F}$ a primitive of f . Then, $\mathcal{L}_c$ is decreasing along trajectories of System (5.7).

Remark 5.45. By abuse of language, we call $\mathcal{L}_c$ a «weak» Lyapunov function because of its decreasing along trajectories and the fact that $\mathcal{L}_c(0,0,0) = 0$. We do not impose $\mathcal{L}_c(x_1, x_2, x_3) > 0$ for $(x_1, x_2, x_3)^t \neq (0,0,0)^t$.

Proof. We multiply by u' the EDO satisfied by the ansatz, Equation (5.5). Denote a primitive of f by $\mathcal{F}$, it holds

$$\begin{aligned} \left[-c \frac{u^2}{2} + \mathcal{F}(u) \right]' &= \Lambda u' u'' - u' u^{(3)} \\ &= \left[\Lambda \frac{(u')^2}{2} \right]' - (u'' u')' + (u'')^2. \end{aligned}$$

It yields

$$\left[c \frac{u^2}{2} - \mathcal{F}(u) + \Lambda \frac{(u')^2}{2} - u'' u' \right]' = -(u'')^2 < 0.$$

We recognize in the left hand side the derivative $\frac{d}{d\tilde{\xi}} \mathcal{L}_c(u(\tilde{\xi}), u'(\tilde{\xi}), u''(\tilde{\xi}))$. The conclusion of Proposition 5.44 is straightforward. $\square$

The successive changes of coordinates correspond to the following relations

$$\begin{cases} c = -\bar{\epsilon}, \\ u(\tilde{\xi}) = \bar{\epsilon} X(\sqrt{\bar{\epsilon}} \tilde{\xi}) + \frac{\bar{\epsilon}^2 \Psi_c((1, X(\sqrt{\bar{\epsilon}} \tilde{\xi}), \sqrt{\bar{\epsilon}} Y(\sqrt{\bar{\epsilon}} \tilde{\xi})); \bar{\epsilon})}{\Lambda^2}, \\ u'(\tilde{\xi}) = \bar{\epsilon}^{\frac{3}{2}} Y(\sqrt{\bar{\epsilon}} \tilde{\xi}) + \frac{\bar{\epsilon}^2 \Psi_c((1, X(\sqrt{\bar{\epsilon}} \tilde{\xi}), \sqrt{\bar{\epsilon}} Y(\sqrt{\bar{\epsilon}} \tilde{\xi})); \bar{\epsilon})}{\Lambda}, \\ u''(\tilde{\xi}) = \bar{\epsilon}^2 \Psi_c((1, X(\sqrt{\bar{\epsilon}} \tilde{\xi}), \sqrt{\bar{\epsilon}} Y(\sqrt{\bar{\epsilon}} \tilde{\xi})); \bar{\epsilon}). \end{cases} \quad (5.27)$$

The function $\mathcal{L}_c$ can be defined also in the (X, Y) -frame, as follows.

Definition 5.46. We define the function $\mathcal{L}_\epsilon$ such that

$$\mathcal{L}_\epsilon := \mathbb{R}^2 \rightarrow \mathbb{R} \\ (X, Y)^t \mapsto \frac{1}{\epsilon^3} \mathcal{L}_{-\epsilon} \left(\epsilon X + \frac{\epsilon^2 \Psi_c((1, X, \sqrt{\epsilon} Y); \epsilon)}{\Lambda^2}, \epsilon^{\frac{3}{2}} Y + \frac{\epsilon^2 \Psi_c((1, X, \sqrt{\epsilon} Y); \epsilon)}{\Lambda}, \epsilon^2 \Psi_c((1, X, \sqrt{\epsilon} Y); \epsilon) \right).$$

Explicitely, it becomes

$$\begin{aligned} \mathcal{L}_\epsilon(X, Y) &= -\frac{(X + \frac{\epsilon \Psi_c((1, X, \sqrt{\epsilon} Y); \epsilon)}{\Lambda^2})^2}{2} - \frac{\mathcal{F}\left(\epsilon X + \frac{\epsilon^2 \Psi_c((1, X, \sqrt{\epsilon} Y); \epsilon)}{\Lambda^2}\right)}{\epsilon^3} \\ &\quad + \frac{\Lambda}{2} \left(Y + \frac{\sqrt{\epsilon} \Psi_c((1, X, \sqrt{\epsilon} Y); \epsilon)}{\Lambda} \right)^2 - \Psi_c((1, X, \sqrt{\epsilon} Y); \epsilon) \left(\sqrt{\epsilon} Y + \frac{\epsilon \Psi_c((1, X, \sqrt{\epsilon} Y); \epsilon)}{\Lambda} \right). \end{aligned}$$

Let us focus on the behavior of $\mathcal{L}_\epsilon(X, Y)$, essentially on the X -line, for the asymptotic case $\epsilon \rightarrow 0$.

Proposition 5.47. Let $X_0 \in (-\infty, -\frac{1}{h(0)})$ and $\omega > 0$ such that $(X_0 - \omega, X_0 + \omega) \subset (-\infty, -\frac{1}{h(0)})$, there exists $\mathcal{V}_{X_0}(0)$ a neighborhood of 0, which depends on X_0 , such that for all $\epsilon \in \mathcal{V}_{X_0}(0)$, the restriction of $\mathfrak{L}_\epsilon$ at the X -axis :

$$\begin{aligned} \mathbb{R} &\rightarrow \mathbb{R} \\ X &\mapsto \mathfrak{L}_\epsilon(X, 0) \end{aligned}$$

is decreasing on $(X_0 - \omega, X_0 + \omega)$.

Proof. • Let us first study the limit case $\epsilon = 0$. By definition of $\mathcal{F}$, one has

$$\frac{d}{dX} \left(\frac{\mathcal{F} \left(\epsilon X + \frac{\epsilon^2 \Psi_c((1, X, 0); \epsilon)}{\Lambda^2} \right)}{\epsilon^3} \right) = \frac{f \left(\epsilon X + \frac{\epsilon^2 \Psi_c((1, X, 0); \epsilon)}{\Lambda^2} \right)}{\epsilon^2} \left(1 + \frac{\epsilon \partial_X \Psi_c((1, X, 0); \epsilon)}{\Lambda^2} \right).$$

Thanks Remark 5.1, we know the asymptotic behavior of this derivative when $\epsilon \rightarrow 0$

$$\lim_{\epsilon \rightarrow 0} \frac{d}{dX} \left(\frac{\mathcal{F} \left(\epsilon X + \frac{\epsilon^2 \Psi_c((1, X, 0); \epsilon)}{\Lambda^2} \right)}{\epsilon^3} \right) = X^2 h(0).$$

Due to the previous calculation, we state

$$\frac{d}{dX} \mathfrak{L}_0(X, 0) = -X - \lim_{\epsilon \rightarrow 0} \frac{d}{dX} \left(\frac{\mathcal{F} \left(\epsilon X + \frac{\epsilon^2 \Psi_c((1, X, 0); \epsilon)}{\Lambda^2} \right)}{\epsilon^3} \right) = -X - h(0)X^2.$$

Since $h(0) > 0$ (we recall that $h(0) = \frac{f''(0)}{2}$ with f strictly convexe in a neighborhood of 0), $X \mapsto \mathfrak{L}_0(X, 0)$ is decreasing on $(-\infty, -\frac{1}{h(0)}) \cup (0, +\infty)$.

• Then, for the perturbed case, assume $X_0 \in (-\infty, -\frac{1}{h(0)})$. Since $(-\infty, -\frac{1}{h(0)})$ is an interval on which the limite case restriction is decreasing, there exists $\omega > 0$ such that $(X_0 - \omega, X_0 + \omega) \subset (-\infty, -\frac{1}{h(0)})$ and a neighborhood $\mathcal{V}_{X_0}(0)$ of ϵ such that the restriction $X \mapsto \mathfrak{L}_\epsilon(X, 0)$ is decreasing on $(X_0 - \omega, X_0 + \omega)$. □

5.7.2 The trapping region

This subsection is devoted to prove that the orbit is trapped in a compact Ω defined in Corollary 5.49. First of all, we exclude the pathologic cases.

Let us assume that there exists an orbit which realizes the pathogenic case on Subfigure 5.5(a). More precisely, we denote $(X(\xi), Y(\xi))^t$ the solution of System (5.20) from $(X(0), Y(0))^t \in \mathcal{W}_{(0,0)}^u$. The unstable manifold theorem, Theorem 5.6, implies automatically that $(X(\xi), Y(\xi))^t \xrightarrow[\xi \rightarrow -\infty]{} (0, 0)^t$, hence, we only have to check the behavior of the solution in the future.

In the limit case $\bar{\epsilon} = 0$, the transversal intersection between the homoclinic orbit $(X_0^{hom}(\xi), Y_0^{hom}(\xi))^t$ and the X -axis, say at $\xi = \xi_0$, implies the existence of $\xi_1 > 0$ such that $(X(\xi_1), Y(\xi_1))^t$ (solution of (5.20)) crosses the X -axis. Without loss of generality, we can assume that $X(\xi_1)$ belongs to a neighborhood of $X_0^{hom}(\xi_0)$, since $\bar{\epsilon} \ll 1$.

Remark 5.48. This transversal intersection argument has already been used in Remark 5.34, to prove the intersection of the perturbed unstable manifold and the X -axis, i.e the existence of X_θ^u , idem for X_θ^s .

Section 5.7. Global «weak» Lyapunov function : elimination of pathogenic cases

Assume, moreover, that the solution again crosses the X -axis in the future in $\xi_2 > \xi_1$, with $X(\xi_2) < X(\xi_1)$ (as illustrated in Subfigure 5.5(a)). As before, the smallness of $\bar{\epsilon}$ enables us to suppose that $X(\xi_2)$ belongs to a neighborhood of $X_0^{hom}(\xi_0)$.

On the one hand, one has

$$\mathcal{L}_{\bar{\epsilon}}(X(\xi_2), 0) > \mathcal{L}_{\bar{\epsilon}}(X(\xi_1), 0). \quad (5.28)$$

Indeed, Proposition 5.31 implies that $(X_0^{hom}(\xi_0), Y_0^{hom}(\xi_0))^t = (-\frac{3}{2h(0)}, 0)^t$. We fix

$$X_0 := -\frac{3}{2h(0)} \in (-\infty, -\frac{1}{h(0)}] \text{ and } \omega \text{ such that } X(\xi_1), X(\xi_2) \in [-\frac{3}{2h(0)} - \omega, -\frac{3}{2h(0)} + \omega]^2$$

in Proposition 5.47. If necessary, we diminish $\bar{\epsilon}$ in order to $-\frac{3}{2h(0)} + \omega < -\frac{1}{h(0)}$. Thus, there exists $\mathcal{V}_{-\frac{3}{2h(0)}}(0)$ a neighborhood of 0 such that, for $\bar{\epsilon} \in \mathcal{V}_{-\frac{3}{2h(0)}}(0)$, the restriction $X \mapsto \mathcal{L}_{\bar{\epsilon}}(X, 0)$ is decreasing on $[-\frac{3}{2h(0)} - \omega, -\frac{3}{2h(0)} + \omega]$.

On the other hand, one has

$$\mathcal{L}_{\bar{\epsilon}}(X(\xi_2), 0) < \mathcal{L}_{\bar{\epsilon}}(X(\xi_1), 0). \quad (5.29)$$

Indeed, the non increasing of $\mathcal{L}_{-\bar{\epsilon}}$ along trajectories, cf Proposition 5.44, implies

$$\mathcal{L}_{-\bar{\epsilon}}(u(\frac{\xi_2}{\sqrt{\bar{\epsilon}}}), u'(\frac{\xi_2}{\sqrt{\bar{\epsilon}}}), u''(\frac{\xi_2}{\sqrt{\bar{\epsilon}}})) < \mathcal{L}_{-\bar{\epsilon}}(u(\frac{\xi_1}{\sqrt{\bar{\epsilon}}}), u'(\frac{\xi_1}{\sqrt{\bar{\epsilon}}}), u''(\frac{\xi_1}{\sqrt{\bar{\epsilon}}})).$$

Relation (5.29) comes from the definition of $\mathcal{L}_{\bar{\epsilon}} : \mathcal{L}_{\bar{\epsilon}}(X(\xi), Y(\xi)) = \frac{1}{\bar{\epsilon}^3} \mathcal{L}_{-\bar{\epsilon}}(u(\frac{\xi}{\sqrt{\bar{\epsilon}}}), u'(\frac{\xi}{\sqrt{\bar{\epsilon}}}), u''(\frac{\xi}{\sqrt{\bar{\epsilon}}}))$.

We obtain a contradiction between Relations (5.28) and (5.29). The pathogenic case of Subfigure 5.5(a) is hence impossible.

Hereafter, we fix the section Σ in the X -axis direction in Remark 5.34 and then, $X_{\sqrt{\bar{\epsilon}}}^s = (X_{\sqrt{\bar{\epsilon}}}^s, 0)$ corresponds to the intersection between the stable manifold $\mathcal{W}_{(0,0)}^s$ and the X -axis, with $\vartheta = \sqrt{\bar{\epsilon}}$.

Corollary 5.49. *The solution $(X, Y)^t$ is confined in a compact Ω whose three boundaries are*

$$(X, Y), \quad \xi \in (-\infty, \xi_1], \\ [X_{\sqrt{\bar{\epsilon}}}^s, X(\xi_1)]$$

and

$$\mathcal{W}_{(0,0)}^s \cap \mathbb{R}^- \times \mathbb{R}^+.$$

Proof. Previously, we have demonstrated that the solution can not cross the segment $[X_{\sqrt{\bar{\epsilon}}}^s, X(\xi_1)]$ (with (5.28) and (5.29)).

As soon as $u'' \neq 0$, the «weak» Lyapunov function is decreasing. However, there exists (at least) a time ξ_3 such that $u''(\xi_3) \neq 0$ on the trajectory, there excludes the case when the solution crosses $\mathcal{W}_{(0,0)}^s \cap \mathbb{R}^+ \times \mathbb{R}^+$, (since in that case, the stable manifold theorem implies the convergence of the solution toward $(0, 0)^t$ in the future). There also excludes the periodic cases (when the solution crosses itself again). Then the solution is strictly in the inside of Ω . $\square$

Remark 5.50. *The «weak» Lyapunov function $\mathcal{L}_{\bar{\epsilon}}$ may also be used to study the global bifurcation of the homoclinic orbit instead of the help of Melnikov integral. Indeed, to eliminate the unfavorable scenario (scenario 2) in Figure 5.4(c) and to prove that only the favorable scenario (scenario 1) in Figure 5.4(b) occurs, we could proceed as follows. Let $(X_{\sqrt{\bar{\epsilon}}}^s, 0)$ and $(X_{\sqrt{\bar{\epsilon}}}^u, 0)$ be the two points defined in Remarks 5.34 and 5.35, with $\vartheta = \sqrt{\bar{\epsilon}}$. Since $\mathcal{L}_{\bar{\epsilon}}$ is decreasing along trajectories (cf Proposition 5.44 and Definition 5.46), one has*

$$\mathcal{L}_{\bar{\epsilon}}(X_{\sqrt{\bar{\epsilon}}}^s, 0) > \mathcal{L}_{\bar{\epsilon}}(\mathbf{0}) \quad \text{and} \quad \mathcal{L}_{\bar{\epsilon}}(\mathbf{0}) > \mathcal{L}_{\bar{\epsilon}}(X_{\sqrt{\bar{\epsilon}}}^u, 0).$$

Therefore, it holds $\mathfrak{L}_{\bar{\epsilon}}(X_{\sqrt{\bar{\epsilon}}}^s, 0) > \mathfrak{L}_{\bar{\epsilon}}(X_{\sqrt{\bar{\epsilon}}}^u, 0)$.

However, for $\bar{\epsilon}$ small enough and ω such that $X_{\sqrt{\bar{\epsilon}}}^s, X_{\sqrt{\bar{\epsilon}}}^u \subset (-\frac{3}{2h(0)} - \omega, -\frac{3}{2h(0)} + \omega)^2$, $\mathfrak{L}_{\bar{\epsilon}}$ is decreasing along $(\min(X_{\sqrt{\bar{\epsilon}}}^s, X_{\sqrt{\bar{\epsilon}}}^u), \max(X_{\sqrt{\bar{\epsilon}}}^s, X_{\sqrt{\bar{\epsilon}}}^u))$ (cf Proposition 5.47). Only the case, $X_{\sqrt{\bar{\epsilon}}}^s < X_{\sqrt{\bar{\epsilon}}}^u$ is thus possible, which bans the scenario 2 of Figure 5.4(c).

We are now able to demonstrate the main theorem, Theorem 5.3

Proof of Theorem 5.3. We recall the classical arguments to prove the existence of a traveling wave for (5.1). These guidelines are significantly the same than those used in [73] to study the traveling wave solution of one-dimensional non conservative hyperbolic system :

- proof of the existence of a local solution around $(0, 0)^t$ (by Proposition 5.51)
- extension to $\mathbb{R}$ of the existence region of the solution (by Corollary 5.53)
- convergence (with oscillations) towards $(\frac{\mathcal{X}(-\bar{\epsilon})}{\bar{\epsilon}}, 0)^t$ (by Corollary 5.55)

Let us diminish η to have $(-\eta, \eta) \subset (-\eta_1, \eta_1) \cap (-\eta_2, \eta_2) \cap (-\eta_3, \eta_3) \cap \mathcal{V}_{-\frac{3}{2h(0)}}(0)$. Let $\bar{\epsilon} \in (0, \eta)$.

Proposition 5.51. *There exists a trajectory $(X, Y)^t$, defined locally on an open set around $(0, 0)^t$ and which converges to $(0, 0)^t$ when $\xi \rightarrow -\infty$.*

Proof. We consider the Cauchy problem composed of System (5.20) and initial condition $(X(0), Y(0))^t \in \mathcal{W}_{(0,0)}^u$. By Cauchy-Lipschitz theorem, there exists a unique maximal solution $(X, Y)^t$ defined locally around $(0, 0)^t$. Moreover, the unstable manifold theorem, Theorem 5.6, implies automatically that $(X(\xi), Y(\xi))^t \xrightarrow[\xi \rightarrow -\infty]{} (0, 0)^t$. $\square$

Thereafter, $(X, Y)^t$ will denote this solution.

Our aim is then to prove that $(X, Y)^t$ converges to $(\frac{\mathcal{X}(-\bar{\epsilon})}{\bar{\epsilon}}, 0)^t$ in $+\infty$. A first step consists of demonstrating that the previous solution can be extended on $\mathbb{R}$.

We recall the well-known «Compact break» theorem, see for example [100],

Theorem 5.52 («Compact break» theorem). *Let $(a, b) \subset \mathbb{R}$ be an open set and $f : (a, b) \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ continuous and locally Lipschitz in its second variable. Let y be a maximal solution defined on (ξ_-, ξ_+) of the Cauchy problem*

$$\begin{cases} x' = f(\xi, x), \\ x|_{t=0} = x_0. \end{cases}$$

If $\xi_+ < b$ then y leaves any compact set as ξ approaches to ξ_+ i.e for any compact set Ω , $y(\xi) \notin \Omega$ for ξ sufficiently close to ξ_+ . The same result occurs for ξ_- .

Corollary 5.53. *The maximal solution $(X, Y)^t$ is global.*

Proof. The maximal solution is trapped in the compact Ω , thanks to Corollary 5.49, then it is global by the «Compact break» theorem. $\square$

The maximal solution $(X, Y)^t$ is then defined on $\mathbb{R}$ and confined in a compact subset : the hypotheses of the following theorem are all true, see [100].

Section 5.8. Afterword : Traveling waves for KdV-Burgers equation

Theorem 5.54 (Poincaré-Bendixson theorem). *Given an autonomous system in $\mathbb{R}^2$*

$$x' = f(x).$$

Let $x(\xi)$ be a solution which stays in a compact set for all ξ . Then either $x(\xi)$ converges to an equilibrium point when $\xi \rightarrow +\infty$, or it converges to a single periodic cycle.

Corollary 5.55. *The maximal solution $(X, Y)^t$ converges in the future to $(\frac{\mathcal{X}(-\bar{\epsilon})}{\bar{\epsilon}}, 0)^t$.*

Proof. By the Poincaré-Bendixson theorem, we have to eliminate the periodic cycle case and the homoclinic case (when $(X, Y)^t$ goes back to $(0, 0)^t$). However, the (strict) decreasing of $\mathcal{L}_c$ in some intervals along trajectories excludes the convergence to a periodic orbit and the convergence to the fixed point $(0, 0)^t$. $\square$

Then we have achieved to connect the unstable equilibrium $(0, 0)^t$ to the stable equilibrium $(\frac{\mathcal{X}(-\bar{\epsilon})}{\bar{\epsilon}}, 0)^t$. Moreover, the traveling wave oscillates around $(\frac{\mathcal{X}(-\bar{\epsilon})}{\bar{\epsilon}}, 0)^t$, since the equilibrium point $(\frac{\mathcal{X}(-\bar{\epsilon})}{\bar{\epsilon}}, 0)^t$ is a stable spiral (see Proposition 5.26).

We have achieved to connect $(0, 0)^t$ to $(\frac{\mathcal{X}(-\bar{\epsilon})}{\bar{\epsilon}}, 0)^t$ in the (X, Y) -frame. The change of coordinates (5.27) implies the connection between $(0, 0, 0, c)^t$ and $(u_+, 0, 0, c)^t$ in the (u, u', u'', c) -frame (thanks to Proposition 5.21 and relation $\mathcal{X}(c) = u_+$). $\square$

5.8 Afterword : Traveling waves for KdV-Burgers equation

A little aside concerning the Korteweg-de Vries-Burgers equation,

$$\partial_t v + \partial_x \left(\frac{v^2}{2} \right) = \lambda \partial_x^2 v + \mu \partial_x^3 v, \quad (\text{KdV}^- - \text{B})$$

with $\lambda > 0$ and $\mu > 0$, is studied in this section.

As before, we focus on the traveling wave solutions and prove their existence and behavior.

Remark 5.56. *Bona and Schonbek in [21] have already answered this question. But their proof is more "geometrical" and constructive than ours. And, it would be interesting to mention our proof which makes the link between the KdV⁻-Burgers equation and the well-known Fisher-KPP equation.*

Desired traveling wave solutions have the form

$$v(t, x) := u \left(\frac{x - ct}{\sqrt{\mu}} \right),$$

where u is the ansatz and $c > 0$ the wave speed. The traveling wave also satisfies the asymptotic conditions

$$\lim_{\xi \rightarrow \pm\infty} u(\xi) = u_{\pm}, \quad \text{and} \quad \lim_{\xi \rightarrow \pm\infty} u^{(i)}(\xi) = 0, \quad \text{for } i \in \{1, 2\}.$$

Remark 5.57. *We are interested in traveling wave relevant to an heteroclinic orbit, therefore, we suppose $u_+ \neq u_-$.*

The following result is the main theorem of the section.

Theorem 5.58. Traveling wave for KdV⁻ - Burgers equation

- If $u_+ > u_-$, there does not exist a traveling wave of speed $c > 0$ connecting u_- to u_+ .

- If $u_+ < u_-$, there exists a traveling wave of speed $c > 0$. Moreover, this traveling wave
 - decreases monotonically to u_+ if $\lambda^2 \geq 4\mu(u_- - c)$
 - and has an oscillation behavior around u_+ if $\lambda^2 < 4\mu(u_- - c)$.

Remark 5.59. Notice that Theorem 5.58 is the one proved by Bona and Schonbek in [21]. However, their oscillations occur around the unstable state u_- whereas for our case, they occur in the future, around u_+ . This difference results in the initial equation. Bona and Schonbek study

$$\partial_t v_1 + \partial_x \left(\frac{v_1^2}{2} \right) + \mu \partial_x^3 v_1 = \lambda \partial_x^2 v_1,$$

whereas we focus on the equation

$$\partial_t v_2 + \partial_x \left(\frac{v_2^2}{2} \right) = \lambda \partial_x^2 v_2 + \mu \partial_x^3 v_2.$$

In both cases, parameters λ and μ are non negative. Since $v_2(t, x) = -v_1(t, -x)$, the behavior of the solutions are reversed.

To illustrate numerically Theorem 5.58, we choose $\lambda = 0.1$ and two dispersive parameters $\mu : \mu = 10.5 \frac{\lambda^2}{4(u_- - c)}$ or $\mu = 0.5 \frac{\lambda^2}{4(u_- - c)}$. The first verifies $\lambda^2 < 4\mu(u_- - c)$ and the seconde $\lambda^2 \geq 4\mu(u_- - c)$. We fix $u_- = 1$ and $u_+ = 0$ and initialize the scheme with

$$u_0 = \begin{cases} 1, & \text{if } x < \frac{L}{2}, \\ 0, & \text{if } x \geq \frac{L}{2}. \end{cases}$$

Finally, we fix the space domain to $[0, L]$ with $L = 50$, the boundary conditions to Dirichlet conditions and the time domain to $[0, 40]$.

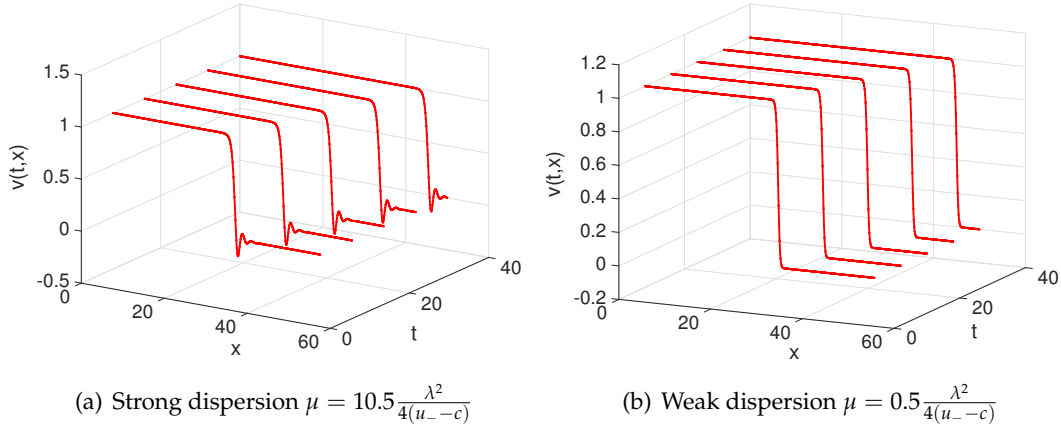


FIGURE 5.7 – Traveling wave solution for the KdV-Burgers equation

5.8.1 From KdV-Burgers equation to Fisher-KPP equation

Note that the ansatz u necessary satisfies the following EDO, where $'$ stands for the derivation with respect to $\xi = x - ct$,

$$-cu'(\xi) + \left(\frac{u^2}{2} \right)'(\xi) = \frac{\lambda}{\sqrt{\mu}} u''(\xi) + u'''(\xi), \quad \xi \in \mathbb{R}.$$

Section 5.8. Afterword : Traveling waves for KdV-Burgers equation

It follows immediately by integrating between $-\infty$ and ξ that

$$-c(u - u_-) + \frac{u^2 - u_-^2}{2} = \frac{\lambda}{\sqrt{\mu}} u' + u''. \quad (5.30)$$

When $\xi \rightarrow +\infty$ and since $u_+ \neq u_-$ (cf Remark 5.57), we obtain a relation between the asymptotic states $u_{\pm}$ and the speed c :

$$c = \frac{u_+ + u_-}{2}.$$

Rearranging some terms in (5.30) gives

$$u'' + \frac{\lambda}{\sqrt{\mu}} u' = \frac{u - u_-}{2} (-2(c - u_-) + (u - u_-)).$$

Let us now perform an affine change of variables combined with a rescaling in ξ by defining

$$y(\xi) = -\frac{1}{2(c - u_-)} u \left(\sqrt{-\frac{1}{c - u_-}} \xi \right) + \frac{2c - u_-}{2(c - u_-)}.$$

Remark 5.60. Since $c = \frac{u_+ + u_-}{2}$ with the hypothesis $u_+ < u_-$, the difference $c - u_- = \frac{u_+ - u_-}{2}$ is negative and the previous change of variables is well defined.

This new variable y verifies the asymptotic behavior

$$\lim_{\xi \rightarrow -\infty} y(\xi) = 1, \quad \lim_{\xi \rightarrow +\infty} y(\xi) = 0 \quad \text{and} \quad \lim_{\xi \rightarrow \pm\infty} y'(\xi) = 0$$

and the following well-known equation

$$y'' + \frac{\lambda}{\sqrt{\mu}} \sqrt{-\frac{1}{c - u_-}} y' = (y - 1)y. \quad (5.31)$$

We recognize the EDO satisfied by the ansatz of a traveling wave of speed

$$\sigma := \frac{\lambda}{\sqrt{\mu}} \sqrt{-\frac{1}{c - u_-}} > 0,$$

solution of Fisher-KPP equation.

Remark 5.61. The Fisher-Kolmogorov-Petrovsky-Piskunov equation, Fisher-KPP equation hereafter, derived in 1937, is a model of species dynamics with diffusion and logistic growth [48, 68]. We consider here the dimensionless equation, namely the diffusion constant, the growth rate of the species and the carrying capacity are all equal to one. Therefore, this equation takes the form

$$\partial_t v - \partial_x^2 v = v(1 - v).$$

A traveling wave moving with speed $\sigma > 0$, say $y(x - \sigma t)$, satisfies the following EDO

$$-\sigma y' - y'' = y(1 - y)$$

and we recognize (5.31) with $\sigma := \frac{\lambda}{\sqrt{\mu}} \sqrt{-\frac{1}{c - u_-}}$.

Proving the existence of a traveling wave for Fisher-KPP is hence enough to prove the existence of traveling wave solution of KdV⁻-Burgers equation.

Proposition 5.62. *There exist traveling waves with speed σ , solutions of Fisher-KPP equation connecting $(1, 0)^t$ to $(0, 0)^t$. These traveling waves decrease monotonically if $\sigma \geq 2$ and have an oscillatory behavior if $0 \leq \sigma < 2$.*

Proof. This proof is initiated in Section 13.2 of [85] and is detailed in Appendix C. □

Proof of Theorem 5.58. Thanks to Proposition 5.62, there exists a monotone traveling wave solution for KdV-Burgers equation if $\sigma = \frac{\lambda}{\sqrt{\mu}} \sqrt{-\frac{1}{c-u_-}} \geq 2$. This traveling wave is oscillating if $\sigma = \frac{\lambda}{\sqrt{\mu}} \sqrt{-\frac{1}{c-u_-}} < 2$. □

CHAPITRE 6

Stability of the traveling wave for the KdV-KS equation

The study of the traveling wave for KdV-KS equation, initiated in Chapter 5, is pursued by the question of the stability of this traveling wave. By using Evans function, we prove numerically that this traveling wave seems to be spectrally stable for the degenerate KdV-KS equation without second order diffusion term and for the degenerate KdV-KS equation without fourth order diffusion term.

Summary

6.1	Introduction (en français)	191
6.2	Linear stability and the Evans function : a revue	192
6.2.1	From the PDE towards the ρ -dependent first-order dynamical system <i>via</i> the spectral problem	192
6.2.2	Spectrum of linearized operator $\mathcal{M}$	193
6.2.2.1	Essential spectrum	193
6.2.2.2	Eigenvalues of $\mathcal{M}$ and Evans function	194
6.3	The ρ -dependent first order system for (KdV ⁻ -4) equation	195
6.3.1	Limit matrices $\mathbb{A}_{\pm}^{KdV-4}(\rho)$	195
6.3.2	The essential spectrum of $\mathcal{M}^{KdV-4}$	196
6.4	Numerical computation of the Evans function	197
6.4.1	In a complex neighborhood of the origin	197
6.4.2	In the real high-frequency limit	200
6.5	Afterword : Stability for the traveling wave of (KdV ⁻ -B) equation	203

6.1 Introduction (en français)

Une fois que l'existence d'ondes progressives est prouvée, une question naturelle est de déterminer la stabilité de ces ondes : c'est le but de ce chapitre. Il s'inscrit dans la suite du chapitre 5 et nous étudions les mêmes équations : l'équation de KdV⁻ avec diffusion d'ordre quatre

$$\partial_t v + \partial_x \left(\frac{v^2}{2} \right) = \mu \partial_x^3 v - \varsigma \partial_x^4 v, \quad (\text{KdV}^- - 4)$$

avec $\mu > 0$ et $\varsigma > 0$, ainsi que l'équation de KdV⁻-Burgers (l'équation de KdV⁻ avec diffusion d'ordre deux)

$$\partial_t v + \partial_x \left(\frac{v^2}{2} \right) = \lambda \partial_x^2 v + \mu \partial_x^3 v, \quad (\text{KdV}^- - \text{B})$$

avec $\lambda > 0$ et $\mu > 0$.

Nous nous focaliserons sur la stabilité spectrale, qui revient à linéariser l'équation autour de l'onde progressive et déterminer le spectre de l'opérateur linéarisé associé.

Définition 6.1. *L'onde progressive est dite spectralement stable, si le spectre de l'opérateur linéaire associé n'intersecte pas le demi-espace $\mathbb{C}_*^+ = \{\rho \in \mathbb{C} : \Re(\rho) \geq 0\} \setminus \{0\}$.*

Il nous faut alors étudier le spectre de cet opérateur et plus particulièrement, son spectre discret (l'ensemble des valeurs propres isolées). Pour cela, nous avons choisi d'étudier numériquement la fonction d'Evans. Il s'agit d'un outil qui permet de localiser les valeurs propres d'un opérateur de dimension infinie et donc qui généralise d'une certaine manière la notion de polynôme caractéristique. Si cette fonction s'annule sur $\mathbb{C}_*^+$, il existera une valeur propre de partie réelle strictement positive : l'onde progressive ne sera pas stable.

En traçant numériquement les fonctions d'Evans, pour les deux équations étudiées (KdV⁻ - 4) et (KdV⁻ - B), Figures 6.2 - 6.3 - 6.4 pour l'équation (KdV⁻ - 4) et Figures 6.5 - 6.6 - 6.7 pour (KdV⁻ - B), elles semblent ne pas s'annuler en dehors de 0 et les ondes progressives être stables pour ces équations. Néanmoins, un calcul théorique de la fonction d'Evans sur le demi-espace $\mathbb{C}_*^+$ serait nécessaire pour conclure catégoriquement quant à la stabilité spectrale des ondes progressives.



Organisation du chapitre. L'analyse faite dans ce chapitre est principalement prospective et l'accent est mis sur la résolution numérique des fonctions d'Evans considérées.

En section 6.2, nous rappelons l'obtention de l'opérateur linéaire associé à l'onde progressive dans le cas général. Nous rappelons aussi les différents éléments qui décomposent un spectre (sous-section 6.2.2.1) ainsi que la définition de la fonction d'Evans (sous-section 6.2.2.2).

Nous appliquons en section 6.3 ces différents rappels à l'équation (KdV⁻ - 4). L'algorithme utilisé pour calculer la fonction d'Evans est détaillé en section 6.4.

Finalement, l'étude de stabilité est effectuée sur l'équation (KdV⁻ - B) en section 6.5.



6.2 Linear stability and the Evans function : a revue

A natural approach to judge the stability of the traveling wave is to study the behavior of the solution near the traveling wave : it is the linearized stability analysis.

6.2.1 From the PDE towards the ρ -dependent first-order dynamical system *via* the spectral problem

Consider the generic PDE

$$\partial_t v + \partial_x (f(v)) = \mathcal{L}_{t, \partial_x, \partial_x^2, \dots, \partial_x^n} v, \quad (6.1)$$

with $\mathcal{L}_{t, \partial_x^i}$ any linear operator (depending on t and ∂_x^i , for $i \in \llbracket 1, n \rrbracket$, $n \in \mathbb{N}$) and f any non linearity. Suppose that (6.1) possesses a traveling wave u and consider a solution v with the following ansatz

$$v(t, x) = \bar{v}(t, x - ct),$$

with

$$\bar{v}(t, x - ct) := u(x - ct) + \bar{w}(t, x - ct),$$

the variable $\bar{w}$ appears like a very small perturbation of the traveling wave. As for Chapter 5, we denote by ξ the change of coordinates $x - ct$ and the ansatz $\bar{v}$ thus satisfies the following PDE

$$\partial_t \bar{v} - c \partial_\xi \bar{v} + \partial_\xi (f(\bar{v})) = \mathcal{L}_{t, \partial_\xi, \partial_\xi^2, \dots, \partial_\xi^n}(\bar{v}),$$

which rewrites at the first order in term of $\bar{w}$ (since u is the traveling wave)

$$\partial_t \bar{w} - c \partial_\xi \bar{w} + \bar{w} \frac{d}{d\xi} (f'(u)) + f'(u) \partial_\xi \bar{w} = \mathcal{L}_{t, \partial_\xi, \partial_\xi^2, \dots, \partial_\xi^n} \bar{w}. \quad (6.2)$$

By introducing a perturbation of the form $\bar{w}(t, x - ct) := e^{\rho t} w(x - ct)$, with $\rho \in \mathbb{C}$, the previous equation becomes the following eigenvalues equation

$$\rho w - c \frac{dw}{d\xi} + w \frac{d}{d\xi} (f'(u)) + f'(u) \frac{dw}{d\xi} = \mathcal{L}_{t, \frac{d}{d\xi}, \frac{d^2}{d\xi^2}, \dots, \frac{d^n}{d\xi^n}} w.$$

The problem may be reformulated in a spectral problem

$$\mathcal{M}w = \rho w, \quad \rho \in \mathbb{C}, \quad w \in H^n(\mathbb{R}), \quad (6.3)$$

with $\mathcal{M}$ the following linearized operator

$$\mathcal{M} := - \frac{d}{d\xi} (f'(u)) I + (c - f'(u)) \frac{d}{d\xi} + \mathcal{L}_{t, \frac{d}{d\xi}, \frac{d^2}{d\xi^2}, \dots, \frac{d^n}{d\xi^n}}.$$

In the previous expression, I stands for the identity operator.

The spectral problem (6.3) may be rewritten in term of a ρ -dependent first-order dynamical system

$$\frac{d}{d\xi} W(\xi) = \mathbb{A}(\rho, \xi) W(\xi), \quad (\rho, \xi) \in \mathbb{C} \times \mathbb{R}, \quad (6.4)$$

where $W := \left(w, \frac{dw}{d\xi}, \frac{d^2 w}{d\xi^2}, \dots, \frac{d^{n-1} w}{d\xi^{n-1}} \right)^t$ and $\mathbb{A}(\rho, \xi)$ is the companion matrix associated to the linearized operator $\mathcal{M}$.

Remark 6.2. There exists a classical method to study the stability of the traveling wave. We heuristically detail the cornerstones of this method. From Equation (6.2), we obtain

$$\partial_t \bar{w} - c \partial_{\xi} \bar{w} + \partial_{\xi} (\bar{w} f'(u)) = \mathcal{L}_{t, \partial_{\xi}, \partial_{\xi}^2, \dots, \partial_{\xi}^n} \bar{w}.$$

By taking $\bar{w} = \partial_{\xi} \bar{y}$ and integrating the equation, we obtain

$$\partial_t \bar{y} - c \partial_{\xi} \bar{y} + f'(u) \partial_{\xi} \bar{y} = \mathcal{L}_{t, \partial_{\xi}, \partial_{\xi}^2, \dots, \partial_{\xi}^n} \bar{y}.$$

We then have to multiply by $\bar{y}$ and integrate it to get

$$\frac{d}{dt} \int_{\mathbb{R}} \frac{\bar{y}^2}{2}(\xi) d\xi - \int_{\mathbb{R}} f''(u) \frac{du}{d\xi}(\xi) \frac{\bar{y}^2}{2}(\xi) d\xi = \int_{\mathbb{R}} (\mathcal{L}_{t, \partial_{\xi}, \partial_{\xi}^2, \dots, \partial_{\xi}^n} \bar{y})(\xi) \bar{y}(\xi) d\xi.$$

In most cases, $\int_{\mathbb{R}} (\mathcal{L}_{t, \partial_{\xi}, \partial_{\xi}^2, \dots, \partial_{\xi}^n} \bar{y})(\xi) \bar{y}(\xi) d\xi$ is signed. Moreover, $f''(u) \geq 0$ because f is supposed to be a convex function. If the traveling wave u is monotone, we may conclude about the finiteness of $\int_{\mathbb{R}} \frac{\bar{y}^2}{2}(\xi) d\xi$. In our case, the oscillations of u prevent us to use that classical method of stability study and we need to study the spectral stability.

6.2.2 Spectrum of linearized operator $\mathcal{M}$

We consider $\mathcal{M}$ as an operator on $L^2(\mathbb{R})$ whose domain is $H^n(\mathbb{R})$.

Remark 6.3. Hereafter, we denote $\sigma(\mathcal{M})$ (resp. $\sigma_{\text{ess}}(\mathcal{M})$ and $\sigma_{\text{disc}}(\mathcal{M})$) the spectrum of $\mathcal{M}$ (resp. the essential spectrum and the discrete spectrum). We recall the definition of spectrum, essential spectrum and discrete spectrum

$$\begin{aligned} \sigma(\mathcal{M}) &:= \{\rho \in \mathbb{C} : \mathcal{M} - \rho I \text{ does not have a bounded inverse}\}, \\ \sigma_{\text{disc}}(\mathcal{M}) &:= \{\rho \in \mathbb{C} : \rho \text{ is an isolated eigenvalue}\}, \\ \sigma_{\text{ess}}(\mathcal{M}) &:= \sigma(\mathcal{M}) \setminus \sigma_{\text{disc}}(\mathcal{M}). \end{aligned}$$

Definition 6.4. The traveling wave u is said to be (spectrally) stable if

$$\sigma(\mathcal{M}) \cap \{\rho \in \mathbb{C}, \Re(\rho) \geq 0\} = \{0\}.$$

According to the previous definition, we have to understand more in details both essential and discrete spectra of $\mathcal{M}$ to study the spectral stability of the traveling wave u .

6.2.2.1 Essential spectrum

We define two limit matrices

$$A_{\pm}(\rho) := \lim_{\xi \rightarrow \pm\infty} A(\rho, \xi), \quad \rho \in \mathbb{C}.$$

In the same way, we may associate two limit operators $\mathcal{M}_{\pm}$. A comprehension of the spectrum of these two limit operators enables to better understand the essential spectrum of $\mathcal{M}$, according to [53] or [88],

Proposition 6.5. The essential spectrum of $\mathcal{M}$ contains the region inside or on the curves $\sigma(\mathcal{M}_+)$ and $\sigma(\mathcal{M}_-)$ and does not intersect the region $\mathbb{C} \setminus \{\sigma(\mathcal{M}_+) \cup \sigma(\mathcal{M}_-)\}$ which includes the right half-plane. We call hereafter this region $\mathbb{C}^+$.

6.2.2.2 Eigenvalues of $\mathcal{M}$ and Evans function

There exist different definitions of Evans function, introduced first by Evans in 1972 [46]. An overview of these possible definitions as well as connections between them can be found in [11]. According to the designation of [11], we use in this section the «mixed» Evans function definition.

We suppose that $\mathbb{A}_\pm(\rho) := \lim_{\xi \rightarrow \pm\infty} \mathbb{A}(\rho, \xi)$ has k eigenvalues with negative real part and $n - k$ eigenvalues with positive real part. We denote $\mathbb{E}_\pm^s(\rho)$ (respectively $\mathbb{E}_\pm^u(\rho)$) the k dimensional stable (respectively $n - k$ dimensional unstable) subspaces of $\mathbb{A}_\pm(\rho)$. We may choose $\Phi_{i,\pm}$ decaying solutions of (6.4) such that

$$\begin{aligned}\mathbb{E}_+^s(\rho) &:= \lim_{\xi \rightarrow +\infty} \text{Span} \{ \Phi_{1,+}(\rho, \xi), \Phi_{2,+}(\rho, \xi), \dots, \Phi_{k,+}(\rho, \xi) \}, \\ \mathbb{E}_-^u(\rho) &:= \lim_{\xi \rightarrow -\infty} \text{Span} \{ \Phi_{1,-}(\rho, \xi), \Phi_{2,-}(\rho, \xi), \dots, \Phi_{n-k,-}(\rho, \xi) \}.\end{aligned}$$

Consider now the adjoint ODE

$$\frac{d}{d\xi} Z(\xi) = -[\mathbb{A}(\rho, \xi)]^* Z(\xi) = -Z^t(\xi) \mathbb{A}(\bar{\rho}, \xi), \quad (\rho, \xi) \in \mathbb{C} \times \mathbb{R}, \quad (6.5)$$

where $\bar{\rho} = \Re(\rho) - i\Im(\rho)$ and $[\mathbb{A}(\rho, \xi)]^* = [\mathbb{A}(\bar{\rho}, \xi)]^t$ is the adjoint matrix.

Likewise, $[\mathbb{A}(\rho, \xi)]^*$ has always k eigenvalues with negative real part and $n - k$ eigenvalues with positive real part. We define in the same way $\mathbb{E}_\pm^{s,*}(\rho)$ (respectively $\mathbb{E}_\pm^{u,*}(\rho)$) for the adjoint matrices $[\mathbb{A}_\pm(\rho)]^*$ and $\Psi_{i,\pm}$ solutions of (6.5) such that

$$\begin{aligned}\mathbb{E}_+^{u,*}(\rho) &:= \lim_{\xi \rightarrow +\infty} \text{Span} \{ \Psi_{1,+}(\rho, \xi), \Psi_{2,+}(\rho, \xi), \dots, \Psi_{n-k,+}(\rho, \xi) \}, \\ \mathbb{E}_-^{s,*}(\rho) &:= \lim_{\xi \rightarrow -\infty} \text{Span} \{ \Psi_{1,-}(\rho, \xi), \Psi_{2,-}(\rho, \xi), \dots, \Psi_{k,-}(\rho, \xi) \}.\end{aligned}$$

Definition 6.6. The Evans function $D_- : \mathbb{C} \rightarrow \mathbb{R}$ is defined by

$$D_-(\rho) := \det \begin{pmatrix} \langle \Psi_{1,-}, \Phi_{1,+} \rangle & \dots & \langle \Psi_{1,-}, \Phi_{k,+} \rangle \\ \vdots & & \vdots \\ \langle \Psi_{k,-}, \Phi_{1,+} \rangle & \dots & \langle \Psi_{k,-}, \Phi_{k,+} \rangle \end{pmatrix} \Bigg|_{\xi=0}.$$

As explained previously, this definition is sometimes called «mixed » Evans function [11]. This function is studied in details in [88], however, we are only interested in the zeros of D_- with positive real part in our stability study. Indeed, the link between eigenvalues and Evans function is detailed in the following proposition.

Proposition 6.7. Assume $\rho \notin \sigma_{\text{ess}}(\mathcal{M})$, then ρ is an eigenvalue of $\mathcal{M}$ if and only if $D_-(\rho) = 0$.

We have to know whether there exists (at least) a zero of D_- of positive real part or not.

Remark 6.8. We could also define the Evans function as $D_+ : \mathbb{C} \rightarrow \mathbb{R}$

$$D_+(\rho) := \det \begin{pmatrix} \langle \Psi_{1,+}, \Phi_{1,-} \rangle & \dots & \langle \Psi_{1,+}, \Phi_{n-k,-} \rangle \\ \vdots & & \vdots \\ \langle \Psi_{n-k,+}, \Phi_{1,-} \rangle & \dots & \langle \Psi_{n-k,+}, \Phi_{n-k,-} \rangle \end{pmatrix} \Bigg|_{\xi=0}.$$

The two previous definitions (D_- and D_+) differ from a function β , see [11] for more details :

$$D_+(\rho) = \beta(\rho) D_-(\rho), \quad \rho \in \mathbb{C}.$$

A fundamental property of β is that this function does not vanish in $\mathbb{C}^+ \cup \mathcal{V}_0$ where $\mathcal{V}_0$ is a neighborhood of 0, where $\mathbb{C}^+$ is defined in Proposition 6.5. Then zeros of D_- coincide with zeros of D_+ in $\{\rho \in \mathbb{C} : \Re(\rho) > 0\}$. This is why, we could use equally D_- or D_+ to study the stability of the traveling wave.

6.3 The ρ -dependent first order system for (KdV⁻ - 4) equation

We apply this method to the (KdV⁻ - 4) equation. For that equation, one chooses a particular ansatz : $u \left(\frac{x-ct}{\zeta^{\frac{1}{3}}} \right)$, in order to obtain the ratio $\Lambda := \frac{\mu}{\zeta^{\frac{2}{3}}}$ in the ODE verified by u , which is more relevant than both factors μ and ζ separated (see Chapter 5). We perform the same choice for the perturbation $\bar{w}$ and then obtain

$$\bar{v} \left(\frac{t}{\zeta^{\frac{1}{3}}}, \frac{x-ct}{\zeta^{\frac{1}{3}}} \right) := u \left(\frac{x-ct}{\zeta^{\frac{1}{3}}} \right) + \bar{w} \left(\frac{t}{\zeta^{\frac{1}{3}}}, \frac{x-ct}{\zeta^{\frac{1}{3}}} \right).$$

The linearized operator is

$$\mathcal{M}^{KdV-4} := -\frac{du}{d\zeta} I + (c-u) \frac{d}{d\zeta} + \frac{\mu}{\zeta^{\frac{2}{3}}} \frac{d^3}{d\zeta^3} - \frac{d^4}{d\zeta^4}.$$

The companion matrix writes

$$\mathbb{A}^{KdV-4}(\rho, \zeta) := \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\rho - \frac{du}{d\zeta}(\zeta) & c - u(\zeta) & 0 & \Lambda \end{pmatrix}, \quad \text{with } \Lambda := \frac{\mu}{\zeta^{\frac{2}{3}}}.$$

6.3.1 Limit matrices $\mathbb{A}_{\pm}^{KdV-4}(\rho)$

This subsection is devoted to the study of the limit matrices $\mathbb{A}_{\pm}^{KdV-4}(\rho)$, when ζ converges towards $\pm\infty$

$$\lim_{\zeta \rightarrow \pm\infty} \mathbb{A}^{KdV-4}(\rho, \zeta) =: \mathbb{A}_{\pm}^{KdV-4}(\rho).$$

More explicitly, one has

$$\mathbb{A}_{\pm}^{KdV-4}(\rho) := \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\rho & c - u_{\pm} & 0 & \Lambda \end{pmatrix}. \quad (6.6)$$

Remark 6.9. In the same way, we have $\mathcal{M}_{\pm}^{KdV-4}$ equals to

$$\mathcal{M}_{\pm}^{KdV-4} := (c - u_{\pm}) \frac{d}{d\zeta} + \Lambda \frac{d^3}{d\zeta^3} - \frac{d^4}{d\zeta^4}.$$

Proposition 6.10. For $\rho > 0$, both limit matrices $\mathbb{A}_{\pm}^{KdV-4}(\rho)$ defined in (6.6) have two eigenvalues with positive real part (named $\kappa_{3,4}^{\pm}(\rho)$) and two eigenvalues with negative real part (named $\kappa_{1,2}^{\pm}(\rho)$).

Proof. The characteristic polynomial of $\mathbb{A}_{\pm}^{KdV-4}(\rho)$ writes

$$\chi(x) := x^4 - \Lambda x^3 + (-c + u_{\pm})x + \rho.$$

As for the proof of Proposition 5.12, we denote x_i for $i \in \{1, 2, 3, 4\}$, the (real or complex) roots related to the coefficients of χ by the following Vieta's formulas

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = \Lambda > 0, & (6.7a) \\ x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4 = 0, & (6.7b) \\ x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + x_2x_3x_4 = c - u_{\pm} \leq 0, & (6.7c) \\ x_1x_2x_3x_4 = \rho > 0. & (6.7d) \end{cases}$$

Section 6.3. The ρ -dependent first order system for (KdV⁻ - 4) equation

Relation (6.7d) implies $x_1, x_2, x_3, x_4 \neq 0$.

- Suppose that x_1 is complex (and that x_2 is its complex conjugate), then $x_1 x_2 = \Re(x_1)^2 + \Im(x_1)^2$ and $x_1 + x_2 = 2\Re(x_1)$. Relation (6.7d) implies $x_3 x_4 > 0$ (thus x_3 and x_4 have the same sign, if they are real) and Relation (6.7b) implies that $x_3 + x_4$ has not the same sign than $\Re(x_1)$. Thus, $\Re(x_1)(= \Re(x_2))$ has not the same sign than $\Re(x_3)(= \Re(x_4))$ if x_3 and x_4 are complex numbers (or the same sign than x_3 and x_4 if they are real numbers).

- Suppose that the four roots are real. Relation (6.7a) implies that (at least) one root (say x_3) is positive. Relation (6.7d) implies that (at least) one root among x_1, x_2, x_4 (say x_4) is positive and that x_1, x_2 have the same sign. Relation (6.7b) rewrites $(x_1 + x_2)(x_3 + x_4) + x_1 x_2 + x_3 x_4 = 0$ and it implies that $x_1 + x_2 < 0$ then $x_1, x_2 < 0 < x_3, x_4$. $\square$

6.3.2 The essential spectrum of $\mathcal{M}^{KdV-4}$

Proposition 6.11. *The essential spectrum of $\mathcal{M}^{KdV-4}$ does not intersect the half-plane $\mathbb{C}_*^+$ with*

$$\mathbb{C}_*^+ := \{\rho \in \mathbb{C} : \Re(\rho) \geq 0\} \setminus \{0\}.$$

Proof. By Fourier transform, we obtain

$$\sigma(\mathcal{M}_{\pm}^{KdV-4}) = \bigcup_{\tau \in \mathbb{R}} \{\rho \in \mathbb{C} : \rho = -\tau^4 + i\tau(c - u_{\pm} - \Lambda\tau^2)\}.$$

Then $\sigma(\mathcal{M}_{\pm}^{KdV-4})$ does not intersect the set $\{\rho \in \mathbb{C} : \Re(\rho) > 0\}$. Moreover, $\rho \in \sigma(\mathcal{M}_{\pm}^{KdV-4}) \cap \{\rho \in \mathbb{C} : \Re(\rho) = 0\}$ implies $\tau = 0$ and so does ρ .

We have proved that $\sigma(\mathcal{M}_{\pm}^{KdV-4}) \cap \mathbb{C}_*^+ = \emptyset$. However, by Proposition 6.5, the essential spectrum of $\mathcal{M}^{KdV-4}$ is contained in the region left or inside the curves $\sigma(\mathcal{M}_+^{KdV-4})$ and $\sigma(\mathcal{M}_-^{KdV-4})$ thus $\mathbb{C}_*^+$ does not intersect $\sigma_{ess}(\mathcal{M}^{KdV-4})$. $\square$

Remark 6.12. In Figure 6.1, we have represented in blue and red lines the spectrum of limit operators $\sigma(\mathcal{M}_{\pm}^{KdV-4})$. The essential spectrum $\sigma_{ess}(\mathcal{M}^{KdV-4})$ belongs to the grey region. Figure 6.1 is obtained by fixing $\Lambda = 1$ and $u_+ = -1, u_- = 1$.

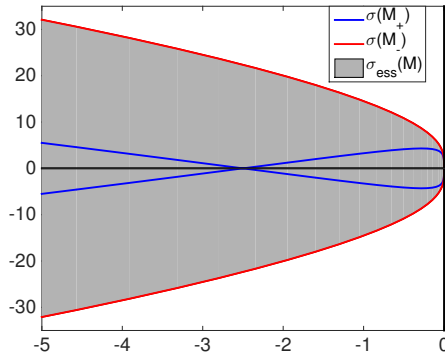


FIGURE 6.1 – Spectrum of the limit operators $\mathcal{M}_{\pm}^{KdV-4}$ and essential spectrum of $\mathcal{M}^{KdV-4}$

Remark 6.13. *The essential spectrum of $\mathcal{M}^{KdV-4}$ lies on the left half plane $\{\Re(\rho) < 0\} \cup \{0\}$, thanks to Proposition 6.11. Thus the stability property of Definition 6.4 is linked to study whether there exists (at least) an isolated eigenvalue of $\mathcal{M}^{KdV-4}$ of positive real part or not.*

6.4 Numerical computation of the Evans function

6.4.1 In a complex neighborhood of the origin

As $\mathbb{A}_{\pm}^{KdV-4}(\rho)$ have two eigenvalues with negative real part, we have $\Phi_{1,2,\pm}$ solutions of (6.4) and $\Psi_{1,2,\pm}$ solutions of (6.5) such that

$$\begin{aligned}\mathbb{E}_+^s(\rho) &:= \lim_{\xi \rightarrow +\infty} \text{Span} \{ \Phi_{1,+}(\rho, \xi), \Phi_{2,+}(\rho, \xi) \}, \\ \mathbb{E}_-^u(\rho) &:= \lim_{\xi \rightarrow -\infty} \text{Span} \{ \Phi_{1,-}(\rho, \xi), \Phi_{2,-}(\rho, \xi) \}, \\ \mathbb{E}_+^{u,*}(\rho) &:= \lim_{\xi \rightarrow +\infty} \text{Span} \{ \Psi_{1,+}(\rho, \xi), \Psi_{2,+}(\rho, \xi) \}, \\ \mathbb{E}_-^{s,*}(\rho) &:= \lim_{\xi \rightarrow -\infty} \text{Span} \{ \Psi_{1,-}(\rho, \xi), \Psi_{2,-}(\rho, \xi) \}.\end{aligned}$$

The Evans function is defined by

$$D_-(\rho) := \det \begin{pmatrix} \langle \Psi_{1,-}, \Phi_{1,+} \rangle & \langle \Psi_{1,-}, \Phi_{2,+} \rangle \\ \langle \Psi_{2,-}, \Phi_{1,+} \rangle & \langle \Psi_{2,-}, \Phi_{2,+} \rangle \end{pmatrix} \Big|_{\xi=0}.$$

By noticing that $\det(\langle \Psi_{i,-}, \Phi_{j,+} \rangle)_{1 \leq i,j \leq 2} \Big|_{\xi=0} = \langle (\Psi_{1,-} \wedge \Psi_{2,-}), (\Phi_{1,+} \wedge \Phi_{2,+}) \rangle \Big|_{\xi=0}$, the wedge products description may be simpler in some cases (especially, in numerical computations).

Remark 6.14. • Starting from the canonical basis of $\mathbb{C}^n$, named $(e_i)_{i \in \llbracket 1, n \rrbracket}$, we can define the exterior algebra, named also Grassmann algebra, $\wedge^k(\mathbb{C}^n)$, which is a $\binom{n}{k}$ -dimensional subspace whose the induced orthonormal basis is $\{e_{i_1} \wedge e_{i_2} \wedge \dots \wedge e_{i_k}, 1 \leq i_1 < i_2 < \dots < i_k \leq n\}$.

• Every linear system $\frac{d}{d\xi} X(\xi) = \mathbb{B}(\xi) X(\xi)$ on $\mathbb{C}^n$, with $\mathbb{B}(\xi)$ a $n \times n$ matrix, may induce a linear system on $\wedge^k(\mathbb{C}^n)$:

$$\frac{d}{d\xi} X^{[k]}(\xi) = \mathbb{B}^{[k]}(\xi) X^{[k]}(\xi),$$

with $X^{[k]}(\xi) := X_1 \wedge X_2 \wedge \dots \wedge X_k$, $X_i \in \mathbb{C}^n$ and $\mathbb{B}^{[k]}(\xi)$ the following operator

$$\mathbb{B}^{[k]}(\xi) (X_1 \wedge X_2 \wedge \dots \wedge X_k) = \sum_{j=1}^k X_1 \wedge \dots \wedge \mathbb{B}(\xi) X_j \wedge \dots \wedge X_k.$$

More details on the construction of $\mathbb{B}^{[k]}(\xi)$ operator may be found in [1] or [26].

Remark 6.15. If $n - k < k$, a better choice would be to compute $\mathbb{B}^{[n-k]}$, as for the (KdV⁻ - B) case.

By applying the previous reminder on the ρ -dependent first-order linear system (6.4) on $\mathbb{C}^4$, we deduce a ρ -dependent linear system on $\wedge^2(\mathbb{C}^4)$:

$$\frac{d}{d\xi} W^{[2]} = \mathbb{A}^{[2]}(\rho, \xi) W^{[2]}(\xi), \quad (\rho, \xi) \in \mathbb{C} \times \mathbb{R}, \quad (6.8)$$

with $\mathbb{A}^{[2]}(\rho, \xi)(W_1 \wedge W_2) = (\mathbb{A}(\rho, \xi) W_1) \wedge W_2 + W_1 \wedge (\mathbb{A}(\rho, \xi) W_2)$.

Section 6.4. Numerical computation of the Evans function

We obtain for $\mathbb{A}^{[2]}(\rho, \xi)$ the following matrix

$$\mathbb{A}^{[2]}(\rho, \xi) := \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ c - u(\xi) & 0 & \Lambda & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \rho + \frac{du}{d\xi}(\xi) & 0 & 0 & 0 & \Lambda & 1 \\ 0 & \rho + \frac{du}{d\xi}(\xi) & 0 & -c + u(\xi) & 0 & \Lambda \end{pmatrix}.$$

Remark 6.16. If $\kappa_i^\pm(\rho)$, $i \in \{1, 2, 3, 4\}$ are the eigenvalues of $\mathbb{A}_\pm(\rho)$. Then, $\kappa_i^\pm + \kappa_j^\pm$, $(i, j) \in \{1, 2, 3, 4\}^2$, $i \neq j$ are the eigenvalues of $\mathbb{A}_\pm^{[2]}(\rho)$. Its biggest real part eigenvalue is thus $\kappa_3^\pm + \kappa_4^\pm$, and its smallest : $\kappa_1^\pm + \kappa_2^\pm$.

The function $\Phi_{1,+} \wedge \Phi_{2,+}$ appears as a solution of (6.8) decreasing at maximal rate $\kappa_1^+ + \kappa_2^+$ in $+\infty$ and $\Psi_{1,-} \wedge \Psi_{2,-}$ as a solution of the adjoint wedge ODE

$$\frac{d}{d\xi} Z^{[2]} = -Z^{[2]}(\xi) \mathbb{A}^{[2]}(\bar{\rho}, \xi). \quad (6.9)$$

This wedge description is used to compute Evans function numerically. We follow the same algorithm of [26] or [1].

Remark 6.17. Numerically, translating the spectrum of $\mathbb{A}^{[2]}(\rho, \xi)$ or $[\mathbb{A}^{[2]}(\rho, \xi)]^*$ by the smallest eigenvalue is more robust than solving directly (6.8) or (6.9). In that case, solutions $\Phi_{1,+} \wedge \Phi_{2,+}$ and $\Psi_{1,-} \wedge \Psi_{2,-}$ are computed as solutions of (6.8) or (6.9) starting from ζ^+ in $+\infty$ or η^- in $-\infty$, which are eigenvectors associated to $\kappa_1^\pm + \kappa_2^\pm$, as explained in the following algorithm.

Algorithm : Computation of D_- for (KdV⁻ 4) equation

- Let $\kappa_1^\pm + \kappa_2^\pm$ be the eigenvalue of $\mathbb{A}_\pm^{[2]}(\rho)$ with the smallest real part. Let $\zeta^+(\rho)$ and $\eta^-(\rho)$ be respectively the associated eigenvectors for $\mathbb{A}_+^{[2]}(\rho)$ and $[\mathbb{A}_-^{[2]}(\rho)]^*$, such that $\langle \zeta^+(\rho), \eta^-(\rho) \rangle = 1$.
- Determination of $\Phi_{1,+} \wedge \Phi_{2,+}$: Compute $W_\rho^{[2]}$ solution of

$$\frac{dW_\rho^{[2]}}{d\xi}(\xi) = [\mathbb{A}^{[2]}(\rho, \xi) - (\kappa_1^+ + \kappa_2^+)I] W_\rho^{[2]}(\xi) \quad \text{such that } W_\rho^{[2]}(\xi) \xrightarrow{\xi \rightarrow +\infty} \zeta^+(\rho)$$

- Determination of $\Psi_{1,-} \wedge \Psi_{2,-}$: Compute $Z_\rho^{[2]}$ solution of

$$\frac{dZ_\rho^{[2]}}{d\xi}(\xi) = -Z_\rho^{[2]}(\xi) [\mathbb{A}^{[2]}(\rho, \xi) - (\kappa_1^- + \kappa_2^-)I] \quad \text{such that } Z_\rho^{[2]}(\xi) \xrightarrow{\xi \rightarrow -\infty} \eta^-(\rho)$$

- Determination of the Evans function : $D_-(\rho) = \langle Z_\rho^{[2]}(0), W_\rho^{[2]}(0) \rangle$.
-

In the following, we present some numerical results for the Evans function applied on the (KdV⁻ 4) equation. We have performed three tests : the first one on the positive real line $\mathbb{R}^+$, the second one on the imaginary axis $\{\rho = it \text{ with } t \in \mathbb{R}\}$ (for computations, we choose $t \in [-10000, 10000]$) and the last one on a complex domain Ω in the neighborhood of the origin

$$\Omega = \{\rho \in \mathbb{C}, \text{ such that } \Re(\rho) \in [0, 6], \Im(\rho) \in [-3, 3]\}.$$

For all cases, we have fixed $u_+ = -1$, $u_- = 1$ and $\Lambda = \frac{\mu}{\xi^3} = 1$. Figure 6.2 corresponds to the Evans function D_- on the positive real line, Figure 6.3 to the computation of D_- on the imaginary axis and Figure 6.4 to the modulus of D_- on the domain Ω .

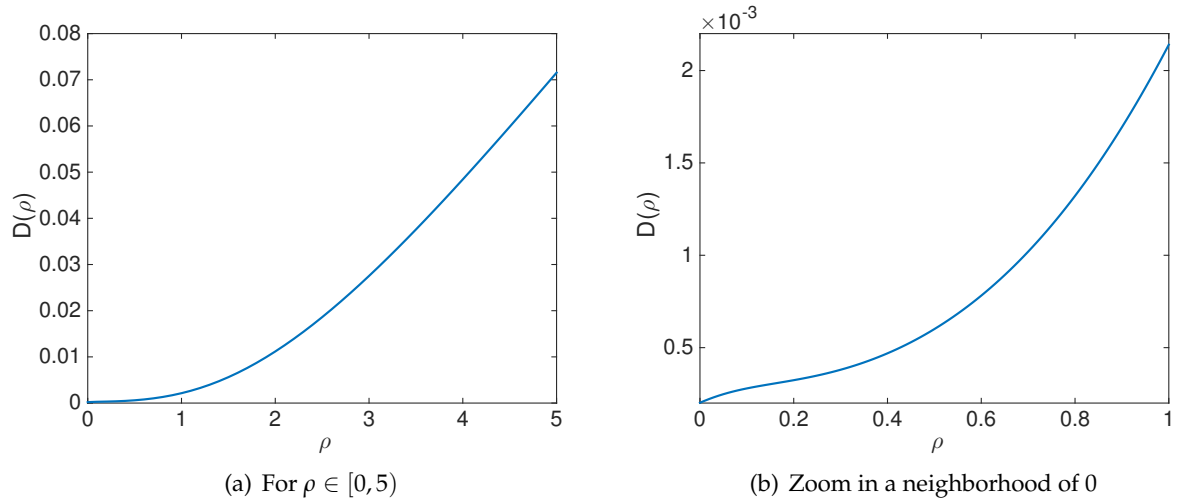


FIGURE 6.2 – Evans function $D_-(\rho)$ for the (KdV⁻ - 4) equation on the real axis

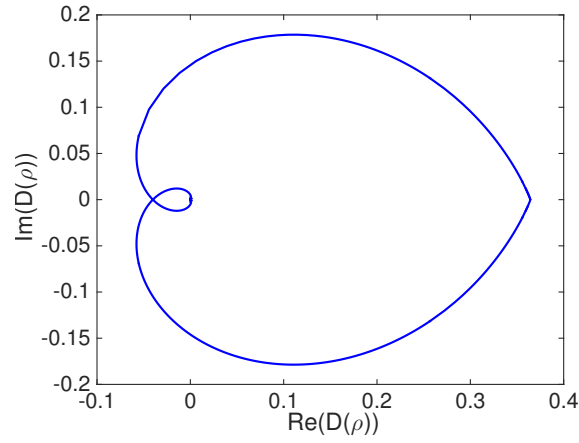


FIGURE 6.3 – Evans function D_- on the imaginary axis, for the (KdV⁻ - 4) equation

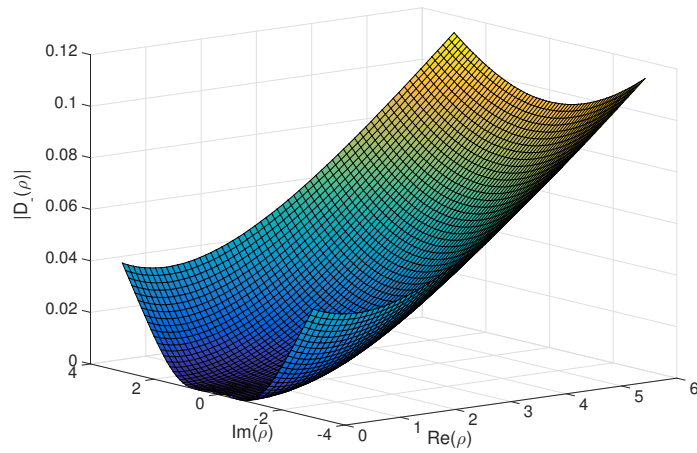


FIGURE 6.4 – Evans function D_- on the domain Ω , for the (KdV⁻ - 4) equation

Section 6.4. Numerical computation of the Evans function

The illustration of the Cauchy problem starting from the indicator function, see Figure 5.1(a), let us suppose that those traveling wave is stable, since the solution seems to converge towards the traveling wave. This intuition is enforced by the previous numerical study of the Evans function : we may see that the Evans function does not vanish on $\mathbb{C}_*^+$ (except for $\rho = 0$), which implies that the linear operator $\mathcal{M}$ has no zero on the real line. The traveling wave seems to be spectrally stable.

However, this is a preliminary result because we have to compute theoretically D_- .

Remark 6.18. The relation $D_-(0) = 0$ is expected because $\frac{du}{d\tilde{\xi}}$ (the derivative of the traveling wave) verifies $\mathcal{M}\frac{du}{d\tilde{\xi}} = 0$. As it belongs to $L^2(\mathbb{R})$, it appears as an eigenfunction of $\mathcal{M}$ for the eigenvalue $\rho = 0$. Therefore, it is normal for D_- to vanish in $\rho = 0$.

Remark 6.19. A property of the Evans function is illustrated by those figures :

$$D_-(\bar{\lambda}) = \overline{D_-(\lambda)}, \quad \lambda \in \mathbb{C},$$

where $\bar{\cdot}$ stands for the complex conjugate.

- Indeed, the Evans function on the imaginary axis is perfectly symmetric in Figure 6.3.
- We recover this symmetry when D_- is computed on a neighborhood of the origin : Figure 6.4.
- From this property, one concludes that $D_-(\lambda) \in \mathbb{R}$ for $\lambda \in \mathbb{R}$, as shown in Figure 6.2 which corresponds to the Evans function on the positive real line.

6.4.2 In the real high-frequency limit

Studying numerically the Evans function D_- only in a neighborhood of the origin is relevant because we can prove that D_- does not vanish in the real high-frequency limit. More precisely, the following proposition is true.

Proposition 6.20. There exists $\rho_0 > 0$ such that for all $\rho > \rho_0$, the Evans function D_- does not vanish.

The proof is based on the homotopy method developed in [11] and used, for instance, for diffusion-dispersion equations in [72].

For $\theta \in [0, 1]$ and $\alpha_0 > 0$ fixed, we build the homotopy operator

$$\mathcal{M}^{KdV-4}(\theta) := -\theta \frac{du}{d\tilde{\xi}} I + \theta(c - u) \frac{d}{d\tilde{\xi}} + \theta \frac{\mu}{\tilde{\xi}^{\frac{2}{3}}} \frac{d^3}{d\tilde{\xi}^3} - \theta \frac{d^4}{d\tilde{\xi}^4} - (1 - \theta)\alpha_0 \frac{d^4}{d\tilde{\xi}^4}.$$

The spectral problem associated to $\mathcal{M}^{KdV-4}(\theta)$ writes

$$-\theta \frac{du}{d\tilde{\xi}} w + \theta(c - u) \frac{dw}{d\tilde{\xi}} + \theta \frac{\mu}{\tilde{\xi}^{\frac{2}{3}}} \frac{d^3 w}{d\tilde{\xi}^3} - \theta \frac{d^4 w}{d\tilde{\xi}^4} - (1 - \theta)\alpha_0 \frac{d^4 w}{d\tilde{\xi}^4} = \rho w, \quad \rho \in \mathbb{C}, \quad w \in H^4(\mathbb{R}), \quad \theta \in [0, 1], \quad (6.10)$$

and the ρ -dependent first-order dynamical system becomes

$$\frac{d}{d\tilde{\xi}} W(\tilde{\xi}, \theta) = \mathbb{A}^{KdV-4}(\rho, \tilde{\xi}, \theta) W(\tilde{\xi}, \theta), \quad (\rho, \tilde{\xi}, \theta) \in \mathbb{C} \times \mathbb{R} \times [0, 1],$$

with the companion matrix

$$\mathbb{A}^{KdV-4}(\rho, \tilde{\xi}, \theta) := \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \frac{-\rho - \theta \frac{du}{d\tilde{\xi}}(\tilde{\xi})}{\theta + (1 - \theta)\alpha_0} & \frac{\theta c - \theta u(\tilde{\xi})}{\theta + (1 - \theta)\alpha_0} & 0 & \frac{\theta \Lambda}{\theta + (1 - \theta)\alpha_0} \end{pmatrix}, \quad \text{with } \Lambda := \frac{\mu}{\tilde{\xi}^{\frac{2}{3}}}. \quad (6.11)$$

As before, both limit matrices $\mathbb{A}_{\pm}^{KdV-4}(\rho, \theta) := \lim_{\xi \rightarrow \pm\infty} \mathbb{A}^{KdV-4}(\rho, \xi, \theta)$ have two eigenvalues with positive real part (named $\kappa_{3,4}^{\pm}(\rho, \theta)$) and two eigenvalues with negative real part (named $\kappa_{1,2}^{\pm}(\rho, \theta)$), this allows to define a two-parameters Evans function

$$D_{-}(\rho, \theta) := \det \begin{pmatrix} \langle \Psi_{1,-}, \Phi_{1,+} \rangle & \langle \Psi_{1,-}, \Phi_{2,+} \rangle \\ \langle \Psi_{2,-}, \Phi_{1,+} \rangle & \langle \Psi_{2,-}, \Phi_{2,+} \rangle \end{pmatrix} (\rho, \xi, \theta) \Big|_{\xi=0}.$$

Remark 6.21. For $\theta = 1$, we recognize $\mathcal{M}^{KdV-4}(1) = \mathcal{M}^{KdV-4}$ and $D_{-}(\rho, 1) = D_{-}(\rho)$, defined in Section 6.3 and Subsection 6.4.1.

The key point of this paragraph is summarized in the following lemma.

Lemma 6.22. *There exists $\rho_0 > 0$ such that, for all $\theta \in [0, 1]$, the operator $\mathcal{M}^{KdV-4}(\theta)$ has no eigenvalue ρ in $(\rho_0, +\infty)$.*

Proof of Lemma 6.22. Let $\rho \in \mathbb{R}$ and $\theta \in (0, 1]$. Suppose that there exists $w \neq 0$, such that $\mathcal{M}^{KdV-4}(\theta)w = \rho w$.

On the one hand, multiplying the spectral problem (6.10) by w and integrating it over $\mathbb{R}$ gives

$$\begin{aligned} - \int_{\mathbb{R}} \theta \frac{du}{d\xi}(\xi) w^2(\xi) d\xi + \int_{\mathbb{R}} \frac{\theta}{2} (c - u(\xi)) \frac{d(w^2)}{d\xi}(\xi) d\xi + \int_{\mathbb{R}} \theta \Lambda \frac{d^3 w}{d\xi^3}(\xi) w(\xi) d\xi \\ - \int_{\mathbb{R}} \theta \frac{d^4 w}{d\xi^4}(\xi) w(\xi) d\xi - \int_{\mathbb{R}} (1 - \theta) \alpha_0 \frac{d^4 w}{d\xi^4}(\xi) w(\xi) d\xi = \rho \int_{\mathbb{R}} w^2(\xi) d\xi. \end{aligned}$$

Some integrations by parts give

$$- \int_{\mathbb{R}} \frac{\theta}{2} \frac{du}{d\xi}(\xi) w^2(\xi) d\xi - \int_{\mathbb{R}} \theta \left(\frac{d^2 w}{d\xi^2}(\xi) \right)^2 d\xi - \int_{\mathbb{R}} (1 - \theta) \alpha_0 \left(\frac{d^2 w}{d\xi^2}(\xi) \right)^2 d\xi = \rho \int_{\mathbb{R}} w^2(\xi) d\xi.$$

Since, $\theta \in (0, 1]$, one simplifies the previous equality into

$$\rho \|w\|_{L^2(\mathbb{R})}^2 \leq \frac{1}{2} \left\| \frac{du}{d\xi} \right\|_{L^\infty(\mathbb{R})} \|w\|_{L^2(\mathbb{R})}^2 + (1 + \alpha_0) \left\| \frac{d^2 w}{d\xi^2} \right\|_{L^2(\mathbb{R})}^2. \quad (6.12)$$

On the other hand, multiplying the spectral problem (6.10) by $\frac{dw}{d\xi}$ and integrating it over $\mathbb{R}$ gives

$$\begin{aligned} - \int_{\mathbb{R}} \theta \frac{du}{d\xi}(\xi) w(\xi) \frac{dw}{d\xi}(\xi) d\xi + \int_{\mathbb{R}} \theta (c - u(\xi)) \left(\frac{dw}{d\xi}(\xi) \right)^2 d\xi + \int_{\mathbb{R}} \theta \Lambda \frac{d^3 w}{d\xi^3}(\xi) \frac{dw}{d\xi}(\xi) d\xi \\ - \int_{\mathbb{R}} \theta \frac{d^4 w}{d\xi^4}(\xi) \frac{dw}{d\xi}(\xi) d\xi - \int_{\mathbb{R}} (1 - \theta) \alpha_0 \frac{d^4 w}{d\xi^4}(\xi) \frac{dw}{d\xi}(\xi) d\xi = \rho \int_{\mathbb{R}} w(\xi) \frac{dw}{d\xi}(\xi) d\xi. \end{aligned}$$

Some integrations by parts give

$$\int_{\mathbb{R}} \frac{\theta}{2} \frac{d^2 u}{d\xi^2}(\xi) w^2(\xi) d\xi + \int_{\mathbb{R}} \theta (c - u(\xi)) \left(\frac{dw}{d\xi}(\xi) \right)^2 d\xi - \int_{\mathbb{R}} \theta \Lambda \left(\frac{d^2 w}{d\xi^2}(\xi) \right)^2 d\xi = 0.$$

Thus, one has

$$\theta \Lambda \left\| \frac{d^2 w}{d\xi^2} \right\|_{L^2(\mathbb{R})}^2 \leq \theta \|c - u\|_{L^\infty(\mathbb{R})} \left\| \frac{dw}{d\xi} \right\|_{L^2(\mathbb{R})}^2 + \frac{\theta}{2} \left\| \frac{d^2 u}{d\xi^2} \right\|_{L^\infty(\mathbb{R})} \|w\|_{L^2(\mathbb{R})}^2.$$

Section 6.4. Numerical computation of the Evans function

Since $\theta > 0$, this implies

$$\Lambda \left\| \frac{d^2 w}{d\tilde{\xi}^2} \right\|_{L^2(\mathbb{R})}^2 \leq \|c - u\|_{L^\infty(\mathbb{R})} \left\| \frac{dw}{d\tilde{\xi}} \right\|_{L^2(\mathbb{R})}^2 + \frac{1}{2} \left\| \frac{d^2 u}{d\tilde{\xi}^2} \right\|_{L^\infty(\mathbb{R})} \|w\|_{L^2(\mathbb{R})}^2.$$

Gagliardo-Nirenberg inequality $\|w'\|_{L^2(\mathbb{R})}^2 \leq \|w\|_{L^2(\mathbb{R})} \|w''\|_{L^2(\mathbb{R})}$ simplifies the previous relation into

$$\Lambda \left\| \frac{d^2 w}{d\tilde{\xi}^2} \right\|_{L^2(\mathbb{R})}^2 \leq \|c - u\|_{L^\infty(\mathbb{R})} \|w\|_{L^2(\mathbb{R})} \left\| \frac{d^2 w}{d\tilde{\xi}^2} \right\|_{L^2(\mathbb{R})} + \frac{1}{2} \left\| \frac{d^2 u}{d\tilde{\xi}^2} \right\|_{L^\infty(\mathbb{R})} \|w\|_{L^2(\mathbb{R})}^2.$$

Let $\mathcal{Q} := \frac{\left\| \frac{d^2 w}{d\tilde{\xi}^2} \right\|_{L^2(\mathbb{R})}}{\|w\|_{L^2(\mathbb{R})}} \geq 0$, then $\mathcal{Q}$ verifies

$$\Lambda \mathcal{Q}^2 \leq \|c - u\|_{L^\infty(\mathbb{R})} \mathcal{Q} + \frac{1}{2} \left\| \frac{d^2 u}{d\tilde{\xi}^2} \right\|_{L^\infty(\mathbb{R})}.$$

Choose $\mathcal{Q}_0 := \frac{\|c - u\|_{L^\infty(\mathbb{R})} + \sqrt{\|c - u\|_{L^\infty(\mathbb{R})}^2 + 2\Lambda \left\| \frac{d^2 u}{d\tilde{\xi}^2} \right\|_{L^\infty(\mathbb{R})}}}{2\Lambda}$, the positive root of

$$\Lambda X^2 - \|c - u\|_{L^\infty(\mathbb{R})} X - \frac{1}{2} \left\| \frac{d^2 u}{d\tilde{\xi}^2} \right\|_{L^\infty(\mathbb{R})},$$

one has $\mathcal{Q} \leq \mathcal{Q}_0$, which implies $\left\| \frac{d^2 w}{d\tilde{\xi}^2} \right\|_{L^2(\mathbb{R})} \leq \mathcal{Q}_0 \|w\|_{L^2(\mathbb{R})}$, by definition of $\mathcal{Q}$. Injecting this upper bound in Equation (6.12) gives

$$\rho \|w\|_{L^2(\mathbb{R})}^2 \leq \left(\frac{1}{2} \left\| \frac{du}{d\tilde{\xi}} \right\|_{L^\infty(\mathbb{R})} + (1 + \alpha_0) \mathcal{Q}_0^2 \right) \|w\|_{L^2(\mathbb{R})}^2.$$

If $\rho > \rho_0 := \left(\frac{1}{2} \left\| \frac{du}{d\tilde{\xi}} \right\|_{L^\infty(\mathbb{R})} + (1 + \alpha_0) \mathcal{Q}_0^2 \right)$, there is not non-zero solution w . Thus the operator $\mathcal{M}^{KdV-4}(\theta)$ has eigenvalue ρ only in $(-\infty, \rho_0]$, for $\theta \in (0, 1]$.

If $\theta = 0$, one has $\mathcal{M}^{KdV-4}(0) := -\alpha_0 \frac{d^4}{d\tilde{\xi}^4}$. Suppose that there exists $w \neq 0 \in H^4(\mathbb{R})$, such that $-\alpha_0 \frac{d^4}{d\tilde{\xi}^4} w = \rho w$. By multiplying by $\frac{d^3 w}{d\tilde{\xi}^3}$ and integrating by part, we obtain

$$0 = -\rho \int_{\mathbb{R}} \left(\frac{d^2 w}{d\tilde{\xi}^2} \right)^2 (\tilde{\xi}) d\tilde{\xi}.$$

Therefore, $\mathcal{M}^{KdV-4}(0)$ has no eigenvalue for $\rho > 0$, *a fortiori*, for $\rho > \rho_0$. □

We are now able to prove Proposition 6.20.

Proof of Proposition 6.20. Thanks to Lemma 6.22, there exists $\rho_0 > 0$, such that $\mathcal{M}^{KdV-4}(\theta)$ has no eigenvalue ρ in $(\rho_0, +\infty)$, for all $\theta \in [0, 1]$. However, the eigenvalues of $\mathcal{M}^{KdV-4}(\theta)$ and the zeros of $D_-(\cdot, \theta)$ coincide, according to Proposition 6.7. Thus, $D_-(\cdot, \theta)$ does not vanish for $\rho \in (\rho_0, +\infty)$, and for all $\theta \in [0, 1]$.

Therefore, thanks to Remark 6.21 and thanks to homotopy, we conclude that, for all $\rho_1 > \rho_0$

$$\text{sgn}(D_-(\rho_1)) = \text{sgn}(D_-(\rho_1, 1)) = \text{sgn}(D_-(\rho_1, 0)) = \lim_{\rho \rightarrow +\infty} \text{sgn}(D_-(\rho, 0)) \neq 0.$$

□

Remark 6.23. We have thus to better understand $D_-(\rho, 0)$, for $\rho > \rho_0$. Using the limit companion matrices

$$\mathbb{A}_{\pm}^{KdV-4}(\rho_1, 0) := \lim_{\xi \rightarrow \pm\infty} \mathbb{A}^{KdV-4}(\rho_1, \xi, 0) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{\rho_1}{\alpha_0} & 0 & 0 & 0 \end{pmatrix},$$

one can prove that $D_-(\rho, 0) > 0$. The method is detailed for a diffusion of second order-dispersion equation in [72].

Remark 6.24. Proposition 6.20 is a preliminary work because it enables to study the Evans function only for the real high-frequency. A further analysis for $\rho \in \mathbb{C}$ with $|\rho| > \rho_0$ is needed.

6.5 Afterword : Stability for the traveling wave of (KdV⁻ - B) equation

For the (KdV⁻ - B) case, the relevant ratio is $\frac{\lambda}{\sqrt{\mu}}$ and so the ansatz is chosen such that

$$\bar{v}\left(\frac{t}{\sqrt{\mu}}, \frac{x-ct}{\sqrt{\mu}}\right) := u\left(\frac{x-ct}{\sqrt{\mu}}\right) + \bar{w}\left(\frac{t}{\sqrt{\mu}}, \frac{x-ct}{\sqrt{\mu}}\right).$$

The linearized operator becomes in the (KdV⁻ - B) case

$$\mathcal{M}^{KdV-B} := -\frac{du}{d\xi}I + (c-u)\frac{d}{d\xi} + \frac{\lambda}{\sqrt{\mu}}\frac{d^2}{d\xi^2} + \frac{d^3}{d\xi^3}$$

and the companion matrix

$$\mathbb{A}^{KdV-B}(\rho, \xi) := \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \rho + \frac{du}{d\xi}(\xi) & -c + u(\xi) & -\frac{\lambda}{\sqrt{\mu}} \end{pmatrix}.$$

For (KdV⁻ - B) equation, the characteristic polynomial of $\mathbb{A}_{\pm}^{KdV-B}(\rho)$ is $\chi(x) := -x^3 - \frac{\lambda}{\sqrt{\mu}}x^2 - (c - u_{\pm})x + \rho$. Thus, the eigenvalues $\kappa_i^{\pm}(\rho)$ satisfy

$$\Re(\kappa_1^{\pm}(\rho)), \Re(\kappa_2^{\pm}(\rho)) < 0 < \Re(\kappa_3^{\pm}(\rho))$$

and

$$\dim(\mathbb{E}_+^s(\rho)) = \dim(\mathbb{E}_-^{s,*}(\rho)) = 2,$$

whereas

$$\dim(\mathbb{E}_-^u(\rho)) = \dim(\mathbb{E}_+^{u,*}(\rho)) = 1.$$

The D_+ definition (cf Remark 6.8) is simpler, in that case, because the determinant $\langle \Psi_{i,+}, \Phi_{j,-} \rangle$ is reduced to a scalar. So we prefer to use it. Moreover, since $\dim(\mathbb{E}_-^u(\rho)) = 1$, the Grassmann algebra is $\wedge^1(\mathbb{C}^3) = \mathbb{C}^3$ and $\mathbb{A}^{[1]}(\rho, \xi) = \mathbb{A}(\rho, \xi)$. The algorithm that we follow is then the same as for (KdV⁻ - 4) with D_+ instead of D_- (and the biggest eigenvalue $\kappa_3^{\pm}(\rho)$ instead of the smallest).

Algorithm : Computation of D_+ for (KdV⁻ - B) equation

- Let $\kappa_3^{\pm}$ be the (only) eigenvalue of $\mathbb{A}_{\pm}(\rho)$ with positive real part. Let $\zeta^-(\rho)$ and $\eta^+(\rho)$ be respectively the associated eigenvectors for $\mathbb{A}_-(\rho)$ and $[\mathbb{A}_+(\rho)]^*$, such that $\langle \zeta^-(\rho), \eta^+(\rho) \rangle = 1$.
- Determination of Φ_- : Compute W_ρ solution of

$$\frac{dW_\rho}{d\xi}(\xi) = [\mathbb{A}(\rho, \xi) - \kappa_3^- I] W_\rho(\xi) \quad \text{such that } W_\rho(\xi) \xrightarrow{\xi \rightarrow -\infty} \zeta^-(\rho)$$

- Determination of Ψ_+ : Compute Z_ρ solution of

$$\frac{dZ_\rho}{d\xi}(\xi) = -Z_\rho(\xi) [\mathbb{A}(\rho, \xi) - \kappa_3^+ I] \quad \text{such that } Z_\rho(\xi) \xrightarrow{\xi \rightarrow +\infty} \eta^+(\rho)$$

- Determination of the Evans function : $D_+(\rho) = \langle Z_\rho(0), W_\rho(0) \rangle$.
-

Remark 6.25. The results are similar to those of [87], the difference being that the authors study (KdV⁻ - B) equation with $\mu < 0$. The situation is thus reverse and they use preferentially D_- instead of D_+ .

We have performed the same type of numerical computations as the ones for the (KdV⁻ - 4) equation, with the parameters $u_+ = -1$, $u_- = 1$ and $\frac{\lambda}{\sqrt{\mu}} = 1$.

In Figure 6.5, we represent D_+ for (KdV⁻ - B) equation, only on the non negative real line $\mathbb{R}^*$. The image by D_+ of the imaginary axis is represented in Figure 6.6. Eventually, in Figure 6.7, the modulus of $D_+(\rho)$ is represented for $\rho \in \Omega$.

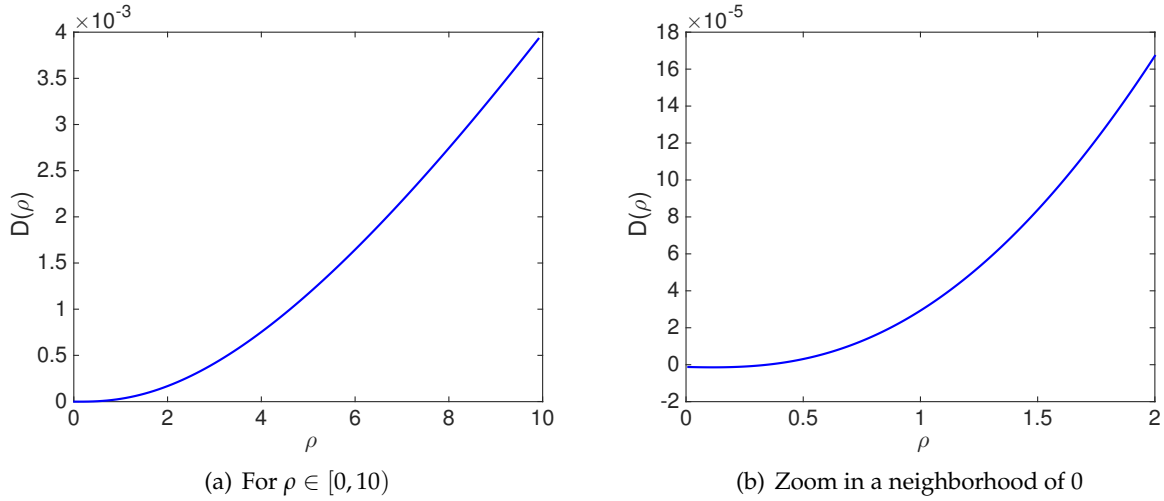


FIGURE 6.5 – Evans function for the (KdV⁻ - B) equation

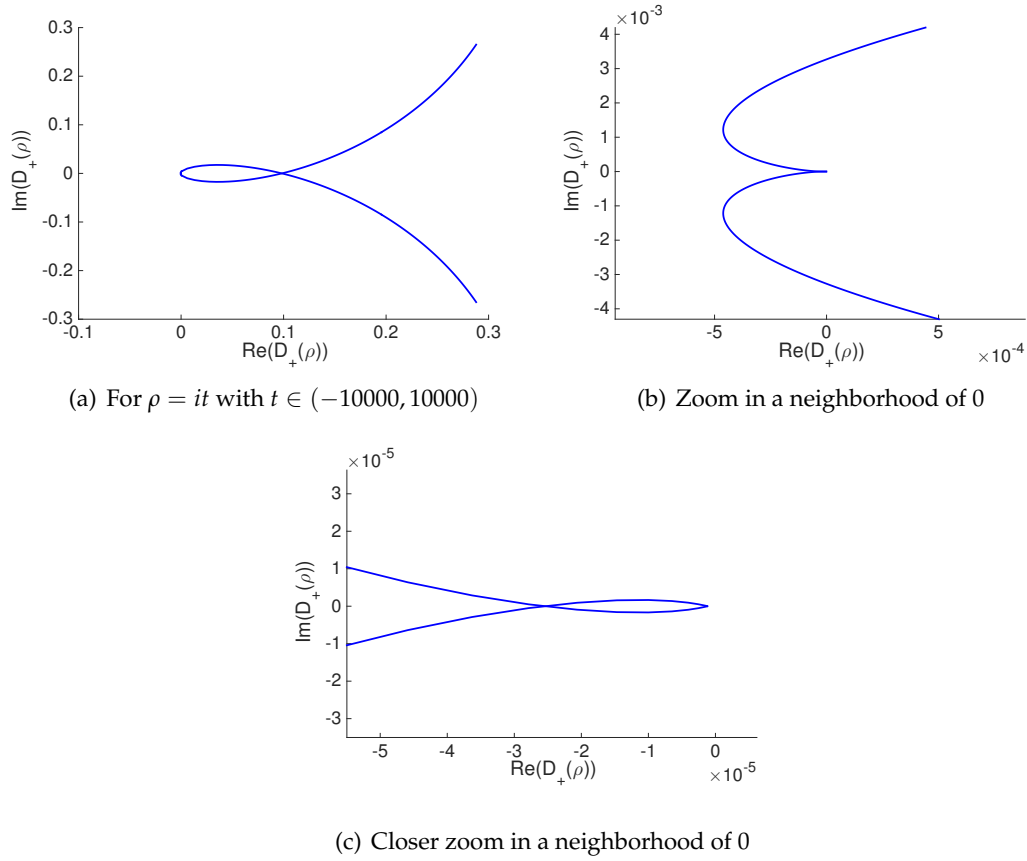


FIGURE 6.6 – Evans function D_+ for the (KdV⁻ - B) equation on the imaginary axis

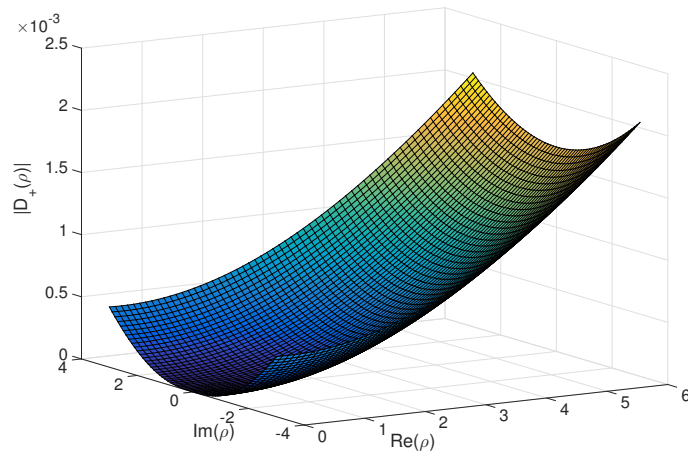


FIGURE 6.7 – Evans function on the domain Ω , for the (KdV⁻ - B) equation

As for (KdV⁻ - 4) equation, Evans function does not vanish on $\mathbb{C}_*^+$ (except for $\rho = 0$) and the linearized operator seems not have zero on $\mathbb{C}_*^+$. However, this intuition has to be enforced by the theoretical study of D_+ on $\mathbb{C}_*^+$ to really conclude on the stability of the traveling wave.

Conclusion et perspectives

Nous avons étudié d'un point de vue essentiellement numérique certaines équations aux dérivées partielles hyperboliques-dispersives. Nous nous sommes focalisés sur des schémas aux différences finies et avons étudié l'ordre de convergence de ces schémas par la méthode classique de consistance et de stabilité.

Ordre de convergence des schémas pour KdV. Tout d'abord la stabilité des schémas décentrés dépend du terme dispersif : lorsque le terme dispersif est de la forme ∂_x^{2p+1} , les schémas décentrés à droite seront stables (sous condition CFL) si p est impair et ceux décentrés à gauche seront stables (sous condition CFL) si p est pair.

La stabilité est la partie la plus délicate de l'étude puisqu'il faut prendre en compte en même temps les termes hyperboliques et dispersifs qui apparaissent dans les équations étudiées. Cette étape est déterminée par une méthode directe d'intégrations par parties discrètes qui revient à transposer d'un point de vue discret la stabilité fort-faible des lois de conservation hyperbolique.

Nous avons déterminé l'ordre de convergence des schémas mais surtout nous les avons quantifiés par rapport à la régularité de la donnée initiale. Lorsque cette dernière était trop peu régulière pour assurer les estimations de consistance, nous l'avons régularisée au moyen d'une suite régularisante. Nous avons dû alors trouver un compromis entre cette régularisation et l'ordre de convergence des schémas pour ne pas trop perdre en convergence. L'étape de stabilité est toujours valable même pour de faibles régularités car nous avons considéré des données au moins dans $H^{\frac{3}{4}}(\mathbb{R})$ ce qui impose à la solution exacte d'être bien définie globalement et au facteur de l'inégalité fort-faible d'être indépendant du paramètre de régularisation.

Nous avons généralisé cette étude aux systèmes $abcd$.

Ondes progressives. Nous avons démontré l'existence d'ondes progressives de petite amplitude pour une équation de KdV-KS sans terme de diffusion d'ordre deux. Nous nous ramè-nons tout d'abord à étudier les trajectoires proches de 0 par projection sur la variété centrale.

Nous avons étudié le système limite (lorsque les états initial et final sont égaux) et démontré que l'homocline du système limite singulier se transformait en orbite hétérocline pour le système perturbé (lorsque les deux états sont distincts).

Dans une étude très prospective, nous avons commencé à étudier numériquement la fonction d'Evans. Les premiers résultats semblent conclure en faveur de la stabilité spectrale des ondes progressives.



Terminons ce manuscrit par quelques perspectives à courts et moyens termes.

Partie III. Le système $abcd$ tel qu'il est décrit dans [16] possède en réalité un petit paramètre ϵ au niveau des termes de dispersion et de non linéarité, qui correspond au paramètre de non linéarité et au paramètre de faible profondeur (cf page 5, avec $\epsilon = \mu$)

$$\begin{cases} (I - \epsilon b \partial_x^2) \partial_t \eta + (I + \epsilon a \partial_x^2) \partial_x u + \epsilon \partial_x (\eta u) = 0, \\ (I - \epsilon d \partial_x^2) \partial_t u + (I + \epsilon c \partial_x^2) \partial_x \eta + \epsilon \partial_x \left(\frac{u^2}{2} \right) = 0. \end{cases}$$

Pour des temps assez faibles, seuls les phénomènes de transport sont vraiment significatifs, ce qui n'est plus le cas à partir d'un certain temps à cause des termes dispersifs et non-linéaires. Une question naturelle est alors de savoir ce que deviennent les résultats numériques de stabilité et de convergence lorsque ce petit paramètre ϵ est pris en compte. La difficulté serait d'obtenir des majorations uniformes en ϵ afin de ne pas avoir un ordre de convergence en $\frac{\Delta x}{\epsilon}$. L'utilisation de schémas de type *Asymptotic Preserving* pourrait être un premier élément de réponse.



Partie IV. L'étude de la stabilité des ondes progressives n'est qu'amorcée dans ce manuscrit et permet de conjecturer que les ondes progressives sont spectralement stables. Une étude théorique plus poussée serait intéressante pour justifier cette conjecture.

Pour connaître le nombre de zéros de la fonction d'Evans dans le demi-espace $\mathbb{C}_*^+$, une idée est de calculer l'indice (ou *winding number* en anglais) c'est-à-dire le nombre de fois que l'image de l'axe imaginaire pur par $D_- : \{D_-(\rho), \rho = it, t \in \mathbb{R}\}$ «tourne» autour de l'origine, cf [87].

Dans un second temps, une étude théorique de la fonction d'Evans serait intéressante. Il serait délicat de la déterminer pour tout ρ et une première approche serait de se restreindre à $\rho \in \mathbb{R}^+$. Une méthode classique consiste à calculer l'indice de stabilité Γ , cf [12], qui revient à comparer les signes de la fonction d'Evans en $+\infty$ et de sa dérivée en 0. Si ces deux signes sont égaux, l'argument est insuffisant et l'étude doit être approfondie. Par contre, s'ils sont différents, la fonction d'Evans s'annule nécessairement en (au moins) un point de l'axe réel positif (d'après le théorème des valeurs intermédiaires) et l'onde progressive n'est dans ce cas pas stable par annulation de la fonction d'Evans sur $\mathbb{R}^+$.



CINQUIÈME PARTIE

Annexes

ANNEXE A

Computations of consistency errors



We recall in this annex the computation of consistency error for several equations. The definition of consistency error is given in the general introduction, in Definition 1.17.

Remark A.1. In this appendix, C will denote any positive constant independent of $u, \eta, \Delta x$ and Δt , which can differ from one line to the other.

A.1 Consistency error for the KdV equation (Chapter 3)

The consistency error is defined by (3.6). For simplicity, we note in this appendix $t^{n+1} = \inf(t^{n+1}, T)$ and $\Delta t = t^{n+1} - t^n$, which corresponds to the following expression, for $(n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}$

$$\begin{aligned}
 \epsilon_j^n &= \frac{1}{\Delta t^2 \Delta x} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} u(s + \Delta t, y) - u(s, y) dy ds \\
 &\quad + \frac{1}{4\Delta x} \overbrace{\left[\left(\frac{1}{\Delta t \Delta x} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} u(s, y + \Delta x) dy ds \right)^2 - \left(\frac{1}{\Delta x \Delta t} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} u(s, y - \Delta x) dy ds \right)^2 \right]}^{=: (a)} \\
 &\quad + \frac{1 - \theta}{\Delta t \Delta x^4} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} u(s, y + 2\Delta x) - 3u(s, y + \Delta x) + 3u(s, y) - u(s, y - \Delta x) dy ds \\
 &\quad + \frac{\theta}{\Delta t \Delta x^4} \int_{t^{n+1}}^{t^{n+2}} \int_{x_j}^{x_{j+1}} u(s, y + 2\Delta x) - 3u(s, y + \Delta x) + 3u(s, y) - u(s, y - \Delta x) dy ds \\
 &\quad - \tau \left(\frac{1}{2\Delta t \Delta x^2} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} u(s, y + \Delta x) - 2u(s, y) + u(s, y - \Delta x) dy ds \right).
 \end{aligned} \tag{A.1}$$

Let us first deal with (a)-term. In order to simplify the following equations, we denote, for v in $\mathbb{R}$

$$K(v) := \left(\frac{1}{\Delta x \Delta t} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} u(s, y + v\Delta x) ds dy \right)^2.$$

A straightforward computation yields

$$\begin{aligned}
 K'(0) &= \frac{2}{\Delta x \Delta t^2} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \partial_x u(\bar{s}, \bar{y}) u(s, y) d\bar{s} d\bar{y} ds dy \\
 &= \frac{2}{\Delta x \Delta t^2} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \left[\partial_x u(s, y) + \int_y^{\bar{y}} \partial_x^2 u(s, v) dv + \int_s^{\bar{s}} \partial_{xt} u(\tau, \bar{y}) d\tau \right] u(s, y) d\bar{s} d\bar{y} ds dy \\
 &= \frac{2}{\Delta t} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} u(s, y) \partial_x u(s, y) ds dy + \frac{2}{\Delta t \Delta x} \int_{x_j}^{x_{j+1}} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} u(s, y) \int_y^{\bar{y}} \partial_x^2 u(s, v) dv ds d\bar{y} dy \\
 &\quad + \frac{2}{\Delta t^2 \Delta x} \int_{x_j}^{x_{j+1}} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{t^n}^{t^{n+1}} u(s, y) \int_s^{\bar{s}} \partial_{xt} u(\tau, \bar{y}) d\tau d\bar{s} ds d\bar{y} dy.
 \end{aligned}$$

Thus, the (a) term in Equation (A.1) rewrites

$$(a) = K(1) - K(-1) = 2K'(0) + \int_0^1 K''(w)(1-w)dw + \int_0^1 K''(-w)(-1+w)dw,$$

and after some Taylor expansions, since u is a solution of (KdV), it holds

$$\begin{aligned}
 |\epsilon_j^n| &\leq \frac{1}{\Delta x \Delta t} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_s^{s+\Delta t} |\partial_t^2 u(w, y)| dw ds dy + \frac{1}{\Delta t \Delta x^2} \int_{x_j}^{x_{j+1}} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} |u(s, y) \int_y^{\bar{y}} \partial_x^2 u(s, v) dv| ds d\bar{y} dy \\
 &\quad + \frac{1}{\Delta t^2 \Delta x^2} \int_{x_j}^{x_{j+1}} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{t^n}^{t^{n+1}} |u(s, y) \int_s^{\bar{s}} \partial_{xt} u(\tau, \bar{y}) d\tau| d\bar{s} ds d\bar{y} dy + \frac{1}{4\Delta x} \int_0^1 |K''(w)|(1-w)dw \\
 &\quad + \frac{1}{4\Delta x} \int_0^1 |K''(-w)|(1-w)dw + \frac{1-\theta}{\Delta t \Delta x} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_y^{y+2\Delta x} |\partial_x^4 u(s, z)| dz + \int_y^{y+\Delta x} |\partial_x^4 u(s, z)| dz \\
 &\quad + \int_y^{y-\Delta x} |\partial_x^4 u(s, z)| dz dy ds + \frac{\theta}{\Delta t \Delta x} \int_{t^{n+1}}^{t^{n+2}} \int_{x_j}^{x_{j+1}} \int_y^{y+2\Delta x} |\partial_x^4 u(s, z)| dz + \int_y^{y+\Delta x} |\partial_x^4 u(s, z)| dz \\
 &\quad + \int_y^{y-\Delta x} |\partial_x^4 u(s, z)| dz dy ds + \frac{\tau}{2\Delta x \Delta t} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_y^{y+\Delta x} |\partial_x^2 u(s, z)| dz + \int_y^{y-\Delta x} |\partial_x^2 u(s, z)| dz dy ds.
 \end{aligned}$$

Thus, once applying Cauchy-Schwarz inequality, the previous inequality becomes

$$\begin{aligned}
 |\epsilon_j^n|^2 &\leq C \left[\frac{\Delta t}{\Delta x} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \left(\partial_t^2 u(w, y) \right)^2 dw dy + \frac{1}{\Delta t \Delta x} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_y^{\bar{y}} \left[u(s, y) \partial_x^2 u(s, v) \right]^2 dv ds dy d\bar{y} \right. \\
 &\quad + \frac{1}{\Delta t \Delta x^2} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_s^{\bar{s}} \left[u(s, y) \partial_{xt} u(\tau, \bar{y}) \right]^2 d\tau d\bar{s} dy ds d\bar{y} + \frac{1}{8\Delta x^2} \|K''\|_{L^\infty([-1,1])}^2 \\
 &\quad + \frac{(1-\theta)^2 \Delta x}{\Delta t} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \left(\partial_x^4 u(s, z) \right)^2 dz ds + \frac{\theta^2 \Delta x}{\Delta t} \int_{t^{n+1}}^{t^{n+2}} \int_{x_j}^{x_{j+1}} \left(\partial_x^4 u(s, z) \right)^2 dz ds \\
 &\quad \left. + \frac{\tau^2 \Delta x}{\Delta t} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \left(\partial_x^2 u(s, z) \right)^2 dz ds \right],
 \end{aligned}$$

with C a constant.

However, one has

$$\begin{aligned}
 K''(v) &= \frac{2}{\Delta t^2} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \partial_x^2 u(s, y + v\Delta x) u(\bar{s}, \bar{y} + v\Delta x) d\bar{s} d\bar{y} ds dy \\
 &\quad + \frac{2}{\Delta t^2} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \partial_x u(s, y + v\Delta x) \partial_x u(\bar{s}, \bar{y} + v\Delta x) d\bar{s} d\bar{y} ds dy,
 \end{aligned}$$

which gives, thanks to Cauchy-Schwarz inequality

$$|K''(\nu)|^2 \leq C \left[\frac{\Delta x^3}{\Delta t^2} \int_{t^n}^{t^{n+1}} \|u(\bar{s}, \cdot)\|_{L_x^\infty}^2 \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} (\partial_x^2 u(s, y + \nu \Delta x))^2 ds dy d\bar{s} \right. \\ \left. + \left(\frac{2\Delta x}{\Delta t} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} (\partial_x u(s, y + \nu \Delta x))^2 ds dy \right)^2 \right].$$

Therefore, it yields

$$\sum_{j \in \mathbb{Z}} \Delta x |\epsilon_j^n|^2 \leq C \left[\Delta t^2 \sup_{t \in [0, T]} \|\partial_t^2 u(t, \cdot)\|_{L_x^2}^2 + \Delta x^2 \sup_{t \in [0, T]} \|u(t, \cdot)\|_{L_x^\infty}^2 \sup_{t \in [0, T]} \|\partial_x^2 u(t, \cdot)\|_{L_x^2}^2 \right. \\ + \Delta t^2 \sup_{t \in [0, T]} \|u(t, \cdot)\|_{L_x^\infty}^2 \sup_{t \in [0, T]} \|\partial_{xt} u(t, \cdot)\|_{L_x^2}^2 + \Delta x^2 \sup_{t \in [0, T]} \|u(t, \cdot)\|_{L_x^\infty}^2 \sup_{t \in [0, T]} \|\partial_x^2 u(t, \cdot)\|_{L_x^2}^2 \\ \left. + \Delta x^2 \sup_{t \in [0, T]} \|\partial_x u(t, \cdot)\|_{L_x^2}^2 \sup_{t \in [0, T]} \|\partial_x u(t, \cdot)\|_{L_x^2}^2 + \Delta x^2 \sup_{n \in \llbracket 0, N \rrbracket} \|\partial_x^4 u\|_{L_x^2}^2 + \Delta x^2 \sup_{n \in \llbracket 0, N \rrbracket} \|\partial_x^2 u\|_{L_x^2}^2 \right]. \quad (\text{A.2})$$

Let us then compute $\|\partial_t^2 u\|_{L_x^2}$ in (A.2). Thanks to the KdV equation, the time derivative is equal to

$$\partial_t^2 u = 2u (\partial_x u)^2 + u^2 \partial_x^2 u + 5\partial_x u \partial_x^3 u + 2u \partial_x^4 u + 3 (\partial_x^2 u)^2 + \partial_x^6 u.$$

For the term $\partial_x u \partial_x^3 u$, we use then the relation, for all u and v in $H^{\alpha+\beta}(\mathbb{R})$

$$\left\| \partial_x^\alpha u \partial_x^\beta v \right\|_{L^2(\mathbb{R})} \leq C \left[\|u\|_{L^\infty(\mathbb{R})} \|v\|_{H^{\alpha+\beta}(\mathbb{R})} + \|v\|_{L^\infty(\mathbb{R})} \|u\|_{H^{\alpha+\beta}(\mathbb{R})} \right]. \quad (\text{A.3})$$

Hence

$$\|\partial_t^2 u\|_{L_x^2} \leq C \left[\|u\|_{L_x^\infty} \|\partial_x u\|_{L_x^4}^2 + \|u\|_{L_x^\infty}^2 \|\partial_x^2 u\|_{L_x^2}^2 + \|u\|_{L_x^\infty} \|\partial_x^4 u\|_{L_x^2}^2 + \|u\|_{L_x^\infty} \|\partial_x^4 u\|_{L_x^2}^2 + \|\partial_x^2 u\|_{L_x^4}^2 + \|\partial_x^6 u\|_{L_x^2}^2 \right].$$

For the term $\|\partial_x u\|_{L_x^4}$, we use an integration by parts and Cauchy-Schwarz inequality to obtain

$$\|\partial_x u\|_{L_x^4}^4 = \int_{\mathbb{R}} (\partial_x u(x))^3 \partial_x u(x) dx = - \int_{\mathbb{R}} 3u(x) \partial_x^2 u(x) (\partial_x u(x))^2 dx \leq 3 \|u\|_{L_x^\infty} \|\partial_x^2 u\|_{L_x^2} \|\partial_x u\|_{L_x^4}^2.$$

We thus conclude $\|\partial_x u\|_{L_x^4}^2 \leq C \|u\|_{L_x^\infty} \|\partial_x^2 u\|_{L_x^2}$.

For the term $\|\partial_x^2 u\|_{L_x^4}^2$, we again use an integration by parts and the Cauchy-Schwarz inequality to write

$$\|\partial_x^2 u\|_{L_x^4}^4 = \int_{\mathbb{R}} (\partial_x^2 u(x))^3 \partial_x^2 u(x) dx = \int_{\mathbb{R}} -3\partial_x^3 u(x) (\partial_x^2 u(x))^2 \partial_x u(x) dx \\ \leq 3 \|\partial_x^2 u\|_{L_x^4}^2 \sqrt{\int_{\mathbb{R}} (\partial_x^3 u(x))^2 (\partial_x u(x))^2 dx},$$

which implies thanks to Relation (A.3) $\|\partial_x^2 u\|_{L_x^4}^2 \leq C \|u\|_{L_x^\infty} \|\partial_x^4 u\|_{L_x^2}$.

For the $\|\partial_{xt} u(t, \cdot)\|_{L_x^2}$ -term in (A.2), it holds

$$\|\partial_{tx} u(t, \cdot)\|_{L_x^2}^2 = \| -(\partial_x u(t, \cdot))^2 - u(t, \cdot) \partial_x^2 u(t, \cdot) - \partial_x^4 u(t, \cdot) \|_{L_x^2}^2 \\ \leq C \left[\|u(t, \cdot)\|_{L_x^\infty}^2 \|\partial_x^2 u(t, \cdot)\|_{L_x^2}^2 + \|\partial_x u(t, \cdot)\|_{L_x^4}^4 + \|\partial_x^4 u(t, \cdot)\|_{L_x^2}^2 \right].$$

Annexe A.2. Consistency error for the *abcd* system (Chapter 4)

To conclude, we obtain with (A.2)

$$\begin{aligned} \|\epsilon^n\|_{\ell^\infty(0,N,\ell_\Delta^2)} &\leq C \left[\Delta t \sup_{t \in [0,T]} \left(\|u\|_{L_x^\infty}^2 \|u\|_{H_x^2} + \|u\|_{L_x^\infty} \|u\|_{H_x^4} + \|u\|_{H_x^6} + \|u\|_{L_x^\infty} \|u\|_{H_x^2} + \|u\|_{H_x^4} \right) \right. \\ &\quad \left. + \Delta x \sup_{t \in [0,T]} \left(\|u\|_{L_x^\infty} \|u\|_{H_x^2} + \|\partial_x u\|_{L_x^\infty} \|u\|_{H_x^1} + \|u\|_{H_x^4} + \|u\|_{H_x^2} \right) \right], \end{aligned}$$

what can be simplified into

$$\begin{aligned} \|\epsilon^n\|_{\ell^\infty(0,N,\ell_\Delta^2)} &\leq C \left[\Delta t \sup_{t \in [0,T]} \left(\|u\|_{L_x^\infty}^2 \|u\|_{H_x^2} + \|u\|_{L_x^\infty} \|u\|_{H_x^4} + \|u\|_{H_x^6} \right) \right. \\ &\quad \left. + \Delta x \sup_{t \in [0,T]} \left(\|u\|_{L_x^\infty} \|u\|_{H_x^2} + \|\partial_x u\|_{L_x^\infty} \|u\|_{H_x^1} + \|u\|_{H_x^4} \right) \right]. \end{aligned}$$

Thus the consistency error is upper bounded by

$$\begin{aligned} \|\epsilon^n\|_{\ell^\infty(0,N,\ell_\Delta^2)} &\leq C \left\{ \Delta t \sup_{t \in [0,T]} \left[\left(1 + \|u\|_{L_x^\infty}^2 \right) \|u\|_{H_x^6} \right] \right. \\ &\quad \left. + \Delta x \sup_{t \in [0,T]} \left[\left(1 + \|u\|_{L_x^\infty} \right) \|u\|_{H_x^4} + \|\partial_x u\|_{L_x^\infty} \|u\|_{H_x^1} \right] \right\}. \end{aligned}$$

Remark A.2. The previous computation enables to determine an upper-bound of the consistency error for the Rusanov scheme

$$\epsilon_j^n = \frac{(u_\Delta)_j^{n+1} - (u_\Delta)_j^n}{\Delta t} + D_c \left(\frac{u_\Delta^2}{2} \right)_j^n - \frac{\tau \Delta x}{2} D_+ D_- (u_\Delta)_j^n,$$

with u_Δ obtained by averaging the solution of Burgers equation, and not KdV equation. We obtain

$$\begin{aligned} \|\epsilon^n\|_{\ell^\infty(0,N,\ell_\Delta^2)} &\leq C \left\{ \Delta t \sup_{t \in [0,T]} \left[\left(\|u\|_{L_x^\infty} + \|u\|_{L_x^\infty}^2 \right) \|u\|_{H_x^2} \right] \right. \\ &\quad \left. + \Delta x \sup_{t \in [0,T]} \left[\left(1 + \|u\|_{L_x^\infty} \right) \|u\|_{H_x^2} + \|\partial_x u\|_{L_x^\infty} \|u\|_{H_x^1} \right] \right\}. \end{aligned}$$

A.2 Consistency error for the *abcd* system (Chapter 4)

We consider the numerical scheme as general as possible, and deal with a θ discretization of the advection term. Once again, for simplicity, we note $t^{n+1} = \inf(t^{n+1}, T)$ and $\Delta t = t^{n+1} - t^n$.

A.2.1 Consistency error for the first equation of the *abcd* system : ϵ_1^n

By definition of the consistency error, on has, for all $(n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}$

$$\begin{aligned} (\epsilon_1)_j^n &= (I - bD_+ D_-) \left(\frac{(\eta_\Delta)_j^{n+1} - (\eta_\Delta)_j^n}{\Delta t} \right) + (I + aD_+ D_-) \left(\theta D_c (u_\Delta)_j^{n+1} + (1 - \theta) D_c (u_\Delta)_j^n \right) \\ &\quad + D_c (\eta_\Delta u_\Delta)_j^n - \frac{\tau_1}{2} \Delta x D_+ D_- (\eta_\Delta)_j^n. \end{aligned}$$

We define

$$\begin{cases} E_{\text{time}} = (I - bD_+D_-) \left(\frac{(\eta_\Delta)_j^{n+1} - (\eta_\Delta)_j^n}{\Delta t} \right), \\ E_{\text{linear}} = (I + aD_+D_-) \left(\theta D_c(u_\Delta)_j^{n+1} + (1 - \theta) D_c(u_\Delta)_j^n \right), \\ E_{\text{non linear}} = D_c(\eta_\Delta u_\Delta)_j^n, \\ E_{\text{viscosity}} = \frac{\tau_1}{2} \Delta x D_+D_- (\eta_\Delta)_j^n. \end{cases}$$

We define $Q = [x_j, x_{j+1}] \times [t^n, t^{n+1}]$. We will only develop the non linear part and enonce the results for the other parts. By Taylor expansions and Cauchy-Schwarz inequality, one has

$$|E_{\text{time}}| \leq C \left(\frac{1}{\Delta t \Delta x} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \partial_t \eta(s, y) - b \partial_x^2 \partial_t \eta(s, y) ds dy + \sqrt{\frac{\Delta t}{\Delta x}} \|\partial_t^2 \eta\|_{L^2(Q)} \right. \\ \left. + b \Delta x \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x^4 \partial_t \eta\|_{L^2(Q)} + b \sqrt{\frac{\Delta t}{\Delta x}} \|\partial_x^2 \partial_t^2 \eta\|_{L^2(Q)} \right).$$

In the same way, we develop the E_{linear} -term.

$$|E_{\text{linear}}| \leq C \left(\frac{1}{\Delta t \Delta x} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \partial_x u(s, y) + a \partial_x^3 u(s, y) ds dy + \theta \sqrt{\frac{\Delta t}{\Delta x}} \|\partial_x \partial_t u\|_{L^2(Q)} + \Delta x \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x^3 u\|_{L^2(Q)} \right. \\ \left. + |a| \Delta x \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x^5 u\|_{L^2(Q)} + |a| \theta \sqrt{\frac{\Delta t}{\Delta x}} \|\partial_t \partial_x^3 u\|_{L^2(Q)} + |a| \Delta x \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x^5 u\|_{L^2(Q)} \right).$$

We will develop the non linear part. As for the KdV equation (Section A.1), we denote K the function on $[0, 1]$

$$K(v) := \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \eta(s, y + v \Delta x) u(t, x + v \Delta x) dx dy ds dt.$$

The non linear part will be computed by a Taylor expansion of K :

$$2\Delta x^3 \Delta t^2 E_{\text{non linear}} = K(1) - K(-1) = 2K'(0) + \int_0^1 K''(w)(1 - w)dw + \int_0^1 K''(-w)(-1 + w)dw.$$

Thus, we have to compute the first derivative of K

$$\begin{aligned} K'(v) &= \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \Delta x \partial_x \eta(s, y + v \Delta x) u(t, x + v \Delta x) + \Delta x \eta(s, y + v \Delta x) \partial_x u(t, x + v \Delta x) dx dy ds dt \\ &= \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \Delta x \partial_x \eta(s, y + v \Delta x) \left[u(s, y + v \Delta x) + \int_y^x \partial_x u(s, z + v \Delta x) dz \right. \\ &\quad \left. + \int_s^t \partial_t u(\tau, x + v \Delta x) d\tau \right] + \Delta x \partial_x u(t, x + v \Delta x) \left[\eta(t, x + v \Delta x) + \int_x^y \partial_x \eta(t, z + v \Delta x) dz \right. \\ &\quad \left. + \int_t^s \partial_t \eta(\tau, y + v \Delta x) d\tau \right] dx dy ds dt \end{aligned}$$

The previous equality leads to

$$\begin{aligned}
 K'(v) = & \Delta x^2 \Delta t \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \partial_x(\eta u)(s, y + v\Delta x) ds dy \\
 & + \Delta x \Delta t \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_y^x \partial_x \eta(s, y + v\Delta x) \partial_x u(s, z + v\Delta x) dz dx dy ds \\
 & + \Delta x \Delta t \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_x^y \partial_x u(t, x + v\Delta x) \partial_x \eta(t, z + v\Delta x) dz dx dy ds \\
 & + \Delta x \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_s^t \partial_x \eta(s, y + v\Delta x) \partial_t u(\tau, x + v\Delta x) d\tau dx dy ds dt \\
 & + \Delta x \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_t^s \partial_x u(t, x + v\Delta x) \partial_t \eta(\tau, y + v\Delta x) d\tau dx dy ds dt
 \end{aligned}$$

Moreover, we have to compute the second derivative of K (exactly as in Section A.1 for the KdV equation).

$$\begin{aligned}
 K''(v) = & \Delta x^2 \Delta t \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} 2\Delta x (\partial_x u \partial_x \eta)(s, y + v\Delta x) + \Delta x (u \partial_x^2 \eta)(s, y + v\Delta x) + \Delta x (\eta \partial_x^2 u)(s, y + v\Delta x) ds dy \\
 & + \Delta x \Delta t \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_y^x \Delta x \partial_x^2 \eta(s, y + v\Delta x) \partial_x u(s, z + v\Delta x) + \Delta x \partial_x \eta(s, y + v\Delta x) \partial_x^2 u(s, z + v\Delta x) dz dx dy ds \\
 & + \Delta x \Delta t \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_x^y \Delta x \partial_x^2 u(t, s + v\Delta x) \partial_x \eta(t, z + v\Delta x) + \Delta x \partial_x u(t, x + v\Delta x) \partial_x^2 \eta(t, z + v\Delta x) dz dx dy ds \\
 & + \Delta x \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_s^t \Delta x \partial_x^2 \eta(s, y + v\Delta x) \partial_t u(\tau, x + v\Delta x) + \Delta x \partial_x \eta(s, y + v\Delta x) \partial_{tx} u(\tau, x + v\Delta x) d\tau dx dy ds dt \\
 & + \Delta x \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_t^s \Delta x \partial_x^2 u(t, x + v\Delta x) \partial_t \eta(\tau, y + v\Delta x) + \Delta x \partial_x u(t, x + v\Delta x) \partial_{tx} \eta(\tau, y + v\Delta x) d\tau dx dy ds dt.
 \end{aligned}$$

Applying once again the Cauchy-Schwarz inequality gives

$$\begin{aligned}
 |K''(v)| \leq & C \left(\Delta x^3 \Delta t \|\partial_x u\|_{L^2(Q)} \|\partial_x \eta\|_{L^2(Q)} + \Delta t \Delta x^3 \|u\|_{L^2(Q)} \|\partial_x^2 \eta\|_{L^2(Q)} + \Delta t \Delta x^3 \|\eta\|_{L^2(Q)} \|\partial_x^2 u\|_{L^2(Q)} \right. \\
 & + \Delta x^4 \Delta t \|\partial_x^2 \eta\|_{L^2(Q)} \|\partial_x u\|_{L^2(Q)} + \Delta x^4 \Delta t \|\partial_x^2 u\|_{L^2(Q)} \|\partial_x \eta\|_{L^2(Q)} + \Delta x^3 \Delta t^2 \|\partial_x^2 \eta\|_{L^2(Q)} \|\partial_t u\|_{L^2(Q)} \\
 & \left. + \Delta x^3 \Delta t^2 \|\partial_x \eta\|_{L^2(Q)} \|\partial_t \partial_x u\|_{L^2(Q)} + \Delta x^3 \Delta t^2 \|\partial_t \eta\|_{L^2(Q)} \|\partial_x^2 u\|_{L^2(Q)} + \Delta x^3 \Delta t^2 \|\partial_t \partial_x \eta\|_{L^2(Q)} \|\partial_x u\|_{L^2(Q)} \right).
 \end{aligned}$$

Thus, the $E_{\text{non linear}}$ -term rewrites

$$\begin{aligned}
 |E_{\text{non linear}}| \leq & C \left(\frac{1}{\Delta x \Delta t} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \partial_x(\eta u)(s, y) ds dy + \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x \eta\|_{L^\infty(Q)} \|\partial_x u\|_{L^2(Q)} \right. \\
 & + \sqrt{\frac{\Delta t}{\Delta x}} \|\partial_x \eta\|_{L^\infty(Q)} \|\partial_t u\|_{L^2(Q)} + \sqrt{\frac{\Delta t}{\Delta x}} \|\partial_t \eta\|_{L^2(Q)} \|\partial_x u\|_{L^\infty(Q)} + \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x u\|_{L^\infty(Q)} \|\partial_x \eta\|_{L^2(Q)} \\
 & + \sqrt{\frac{\Delta x}{\Delta t}} \|u\|_{L^\infty(Q)} \|\partial_x^2 \eta\|_{L^2(Q)} + \sqrt{\frac{\Delta x}{\Delta t}} \|\eta\|_{L^\infty(Q)} \|\partial_x^2 u\|_{L^2(Q)} + \Delta x \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x^2 \eta\|_{L^2(Q)} \|\partial_x u\|_{L^\infty(Q)} \\
 & + \Delta x \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x^2 u\|_{L^2(Q)} \|\partial_x \eta\|_{L^\infty(Q)} + \sqrt{\Delta t \Delta x} \|\partial_x^2 \eta\|_{L^\infty(Q)} \|\partial_t u\|_{L^2(Q)} + \sqrt{\Delta t \Delta x} \|\partial_x \eta\|_{L^\infty(Q)} \|\partial_t \partial_x u\|_{L^2(Q)} \\
 & \left. + \sqrt{\Delta t \Delta x} \|\partial_t \eta\|_{L^2(Q)} \|\partial_x^2 u\|_{L^\infty(Q)} + \sqrt{\Delta t \Delta x} \|\partial_t \partial_x \eta\|_{L^2(Q)} \|\partial_x u\|_{L^\infty(Q)} \right).
 \end{aligned}$$

Finally, one has

$$|E_{\text{viscosity}}| \lesssim \frac{\tau_1 \sqrt{\Delta x}}{2\sqrt{\Delta t}} \|\partial_x^2 \eta\|_{L^2(Q)}.$$

Then, when we sum up all the previous results, we obtain

$$\begin{aligned}
 (\epsilon_1)_j^n &\leq \frac{1}{\Delta t \Delta x} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \partial_t \eta(s, y) - b \partial_x^2 \partial_t \eta(s, y) + \partial_x u(s, y) + a \partial_x^3 u(s, y) + \partial_x(\eta u)(s, y) ds dy \\
 &\quad + \sqrt{\frac{\Delta t}{\Delta x}} \|\partial_t^2 \eta\|_{L^2(Q)} + b \Delta x \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x^4 \partial_t \eta\|_{L^2(Q)} + b \sqrt{\frac{\Delta t}{\Delta x}} \|\partial_x^2 \partial_t^2 \eta\|_{L^2(Q)} + \theta \sqrt{\frac{\Delta t}{\Delta x}} \|\partial_x \partial_t u\|_{L^2(Q)} \\
 &\quad + \Delta x \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x^3 u\|_{L^2(Q)} + |a| \Delta x \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x^5 u\|_{L^2(Q)} + |a| \theta \sqrt{\frac{\Delta t}{\Delta x}} \|\partial_t \partial_x^3 u\|_{L^2(Q)} + |a| \Delta x \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x^5 u\|_{L^2(Q)} \\
 &\quad + \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x \eta\|_{L^2(Q)} \|\partial_x u\|_{L^\infty(Q)} + \sqrt{\frac{\Delta t}{\Delta x}} \|\partial_x \eta\|_{L^\infty(Q)} \|\partial_t u\|_{L^2(Q)} + \sqrt{\frac{\Delta t}{\Delta x}} \|\partial_t \eta\|_{L^2(Q)} \|\partial_x u\|_{L^\infty(Q)} \\
 &\quad + \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x u\|_{L^\infty(Q)} \|\partial_x \eta\|_{L^2(Q)} + \sqrt{\frac{\Delta x}{\Delta t}} \|u\|_{L^\infty(Q)} \|\partial_x^2 \eta\|_{L^2(Q)} + \sqrt{\frac{\Delta x}{\Delta t}} \|\eta\|_{L^\infty(Q)} \|\partial_x^2 u\|_{L^2(Q)} \\
 &\quad + \Delta x \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x^2 \eta\|_{L^2(Q)} \|\partial_x u\|_{L^\infty(Q)} + \Delta x \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x^2 u\|_{L^2(Q)} \|\partial_x \eta\|_{L^\infty(Q)} + \sqrt{\Delta t \Delta x} \|\partial_x^2 \eta\|_{L^\infty(Q)} \|\partial_t u\|_{L^2(Q)} \\
 &\quad + \sqrt{\Delta t \Delta x} \|\partial_x \eta\|_{L^\infty(Q)} \|\partial_t \partial_x u\|_{L^2(Q)} + \sqrt{\Delta t \Delta x} \|\partial_t \eta\|_{L^2(Q)} \|\partial_x^2 u\|_{L^\infty(Q)} + \sqrt{\Delta t \Delta x} \|\partial_t \partial_x \eta\|_{L^2(Q)} \|\partial_x u\|_{L^\infty(Q)} \\
 &\quad + \frac{\tau_1 \sqrt{\Delta x}}{2 \sqrt{\Delta t}} \|\partial_x^2 \eta\|_{L^2(Q)}.
 \end{aligned}$$

We recognize the initial equation on the first line. We recall the relation

$$\sum_{j \in \mathbb{Z}} \Delta x \|f\|_{L^2(Q)}^2 = \Delta x \int_{t^n}^{t^{n+1}} \int_{\mathbb{R}} |f(s, y)|^2 dy ds \leq \Delta x \Delta t \sup_{t \in [0, T]} \|f(t, \cdot)\|_{L^2(\mathbb{R})}^2 = \Delta x \Delta t \|f(t, \cdot)\|_{L_t^\infty L_x^2}^2. \quad (\text{A.4})$$

Eventually, when we compute the ℓ_Δ^2 -norm, we obtain

$$\begin{aligned}
 \|\epsilon_1^n\|_{\ell_\Delta^2}^2 &\leq C \left(\Delta t^2 \|\partial_t^2 \eta\|_{L_t^\infty L_x^2}^2 + b^2 \Delta x^4 \|\partial_x^4 \partial_t \eta\|_{L_t^\infty L_x^2}^2 + b^2 \Delta t^2 \|\partial_x^2 \partial_t^2 \eta\|_{L_t^\infty L_x^2}^2 + \theta^2 \Delta t^2 \|\partial_x \partial_t u\|_{L_t^\infty L_x^2}^2 \right. \\
 &\quad + \Delta x^4 \|\partial_x^3 u\|_{L_t^\infty L_x^2}^2 + |a|^2 \Delta x^4 \|\partial_x^5 u\|_{L_t^\infty L_x^2}^2 + |a|^2 \theta^2 \Delta t^2 \|\partial_t \partial_x^3 u\|_{L_t^\infty L_x^2}^2 + |a|^2 \Delta x^4 \|\partial_x^5 u\|_{L_t^\infty L_x^2}^2 \\
 &\quad + \Delta x^2 \|\partial_x \eta\|_{L_t^\infty L_x^2}^2 \|\partial_x u\|_{L_t^\infty L_x^\infty}^2 + \Delta t^2 \|\partial_x \eta\|_{L_t^\infty L_x^\infty}^2 \|\partial_t u\|_{L_t^\infty L_x^2}^2 + \Delta t^2 \|\partial_t \eta\|_{L_t^\infty L_x^2}^2 \|\partial_x u\|_{L_t^\infty L_x^\infty}^2 \\
 &\quad + \Delta x^2 \|\partial_x u\|_{L_t^\infty L_x^\infty}^2 \|\partial_x \eta\|_{L_t^\infty L_x^2}^2 + \Delta x^2 \|u\|_{L_t^\infty L_x^\infty}^2 \|\partial_x^2 \eta\|_{L_t^\infty L_x^2}^2 + \Delta x^2 \|\eta\|_{L_t^\infty L_x^\infty}^2 \|\partial_x^2 u\|_{L_t^\infty L_x^2}^2 \\
 &\quad + \Delta x^4 \|\partial_x^2 \eta\|_{L_t^\infty L_x^2}^2 \|\partial_x u\|_{L_t^\infty L_x^\infty}^2 + \Delta x^4 \|\partial_x^2 u\|_{L_t^\infty L_x^2}^2 \|\partial_x \eta\|_{L_t^\infty L_x^\infty}^2 + \Delta t^2 \Delta x^2 \|\partial_x^2 \eta\|_{L_t^\infty L_x^\infty}^2 \|\partial_t u\|_{L_t^\infty L_x^2}^2 \\
 &\quad + \Delta t^2 \Delta x^2 \|\partial_x \eta\|_{L_t^\infty L_x^\infty}^2 \|\partial_t \partial_x u\|_{L_t^\infty L_x^2}^2 + \Delta t^2 \Delta x^2 \|\partial_t \eta\|_{L_t^\infty L_x^2}^2 \|\partial_x^2 u\|_{L_t^\infty L_x^\infty}^2 \\
 &\quad \left. + \Delta t^2 \Delta x^2 \|\partial_t \partial_x \eta\|_{L_t^\infty L_x^2}^2 \|\partial_x u\|_{L_t^\infty L_x^\infty}^2 + (\tau_1)^2 \Delta x^2 \|\partial_x^2 \eta\|_{L_t^\infty L_x^2}^2 \right).
 \end{aligned}$$

In some cases, $D_+ \epsilon_1^n$ is needed. To obtain an upper bound, we perform the same computations and find

$$\begin{aligned}
 \|D_+ \epsilon_1^n\|_{\ell_\Delta^2}^2 &\leq C \left(\Delta x^2 \|\partial_x^2 \partial_t \eta\|_{L_t^\infty L_x^2}^2 + \Delta t^2 \|\partial_x \partial_t^2 \eta\|_{L_t^\infty L_x^2}^2 + b^2 \Delta x^4 \|\partial_x^4 \partial_t \eta\|_{L_t^\infty L_x^2}^2 + b^2 \Delta t^2 \|\partial_x^3 \partial_t^2 \eta\|_{L_t^\infty L_x^2}^2 \right. \\
 &\quad + \Delta x^2 \|\partial_x^3 u\|_{L_t^\infty L_x^2}^2 + |a|^2 \Delta x^2 \|\partial_x^5 u\|_{L_t^\infty L_x^2}^2 + \theta^2 \Delta t^2 \|\partial_x^2 \partial_t u\|_{L_t^\infty L_x^2}^2 + |a|^2 \theta^2 \Delta t^2 \|\partial_t \partial_x^4 u\|_{L_t^\infty L_x^2}^2 \\
 &\quad + \Delta x^4 \|\partial_x^4 u\|_{L_t^\infty L_x^2}^2 + |a|^2 \Delta x^4 \|\partial_x^6 u\|_{L_t^\infty L_x^2}^2 + \Delta x^2 \|\partial_x^2 \eta\|_{L_t^\infty L_x^2}^2 \|\partial_x u\|_{L^\infty(Q)} \\
 &\quad + \Delta x^2 \|\partial_x \eta\|_{L_t^\infty L^\infty(Q)}^2 \|\partial_x^2 u\|_{L_t^\infty L_x^2}^2 + \Delta t^2 \|\partial_x^2 \eta\|_{L_t^\infty L_x^\infty}^2 \|\partial_t u\|_{L_t^\infty L_x^2}^2 + \Delta t^2 \|\partial_x \eta\|_{L_t^\infty L_x^\infty}^2 \|\partial_t u\|_{L_t^\infty L_x^2}^2 \\
 &\quad + \Delta t^2 \|\partial_x^2 u\|_{L_t^\infty L_x^\infty}^2 \|\partial_t \eta\|_{L_t^\infty L_x^2}^2 + \Delta t^2 \|\partial_x u\|_{L_t^\infty L_x^\infty}^2 \|\partial_t \eta\|_{L_t^\infty L_x^2}^2 + \Delta x^2 \|\partial_x^2 u\|_{L_t^\infty L_x^2}^2 \|\partial_x \eta\|_{L_t^\infty L_x^\infty}^2 \\
 &\quad + \Delta x^2 \|\partial_x u\|_{L_t^\infty L_x^\infty}^2 \|\partial_x^2 \eta\|_{L_t^\infty L_x^2}^2 + \Delta x^2 \|u\|_{L_t^\infty L_x^\infty}^2 \|\partial_x^3 \eta\|_{L_t^\infty L_x^2}^2 + \Delta x^2 \|\eta\|_{L_t^\infty L_x^\infty}^2 \|\partial_x^3 u\|_{L_t^\infty L_x^2}^2 \\
 &\quad + \Delta x^4 \|\partial_x^3 \eta\|_{L_t^\infty L_x^2}^2 \|\partial_x u\|_{L_t^\infty L_x^\infty}^2 + \Delta x^4 \|\partial_x^2 u\|_{L_t^\infty L_x^2}^2 \|\partial_x^2 \eta\|_{L_t^\infty L_x^\infty}^2 + \Delta x^4 \|\partial_x \eta\|_{L_t^\infty L_x^\infty}^2 \|\partial_x^3 u\|_{L_t^\infty L_x^2}^2 \\
 &\quad + \Delta t^2 \Delta x^2 \|\partial_x^3 \eta\|_{L_t^\infty L_x^\infty}^2 \|\partial_t u\|_{L_t^\infty L_x^2}^2 + \Delta t^2 \Delta x^2 \|\partial_x^2 \eta\|_{L_t^\infty L_x^\infty}^2 \|\partial_t \partial_x u\|_{L_t^\infty L_x^2}^2 \\
 &\quad \left. + \Delta t^2 \Delta x^2 \|\partial_x \eta\|_{L_t^\infty L_x^\infty}^2 \|\partial_t \partial_x^2 u\|_{L_t^\infty L_x^2}^2 + \Delta t^2 \Delta x^2 \|\partial_t \eta\|_{L_t^\infty L_x^2}^2 \|\partial_x^3 u\|_{L_t^\infty L_x^\infty}^2 \right)
 \end{aligned}$$

$$+ \Delta t^2 \Delta x^2 \|\partial_t \partial_x \eta\|_{L_t^\infty L_x^2}^2 \|\partial_x^2 u\|_{L_t^\infty L_x^\infty}^2 + \Delta t^2 \Delta x^2 \|\partial_x u\|_{L_t^\infty L_x^\infty}^2 \|\partial_t \partial_x^2 \eta\|_{L_t^\infty L_x^2}^2 \\ + \tau_1^2 \Delta x^2 \|\partial_x^3 \eta\|_{L_t^\infty L_x^2}^2 \Big).$$

A.2.2 Consistency error for the second equation of the $abcd$ system : ϵ_2^n

By definition of the consistency error, one has for the second equation

$$(\epsilon_2)_j^n = (I - dD_+ D_-) \left(\frac{(u_\Delta)_j^{n+1} - (u_\Delta)_j^n}{\Delta t} \right) + (I + cD_+ D_-) \left(\theta D_c (\eta_\Delta)_j^{n+1} + (1 - \theta) D_c (\eta_\Delta)_j^n \right) \\ + D_c \left(\frac{u_\Delta^2}{2} \right)_j^n - \frac{\tau_2}{2} \Delta x D_+ D_- u_\Delta^n, \quad (n, j) \in \llbracket 0, N \rrbracket \times \mathbb{Z}.$$

We adapt the previous computations with $(d, c, \eta_\Delta^n, u_\Delta^n)$ instead of $(b, a, u_\Delta^n, \eta_\Delta^n)$. The only difference is concerning the non linear term

$$K_2(v) := \frac{1}{2} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} u(s, y + v \Delta x) u(t, x + v \Delta x) dx dy ds dt.$$

So one has

$$E_{\text{non linear}} \leq C \left(\frac{1}{\Delta x \Delta t} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \partial_x \frac{u^2}{2}(s, y) ds dy + \frac{\sqrt{\Delta x}}{\sqrt{\Delta t}} \|\partial_x u\|_{L^\infty(Q)} \|\partial_x u\|_{L^2(Q)} \right. \\ \left. + \frac{\sqrt{\Delta t}}{\sqrt{\Delta x}} \|\partial_x u\|_{L^\infty(Q)} \|\partial_t u\|_{L^2(Q)} + \frac{\sqrt{\Delta x}}{2\sqrt{\Delta t}} \|u\|_{L^\infty(Q)} \|\partial_x^2 u\|_{L^2(Q)} + \Delta x \frac{\sqrt{\Delta x}}{2\sqrt{\Delta t}} \|\partial_x^2 u\|_{L^2(Q)} \|\partial_x u\|_{L^\infty(Q)} \right. \\ \left. + \sqrt{\Delta t \Delta x} \|\partial_x^2 u\|_{L^\infty(Q)} \|\partial_t u\|_{L^2(Q)} + \sqrt{\Delta t \Delta x} \|\partial_x u\|_{L^\infty(Q)} \|\partial_t \partial_x u\|_{L^2(Q)} \right).$$

The consistency error verifies

$$(\epsilon_2)_j^n \leq C \left(\frac{1}{\Delta t \Delta x} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \partial_t u(s, y) - d \partial_x^2 \partial_t u(s, y) + \partial_x \eta(s, y) + c \partial_x^3 \eta(s, y) + \partial_x \frac{u^2}{2}(s, y) ds dy \right. \\ \left. + \sqrt{\frac{\Delta t}{\Delta x}} \|\partial_t^2 u\|_{L^2(Q)} + d \Delta x \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x^4 \partial_t u\|_{L^2(Q)} + d \sqrt{\frac{\Delta t}{\Delta x}} \|\partial_x^2 \partial_t^2 u\|_{L^2(Q)} + \theta \sqrt{\frac{\Delta t}{\Delta x}} \|\partial_x \partial_t \eta\|_{L^2(Q)} \right. \\ \left. + \Delta x \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x^3 \eta\|_{L^2(Q)} + |c| \Delta x \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x^5 \eta\|_{L^2(Q)} + |c| \theta \sqrt{\frac{\Delta t}{\Delta x}} \|\partial_t \partial_x^3 \eta\|_{L^2(Q)} + |c| \Delta x \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x^5 \eta\|_{L^2(Q)} \right. \\ \left. + \sqrt{\frac{\Delta x}{\Delta t}} \|\partial_x u\|_{L^2(Q)} \|\partial_x u\|_{L^\infty(Q)} + \sqrt{\frac{\Delta t}{\Delta x}} \|\partial_x u\|_{L^\infty(Q)} \|\partial_t u\|_{L^2(Q)} + \frac{\sqrt{\Delta x}}{2\sqrt{\Delta t}} \|u\|_{L^\infty(Q)} \|\partial_x^2 u\|_{L^2(Q)} \right. \\ \left. + \Delta x \frac{\sqrt{\Delta x}}{2\sqrt{\Delta t}} \|\partial_x^2 u\|_{L^2(Q)} \|\partial_x u\|_{L^\infty(Q)} + \sqrt{\Delta t \Delta x} \|\partial_x^2 u\|_{L^\infty(Q)} \|\partial_t u\|_{L^2(Q)} + \sqrt{\Delta t \Delta x} \|\partial_x u\|_{L^\infty(Q)} \|\partial_t \partial_x u\|_{L^2(Q)} \right. \\ \left. + \frac{\tau_2 \sqrt{\Delta x}}{2\sqrt{\Delta t}} \|\partial_x^2 u\|_{L^2(Q)} \right).$$

Finally, one has, thanks to the relation (A.4)

$$\|\epsilon_2^n\|_{\ell_\Delta^2}^2 \leq C \left(\Delta t^2 \|\partial_t^2 u\|_{L_t^\infty L_x^2}^2 + d^2 \Delta x^4 \|\partial_x^4 \partial_t u\|_{L_t^\infty L_x^2}^2 + d^2 \Delta t^2 \|\partial_x^2 \partial_t^2 u\|_{L_t^\infty L_x^2}^2 + \theta^2 \Delta t^2 \|\partial_x \partial_t \eta\|_{L_t^\infty L_x^2}^2 \right. \\ \left. + \Delta x^4 \|\partial_x^3 \eta\|_{L_t^\infty L_x^2}^2 + |c|^2 \Delta x^4 \|\partial_x^5 \eta\|_{L_t^\infty L_x^2}^2 + |c|^2 \theta^2 \Delta t^2 \|\partial_t \partial_x^3 \eta\|_{L_t^\infty L_x^2}^2 + |c|^2 \Delta x^4 \|\partial_x^5 \eta\|_{L_t^\infty L_x^2}^2 \right. \\ \left. + \Delta x^2 \|\partial_x u\|_{L_t^\infty L_x^2}^2 \|\partial_x u\|_{L_t^\infty L_x^\infty}^2 + \Delta t^2 \|\partial_x u\|_{L_t^\infty L_x^\infty}^2 \|\partial_t u\|_{L_t^\infty L_x^2}^2 + \Delta x^2 \|u\|_{L_t^\infty L_x^\infty}^2 \|\partial_x^2 u\|_{L_t^\infty L_x^2}^2 \right. \\ \left. + \Delta x^4 \|\partial_x^2 u\|_{L_t^\infty L_x^2}^2 \|\partial_x u\|_{L_t^\infty L_x^\infty}^2 + \Delta t^2 \Delta x^2 \|\partial_x^2 u\|_{L_t^\infty L_x^\infty}^2 \|\partial_t u\|_{L_t^\infty L_x^2}^2 \right. \\ \left. + \Delta t^2 \Delta x^2 \|\partial_x u\|_{L_t^\infty L_x^\infty}^2 \|\partial_t \partial_x u\|_{L_t^\infty L_x^2}^2 + \tau_2^2 \Delta x^2 \|\partial_x^2 u\|_{L_t^\infty L_x^2}^2 \right).$$

For $D_+\epsilon_2^n$, the results are similar to those for $D_+\epsilon_1^n$.

Remark A.3. *If we use a Taylor expansion a step further for $E_{\text{non linear}}$ we recover*

$$\|\epsilon_1^n\|_{\ell_\Delta^2}^2 \leq \mathcal{O}(\Delta t^2 + \Delta x^4) \quad \text{and} \quad \|\epsilon_2^n\|_{\ell_\Delta^2}^2 \leq \mathcal{O}(\Delta t^2 + \Delta x^4).$$

ANNEXE B

Proofs of technical Lemmas dealing with the operators D_+ , D_- , D_c



We gather in this appendix all the proofs of the useful lemmas relating to the discrete operators D_+ , D_- and D_c , introduced first in Equation (1.12) in general introduction and the cornerstone of Section 3.4.1 of Chapter 3.

B.1 Proofs of useful Lemmas of Section 3.4.1

Proof of Lemma 3.19. Identities (3.18) to (3.22b) are proved by developing the expressions. □

Proof of Lemma 3.20. • Relation (3.23) is obtained by developing the sums.

- Relation (3.24) is proved by summing the square of Identity (3.20) with $b_j = \frac{a_j}{2}$.
- Relation (3.25) is obtained by developing the expression :

$$\begin{aligned}
 \sum_{j \in \mathbb{Z}} \left[D_+ D_- (a)_j \right]^2 &= \sum_{j \in \mathbb{Z}} \left[\frac{a_{j+1} - 2a_j + a_{j-1}}{\Delta x^2} \right]^2 \\
 &= \sum_{j \in \mathbb{Z}} \frac{a_{j+1}^2 - 4a_{j+1}a_j + 2a_{j+1}a_{j-1} + 4a_j^2 - 4a_ja_{j-1} + a_{j-1}^2}{\Delta x^4} \\
 &= \sum_{j \in \mathbb{Z}} 4 \frac{a_{j+1}^2 - 2a_{j+1}a_j + a_j^2}{\Delta x^4} - \frac{a_{j+1}^2 - 2a_{j-1}a_{j+1} + a_{j-1}^2}{\Delta x^4} \\
 &= \sum_{j \in \mathbb{Z}} \frac{4}{\Delta x^2} \frac{a_{j+1}^2 - 2a_{j+1}a_j + a_j^2}{\Delta x^2} - \frac{4}{\Delta x^2} \frac{a_{j+1}^2 - 2a_{j-1}a_{j+1} + a_{j-1}^2}{4\Delta x^2}.
 \end{aligned}$$

□

Proof of Lemma 3.22. • Identities (3.27) and (3.28) derive from developing the sums.

- Relation (3.29) is a straightforward consequence of (3.22a).

- For (3.30), one has

$$\sum_{j \in \mathbb{Z}} D_+(a)_j a_j a_{j+1} = \sum_{j \in \mathbb{Z}} \frac{a_{j+1}^2 a_j - a_j^2 a_{j+1}}{\Delta x} = \sum_{j \in \mathbb{Z}} -\frac{a_{j+1}^3 - 3a_{j+1}^2 a_j + 3a_{j+1} a_j^2 - a_j^3}{3\Delta x} = -\sum_{j \in \mathbb{Z}} \frac{\Delta x^2}{3} \left[\frac{a_{j+1} - a_j}{\Delta x} \right]^3.$$

- Likewise, Identity (3.31) is proved by the same way

$$\begin{aligned} \sum_{j \in \mathbb{Z}} D_c(a)_j a_{j-1} a_{j+1} &= \sum_{j \in \mathbb{Z}} \frac{a_{j+1}^2 a_{j-1} - a_{j-1}^2 a_{j+1}}{2\Delta x} = \sum_{j \in \mathbb{Z}} -\frac{4}{3} \frac{a_{j+1}^3 - 3a_{j+1}^2 a_{j-1} + 3a_{j+1} a_{j-1}^2 - a_{j-1}^3}{8\Delta x} \\ &= \sum_{j \in \mathbb{Z}} -\frac{4\Delta x^2}{3} \left[\frac{a_{j+1} - a_{j-1}}{2\Delta x} \right]^3. \end{aligned}$$

- The proof of (3.32) is based on a development of the expression of $D_c(ab)_j$ as explained below

$$\sum_{j \in \mathbb{Z}} a_j D_c(ab)_j = \sum_{j \in \mathbb{Z}} a_j \frac{a_{j+1} b_{j+1} - a_{j-1} b_{j-1}}{2\Delta x} = \sum_{j \in \mathbb{Z}} \frac{b_{j+1} a_{j+1} a_j - b_j a_{j+1} a_j}{2\Delta x} = \sum_{j \in \mathbb{Z}} \frac{b_{j+1} - b_j}{2\Delta x} a_{j+1} a_j.$$

- For Relation (3.33), we expand

$$\begin{aligned} &\sum_{j \in \mathbb{Z}} D_+ D_- (a)_j D_c(ab)_j \\ &= \sum_{j \in \mathbb{Z}} \frac{(a_{j+1} - 2a_j + a_{j-1}) [a_{j+1} b_{j+1}] - (a_{j+1} - 2a_j + a_{j-1}) [a_{j-1} b_{j-1}]}{2\Delta x^3} \\ &= \sum_{j \in \mathbb{Z}} \frac{(a_{j+1})^2 b_{j+1} - 2a_j a_{j+1} b_{j+1} + a_{j-1} a_{j+1} b_{j+1} - a_{j+1} a_{j-1} b_{j-1} + 2a_j a_{j-1} b_{j-1} - (a_{j-1})^2 b_{j-1}}{2\Delta x^3} \\ &= \sum_{j \in \mathbb{Z}} \frac{-2a_j a_{j+1} b_{j+1} + 2a_{j+1} a_j b_j}{2\Delta x^3} + \frac{a_{j-1} a_{j+1} b_{j+1} - a_{j+1} a_{j-1} b_{j-1}}{2\Delta x^3} \\ &= \sum_{j \in \mathbb{Z}} -\frac{b_{j+1} - b_j}{\Delta x^3} a_{j+1} a_j + \sum_{j \in \mathbb{Z}} \frac{b_{j+1} - b_{j-1}}{2\Delta x^3} a_{j+1} a_{j-1}. \end{aligned}$$

□

Proof of Corollary 3.23. • Identity (3.34) is based on (3.32) with $b_j = \frac{a_j}{2}$ and Relation (3.30).

- Proving (3.35) corresponds to choose $b_j = \frac{a_j}{2}$ in Identity (3.33) and using Relations (3.30) and (3.31).

□

Proof of Lemma 3.27. Inequality (3.40) is based on Relation (3.21)

$$\begin{aligned} \sum_{j \in \mathbb{Z}} \left[D_c(ab)_j \right]^2 &= \sum_{j \in \mathbb{Z}} \left[b_j D_c(a)_j + \frac{a_{j+1}}{2} D_+(b)_j + \frac{a_{j-1}}{2} D_-(b)_j \right]^2 \\ &= \sum_{j \in \mathbb{Z}} (b_j)^2 (D_c(a)_j)^2 + \sum_{j \in \mathbb{Z}} b_j a_{j+1} D_c(a)_j D_+(b)_j + \sum_{j \in \mathbb{Z}} b_j D_c(a)_j a_{j-1} D_-(b)_j \\ &\quad + \sum_{j \in \mathbb{Z}} \left(\frac{a_{j+1}}{2} D_+(b)_j \right)^2 + \sum_{j \in \mathbb{Z}} \frac{a_{j+1} a_{j-1}}{2} D_+(b)_j D_-(b)_j + \sum_{j \in \mathbb{Z}} \left(\frac{a_{j-1}}{2} D_-(b)_j \right)^2. \end{aligned}$$

Young inequality implies

$$\begin{aligned} \sum_{j \in \mathbb{Z}} \left(D_c(ab)_j \right)^2 &\leq \sum_{j \in \mathbb{Z}} (b_j)^2 \left(D_c(a)_j \right)^2 + \sum_{j \in \mathbb{Z}} \frac{(b_j)^2}{2\Delta t} (a_{j+1})^2 + \sum_{j \in \mathbb{Z}} \frac{\Delta t}{2} \left(D_c(a)_j \right)^2 \left(D_+(b)_j \right)^2 \\ &+ \sum_{j \in \mathbb{Z}} \frac{(b_j)^2}{2\Delta t} (a_{j-1})^2 + \sum_{j \in \mathbb{Z}} \frac{\Delta t}{2} \left(D_c(a)_j \right)^2 \left(D_-(b)_j \right)^2 + \sum_{j \in \mathbb{Z}} \frac{3}{2} \left(\frac{a_{j+1}}{2} D_+(b)_j \right)^2 + \sum_{j \in \mathbb{Z}} \frac{3}{2} \left(\frac{a_{j-1}}{2} D_-(b)_j \right)^2. \end{aligned}$$

□

Proof of Lemma 3.28. We use relation (3.21) to develop $D_+ D_+ D_- (a)_j D_c(ab)_j$ which gives

$$\begin{aligned} \sum_{j \in \mathbb{Z}} D_+ D_+ D_- (a)_j D_c(ab)_j &= \sum_{j \in \mathbb{Z}} D_+ D_+ D_- (a)_j \left[b_j D_c(a)_j + \frac{a_{j+1}}{2} D_+(b)_j + \frac{a_{j-1}}{2} D_-(b)_j \right] \\ &\leq \sum_{j \in \mathbb{Z}} b_j D_+ D_+ D_- (a)_j D_c(a)_j + \sum_{j \in \mathbb{Z}} \frac{\Delta t}{4} \left(D_+ D_+ D_- (a)_j \right)^2 |D_+(b)_j| \\ &+ \sum_{j \in \mathbb{Z}} \frac{(a_{j+1})^2}{4\Delta t} |D_+(b)_j| + \sum_{j \in \mathbb{Z}} \frac{\Delta t}{4} \left(D_+ D_+ D_- (a)_j \right)^2 |D_-(b)_j| + \sum_{j \in \mathbb{Z}} \frac{(a_{j-1})^2}{4\Delta t} |D_-(b)_j|. \end{aligned} \quad (\text{B.1})$$

The conclusion comes from the following Lemma.

Lemma B.1. Let $(a_j)_{j \in \mathbb{Z}}$ and $(b_j)_{j \in \mathbb{Z}}$ be two sequences in $\ell_\Delta^2(\mathbb{Z})$, σ be in $\{0, 1\}$ and ν be non negative. Then, it holds

$$\begin{aligned} \sum_{j \in \mathbb{Z}} b_j D_+ D_+ D_- (a)_j D_c(a)_j &\leq \sum_{j \in \mathbb{Z}} \frac{1}{2} \left\{ \Delta x^\nu \left(\frac{|D_-(b)_j|^\sigma}{2} + \frac{|D_-(b)_j|^\sigma}{2} \right) - \frac{\Delta x}{2} D_- b_j \right\} \left(D_+ D_- (a)_j \right)^2 \\ &+ \sum_{j \in \mathbb{Z}} \frac{1}{2\Delta x^\nu} |D_+(b)_j|^{2-\sigma} \left(D_+(a)_j \right)^2 - \sum_{j \in \mathbb{Z}} b_j \left(D_+ D_c(a)_j \right)^2. \end{aligned} \quad (\text{B.2})$$

Proof of Lemma B.1. By developing $D_c(a)_j$ and using Relation (3.27), it holds

$$\begin{aligned} \sum_{j \in \mathbb{Z}} b_j D_+ D_+ D_- (a)_j D_c(a)_j &= \sum_{j \in \mathbb{Z}} \frac{b_j}{2} D_+ D_+ D_- (a)_j D_+(a)_j + \sum_{j \in \mathbb{Z}} \frac{b_j}{2} D_+ D_+ D_- (a)_j D_-(a)_j \\ &= - \sum_{j \in \mathbb{Z}} D_+ D_- (a)_j D_- \left(\frac{b_j}{2} D_+(a)_j \right) - \sum_{j \in \mathbb{Z}} D_+ D_- (a)_j D_- \left(\frac{b_j}{2} D_-(a)_j \right). \end{aligned}$$

We focus first on the term $-\sum_{j \in \mathbb{Z}} D_+ D_- (a)_j D_- \left(\frac{b_j}{2} D_+(a)_j \right)$. Equality (3.19b) gives

$$- \sum_{j \in \mathbb{Z}} D_+ D_- (a)_j D_- \left(\frac{b_j}{2} D_+(a)_j \right) = - \sum_{j \in \mathbb{Z}} D_+ D_- (a)_j \left(\frac{D_-(b)_j}{2} D_-(a)_j + \frac{b_j}{2} D_+ D_- (a)_j \right).$$

Eventually, Young inequality provides

$$\begin{aligned} - \sum_{j \in \mathbb{Z}} D_+ D_- (a)_j D_- \left(\frac{b_j}{2} D_+(a)_j \right) &\leq \sum_{j \in \mathbb{Z}} \frac{\Delta x^\nu |D_- b_j|^\sigma}{4} \left(D_+ D_- (a)_j \right)^2 + \sum_{j \in \mathbb{Z}} \frac{|D_+(b)_j|^{2-\sigma}}{4\Delta x^\nu} \left(D_+(a)_j \right)^2 \\ &- \sum_{j \in \mathbb{Z}} \frac{b_j}{2} \left(D_+ D_- (a)_j \right)^2. \end{aligned}$$

Annexe B.1. Proofs of useful Lemmas of Section 3.4.1

For the term $-\sum_{j \in \mathbb{Z}} D_+ D_- (a)_j D_- \left(\frac{b_j}{2} D_- (a)_j \right)$, one has thanks to Equality (3.19b),

$$-\sum_{j \in \mathbb{Z}} D_+ D_- (a)_j D_- \left(\frac{b_j}{2} D_- (a)_j \right) = -\sum_{j \in \mathbb{Z}} D_+ D_- (a)_j \left(\frac{D_- (b)_j}{2} D_- (a)_j + \frac{b_{j-1}}{2} D_- D_- (a)_j \right).$$

Hence, it holds (by Young inequality)

$$\begin{aligned} & -\sum_{j \in \mathbb{Z}} D_+ D_- (a)_j D_- \left(\frac{b_j}{2} D_- (a)_j \right) \\ & \leq \sum_{j \in \mathbb{Z}} \frac{\Delta x^\nu |D_- b_j|^\sigma}{4} (D_+ D_- (a)_j)^2 + \sum_{j \in \mathbb{Z}} \frac{|D_+ (b)_j|^{2-\sigma}}{4\Delta x^\nu} (D_+ a_j)^2 - \sum_{j \in \mathbb{Z}} \frac{b_{j-1}}{2} D_+ D_- (a)_j D_- D_- (a)_j \\ & \leq \sum_{j \in \mathbb{Z}} \frac{\Delta x^\nu |D_- b_j|^\sigma}{4} (D_+ D_- a_j)^2 + \sum_{j \in \mathbb{Z}} \frac{|D_+ (b)_j|^{2-\sigma}}{4\Delta x^\nu} (D_+ a_j)^2 \\ & - \sum_{j \in \mathbb{Z}} b_{j-1} \left(\frac{D_+ D_- a_j + D_- D_- a_j}{2} \right)^2 + \sum_{j \in \mathbb{Z}} \frac{b_{j-1}}{4} (D_+ D_- a_j)^2 + \sum_{j \in \mathbb{Z}} \frac{b_{j-1}}{4} (D_- D_- a_j)^2 \\ & \leq \sum_{j \in \mathbb{Z}} \left(\frac{\Delta x^\nu |D_- b_j|^\sigma}{4} + \frac{b_{j-1} + b_j}{4} \right) (D_+ D_- a_j)^2 - \sum_{j \in \mathbb{Z}} b_j (D_+ D_- a_j)^2 + \sum_{j \in \mathbb{Z}} \frac{|D_+ (b)_j|^{2-\sigma}}{4\Delta x^\nu} (D_+ (a)_j)^2. \end{aligned}$$

By collecting the previous results, one has

$$\begin{aligned} \sum_{j \in \mathbb{Z}} b_j D_+ D_+ D_- (a)_j D_- (a)_j & \leq \sum_{j \in \mathbb{Z}} \left\{ \frac{\Delta x^\nu |D_- b_j|^\sigma}{4} + \frac{\Delta x^\nu |D_- b_j|^\sigma}{4} + \frac{b_{j-1} - b_j}{4} \right\} (D_+ D_- a_j)^2 \\ & + \sum_{j \in \mathbb{Z}} \frac{|D_+ (b)_j|^{2-\sigma}}{2\Delta x^\nu} (D_+ (a)_j)^2 - \sum_{j \in \mathbb{Z}} b_j (D_+ D_- a_j)^2. \end{aligned}$$

Lemma B.1 is then proved. $\square$

Lemma 3.28 is a consequence, with $\nu = 0$. $\square$

Proof of Lemma 3.29. To prove Lemma 3.29, the following lemma is useful.

Lemma B.2. Let $(a_j)_{j \in \mathbb{Z}}$ be a sequence in $\ell_\Delta^2(\mathbb{Z})$, then one has

$$\|D_+ a\|_{\ell_\Delta^4} \leq \sqrt{3} \|a\|_{\ell^\infty} \|D_+ D_- a\|_{\ell_\Delta^2}.$$

Proof of Lemma B.2. We compute the following sum thanks to the discret integration by parts (3.27)

$$\sum_{j \in \mathbb{Z}} (D_+ a_j)^4 = \sum_{j \in \mathbb{Z}} (D_+ a_j)^3 D_+ a_j = - \sum_{j \in \mathbb{Z}} D_- \left[(D_+ a_j)^3 \right] a_j.$$

We use Relation (3.19b) to develop $D_- \left[(D_+ a_j)^3 \right]$ (line 1, hereafter) and $D_- \left[(D_+ (a)_j)^2 \right]$ (line 2, hereafter)

$$\begin{aligned} - \sum_{j \in \mathbb{Z}} D_- \left[(D_+ a_j)^3 \right] a_j & = - \sum_{j \in \mathbb{Z}} D_- \left[(D_+ (a)_j)^2 \right] D_- (a)_j a_j - \sum_{j \in \mathbb{Z}} (D_+ a_j)^2 D_- D_+ (a)_j a_j \\ & = - \sum_{j \in \mathbb{Z}} D_+ D_- (a)_j (D_- a_j)^2 a_j - \sum_{j \in \mathbb{Z}} D_+ (a)_j D_+ D_- (a)_j D_- (a)_j a_j - \sum_{j \in \mathbb{Z}} (D_+ a_j)^2 D_+ D_- (a)_j a_j. \end{aligned}$$

We recover $D_c(a)_j$ in the second sum, which gives

$$\begin{aligned} \sum_{j \in \mathbb{Z}} (D_+ a_j)^4 &= - \sum_{j \in \mathbb{Z}} D_+ D_- a_j (D_- a_j)^2 a_j - \sum_{j \in \mathbb{Z}} 2 D_+ D_- (a)_j a_j (D_c(a)_j)^2 + \sum_{j \in \mathbb{Z}} D_+ D_- (a)_j a_j \frac{(D_+ a_j)^2}{2} \\ &\quad + \sum_{j \in \mathbb{Z}} D_+ D_- (a)_j a_j \frac{(D_- a_j)^2}{2} - \sum_{j \in \mathbb{Z}} (D_+ a_j)^2 D_+ D_- (a)_j a_j \\ &= - \sum_{j \in \mathbb{Z}} \frac{D_+ D_- (a)_j a_j (D_- a_j)^2}{2} - 2 \sum_{j \in \mathbb{Z}} D_+ D_- (a)_j a_j (D_c a_j)^2 - \sum_{j \in \mathbb{Z}} D_+ D_- (a)_j \frac{a_j (D_+ a_j)^2}{2}. \end{aligned}$$

Multiplying by Δx and applying Cauchy-Schwarz inequality simplify the ℓ_Δ^4 -norm

$$\begin{aligned} \|D_+ a\|_{\ell_\Delta^4}^4 &\leq \frac{\|a\|_{\ell^\infty}}{2} \|D_+ D_- (a)\|_{\ell_\Delta^2} \|D_- a\|_{\ell_\Delta^4}^2 + 2\|a\|_{\ell^\infty} \|D_+ D_- (a)\|_{\ell_\Delta^2} \|D_c a\|_{\ell_\Delta^4}^2 \\ &\quad + \frac{\|a\|_{\ell^\infty}}{2} \|D_+ D_- (a)\|_{\ell_\Delta^2} \|D_+ a\|_{\ell_\Delta^4}^2. \end{aligned} \quad (\text{B.3})$$

However, one has

$$\begin{aligned} \|D_c a\|_{\ell_\Delta^4}^4 &= \sum_{j \in \mathbb{Z}} \Delta x \left(\frac{a_{j+1} - a_j + a_j - a_{j-1}}{2\Delta x} \right)^4 \\ &= \sum_{j \in \mathbb{Z}} \Delta x \left(\frac{a_{j+1} - a_j + a_j - a_{j-1}}{2\Delta x} \right)^2 \left(\frac{a_{j+1} - a_j + a_j - a_{j-1}}{2\Delta x} \right)^2 \\ &= \sum_{j \in \mathbb{Z}} \Delta x \left[2 \left(\frac{a_{j+1} - a_j}{2\Delta x} \right)^2 + 2 \left(\frac{a_j - a_{j-1}}{2\Delta x} \right)^2 \right] \left[2 \left(\frac{a_{j+1} - a_j}{2\Delta x} \right)^2 + 2 \left(\frac{a_j - a_{j-1}}{2\Delta x} \right)^2 \right] \\ &\leq \sum_{j \in \mathbb{Z}} \Delta x \frac{1}{4} \left(\frac{a_{j+1} - a_j}{\Delta x} \right)^4 + \frac{1}{2} \Delta x \left(\frac{a_{j+1} - a_j}{\Delta x} \right)^2 \left(\frac{a_j - a_{j-1}}{\Delta x} \right)^2 + \frac{1}{4} \Delta x \left(\frac{a_j - a_{j-1}}{\Delta x} \right)^4 \\ &\leq \sum_{j \in \mathbb{Z}} \Delta x \left(\frac{a_{j+1} - a_j}{\Delta x} \right)^4 \\ &\leq \|D_+ a\|_{\ell_\Delta^4}^4. \end{aligned}$$

Equation (B.3) is changed as follows

$$\|D_+ a\|_{\ell_\Delta^4}^4 \leq 3\|a\|_{\ell^\infty} \|D_+ D_- (a)\|_{\ell_\Delta^2} \|D_+ a\|_{\ell_\Delta^4}^2.$$

Then it holds

$$\|D_+ a\|_{\ell_\Delta^4}^2 \leq 3\|a\|_{\ell^\infty} \|D_+ D_- (a)\|_{\ell_\Delta^2}.$$

Lemma B.2 is thus proved. $\square$

Remark B.3. Lemma B.2 is also valid for the classical discret ℓ^4 -norm instead of ℓ_Δ^4 -norm.

To prove Lemma 3.29, we first develop the left hand side thanks to (3.21)

$$\sum_{j \in \mathbb{Z}} D_+ D_+ D_- (a)_j D_c \left(\frac{a^2}{2} \right)_j = \sum_{j \in \mathbb{Z}} D_+ D_+ D_- (a)_j \left[\frac{a_j}{2} D_c(a)_j + \frac{a_{j+1}}{4} D_+(a)_j + \frac{a_{j-1}}{4} D_-(a)_j \right].$$

Annexe B.1. Proofs of useful Lemmas of Section 3.4.1

- The first term $\sum_{j \in \mathbb{Z}} \frac{a_j}{2} D_+ D_+ D_- (a)_j D_c (a)_j$ is treated with Lemma B.1 of Appendix B with $\nu = \frac{1}{2} - \gamma$ and $\sigma = 0$, which rewrites

$$\sum_{j \in \mathbb{Z}} \frac{a_j}{2} D_+ D_+ D_- (a)_j D_c (a)_j \leq \sum_{j \in \mathbb{Z}} \frac{1}{4} \left\{ \Delta x^{\frac{1}{2}-\gamma} - \frac{\Delta x}{2} D_- a_j \right\} (D_+ D_- a_j)^2 + \sum_{j \in \mathbb{Z}} \frac{1}{4 \Delta x^{\frac{1}{2}-\gamma}} (D_+ a_j)^4 - \sum_{j \in \mathbb{Z}} \frac{a_j}{2} (D_+ D_c a_j)^2.$$

- For the second term, we integrate by parts thanks to (3.27) and (3.19b)

$$\begin{aligned} \sum_{j \in \mathbb{Z}} D_+ D_+ D_- (a)_j \frac{a_{j+1}}{4} D_+ (a)_j &= \sum_{j \in \mathbb{Z}} -D_+ D_- (a)_j D_- \left(\frac{a_{j+1}}{4} D_+ (a)_j \right) \\ &= \sum_{j \in \mathbb{Z}} -D_+ D_- (a)_j \left\{ \frac{a_j}{4} D_+ D_- (a)_j + \frac{(D_+ (a)_j)^2}{4} \right\}. \end{aligned}$$

Young inequality completes the upper bound

$$\begin{aligned} \sum_{j \in \mathbb{Z}} D_+ D_+ D_- (a)_j \frac{a_{j+1}}{4} D_+ (a)_j &\leq \sum_{j \in \mathbb{Z}} - (D_+ D_- (a)_j)^2 \frac{a_j}{4} + \sum_{j \in \mathbb{Z}} \frac{1}{8} (D_+ D_- (a)_j)^2 \Delta x^{\frac{1}{2}-\gamma} \\ &\quad + \sum_{j \in \mathbb{Z}} \frac{1}{8 \Delta x^{\frac{1}{2}-\gamma}} (D_+ (a)_j)^4. \end{aligned}$$

- For the third term, Relation (3.27) together with (3.19a) gives

$$\begin{aligned} \sum_{j \in \mathbb{Z}} D_+ D_+ D_- (a)_j \frac{a_{j-1}}{4} D_- (a)_j &= \sum_{j \in \mathbb{Z}} -D_+ D_+ (a)_j D_+ \left(\frac{a_{j-1}}{4} D_- (a)_j \right) \\ &= \sum_{j \in \mathbb{Z}} -D_+ D_+ (a)_j \left\{ \frac{a_j}{4} D_+ D_- (a)_j + \frac{D_+ (a)_{j-1}}{4} D_- (a)_j \right\} \\ &= \sum_{j \in \mathbb{Z}} -\frac{a_j}{2} \left(\frac{D_+ D_+ (a)_j + D_+ D_- (a)_j}{2} \right)^2 + \sum_{j \in \mathbb{Z}} \frac{a_j}{8} (D_+ D_+ (a)_j)^2 + \sum_{j \in \mathbb{Z}} \frac{a_j}{8} (D_+ D_- (a)_j)^2 \\ &\quad - \sum_{j \in \mathbb{Z}} D_+ D_+ (a)_j \frac{(D_- (a)_j)^2}{4} \\ &\leq \sum_{j \in \mathbb{Z}} -\frac{a_j}{2} (D_+ D_c (a)_j)^2 + \sum_{j \in \mathbb{Z}} \frac{a_j + a_{j-1}}{8} (D_+ D_- (a)_j)^2 + \sum_{j \in \mathbb{Z}} \frac{\Delta x^{\frac{1}{2}-\gamma}}{8} (D_+ D_- (a)_j)^2 \\ &\quad + \sum_{j \in \mathbb{Z}} \frac{1}{8 \Delta x^{\frac{1}{2}-\gamma}} (D_+ (a)_j)^4. \end{aligned}$$

Gathering all these results yields

$$\begin{aligned} \sum_{j \in \mathbb{Z}} D_+ D_+ D_- (a)_j D_c \left(\frac{a^2}{2} \right)_j &\leq \sum_{j \in \mathbb{Z}} \left\{ \frac{\Delta x^{\frac{1}{2}-\gamma}}{2} - \frac{\Delta x}{8} D_- (a)_j + \frac{a_{j-1} - a_j}{8} \right\} (D_+ D_- (a)_j)^2 \\ &\quad + \sum_{j \in \mathbb{Z}} \frac{1}{2 \Delta x^{\frac{1}{2}-\gamma}} (D_+ a_j)^4 - \sum_{j \in \mathbb{Z}} a_j (D_+ D_c (a)_j)^2. \end{aligned}$$

Lemma B.2 concludes this proof

$$\sum_{j \in \mathbb{Z}} D_+ D_+ D_- (a)_j D \left(\frac{a^2}{2} \right)_j \leq \sum_{j \in \mathbb{Z}} \left\{ \frac{\Delta x^{\frac{1}{2}-\gamma} + \|a\|_{\ell^\infty} + 9 \frac{\|a\|_{\ell^\infty}^2}{\Delta x^{\frac{1}{2}-\gamma}}}{2} \right\} (D_+ D_- (a)_j)^2 + \sum_{j \in \mathbb{Z}} \|a\|_{\ell^\infty} (D_+ D_c (a)_j)^2.$$

□

Proof of Lemma 3.31. For Identity (3.43), the right hand side yields

$$\begin{aligned} & \sum_{j \in \mathbb{Z}} \left[D_c(a)_j \right]^2 \frac{a_{j+1}b_{j+1} + a_{j-1}b_{j-1}}{2} - \sum_{j \in \mathbb{Z}} D_c(b)_j \left[D_c(a)_j \right]^3 \frac{4\Delta x^2}{3} - \sum_{j \in \mathbb{Z}} \frac{1}{3} D_c D_c(b)_j a_j^3 \\ &= \sum_{j \in \mathbb{Z}} \frac{a_{j+1}^3 b_{j+1} + a_{j+1}^2 a_{j-1} b_{j-1} - 2a_{j+1}^2 a_{j-1} b_{j+1} - 2a_{j+1} a_{j-1}^2 b_{j-1} + a_{j-1}^2 a_{j+1} b_{j+1} + a_{j-1}^3 b_{j-1}}{8\Delta x^2} \\ & - \sum_{j \in \mathbb{Z}} \left(\frac{b_{j+1} - b_{j-1}}{2\Delta x} \right) \frac{a_{j+1}^3 - 3a_{j+1}^2 a_{j-1} + 3a_{j+1} a_{j-1}^2 - a_{j-1}^3}{8\Delta x^3} \left(\frac{4\Delta x^2}{3} \right) - \sum_{j \in \mathbb{Z}} \frac{1}{3} \frac{b_{j+2} - 2b_j + b_{j-2}}{4\Delta x^2} a_j^3. \end{aligned}$$

Adding the first two sums gives

$$\sum_{j \in \mathbb{Z}} \frac{a_{j+1}^3 b_{j+1} - 3a_{j+1}^2 a_{j-1} b_{j-1} - 3a_{j-1}^2 a_{j+1} b_{j+1} + a_{j-1}^3 b_{j-1} + 2a_{j-1}^3 b_{j+1} + 2b_{j-1} a_{j+1}^3}{24\Delta x^2}.$$

The sum yields then

$$\begin{aligned} & \sum_{j \in \mathbb{Z}} \left[D_c(a)_j \right]^2 \frac{a_{j+1}b_{j+1} + a_{j-1}b_{j-1}}{2} - \sum_{j \in \mathbb{Z}} D_c(b)_j \left[D_c(a)_j \right]^3 \frac{4\Delta x^2}{3} - \sum_{j \in \mathbb{Z}} \frac{1}{3} D_c D_c(b)_j a_j^3 \\ &= \sum_{j \in \mathbb{Z}} \frac{a_{j+1}^3 b_{j+1} - 3a_{j+1}^2 a_{j-1} b_{j-1} - 3a_{j-1}^2 a_{j+1} b_{j+1} + a_{j-1}^3 b_{j-1} + 2a_{j-1}^3 b_{j+1} + 2b_{j-1} a_{j+1}^3}{24\Delta x^2} \\ & - \sum_{j \in \mathbb{Z}} \frac{1}{3} \frac{b_{j+1} a_{j-1}^3 - 2b_j a_j^3 + b_{j-1} a_{j+1}^3}{4\Delta x^2} \\ &= \sum_{j \in \mathbb{Z}} \frac{a_{j+1}^3 b_{j+1} - a_{j+1}^2 a_{j-1} b_{j-1} - a_{j-1}^2 a_{j+1} b_{j+1} + a_{j-1}^3 b_{j-1}}{8\Delta x^2} \\ &= \sum_{j \in \mathbb{Z}} \left(\frac{a_{j+1} b_{j+1} - a_{j-1} b_{j-1}}{2\Delta x} \right) \left(\frac{a_{j+1}^2 - a_{j-1}^2}{4\Delta x} \right) \\ &= \sum_{j \in \mathbb{Z}} D_c(ab)_j D_c \left(\frac{a^2}{2} \right)_j. \end{aligned}$$

□

B.2 Proofs of useful Lemmas of Section 3.8.2.2

Proof of Lemma 3.62. • Relation (3.82) is obtained thanks to

$$\left[D_c \left(\frac{a^2}{2} \right)_j \right]^2 = \frac{(a_{j+1} - a_{j-1})^2 (a_{j+1} + a_{j-1})^2}{16\Delta x^2}$$

• Identity (3.83) comes from developing the sum

$$\begin{aligned}
 [D_+ D_- (a)]_j^2 &= \left(\frac{a_{j+1} - 2a_j + a_{j-1}}{\Delta x^2} \right)^2 \\
 &= \frac{4a_{j+1}^2 - 8a_{j+1}a_j + 4a_j^2}{\Delta x^4} + \frac{4a_{j+1}a_j - 4a_ja_{j-1}}{\Delta x^4} - \frac{3a_{j+1}^2}{\Delta x^4} + \frac{2a_{j+1}a_{j-1}}{\Delta x^4} + \frac{a_{j-1}^2}{\Delta x^4} \\
 &= \frac{4(a_{j+1} - a_j)^2}{\Delta x^4} + \frac{4}{\Delta x^3} \left(\frac{a_{j+1}a_j - a_ja_{j-1}}{\Delta x} \right) - \frac{a_{j+1}^2 - 2a_{j+1}a_{j-1} + a_{j-1}^2}{\Delta x^4} + 2 \frac{a_{j-1}^2 - a_{j+1}^2}{\Delta x^4} \\
 &= \frac{4}{\Delta x^2} [D_+ (a)]_j^2 + \frac{4}{\Delta x^3} D_+ (aS^{-1}a)_j - \frac{4}{\Delta x^2} [D_c (a)]_j^2 - \frac{4}{\Delta x^3} D_c (a^2)_j.
 \end{aligned}$$

□

Proof of Lemma 3.63. • To prove Equation (3.84), we develop the expression

$$\begin{aligned}
 a_j D_c \left(\frac{a^2}{2} \right)_j &= \frac{a_j a_{j+1}^2 - a_j a_{j-1}^2}{4\Delta x} \\
 &= \frac{a_j a_{j+1}^2 - a_{j+1} a_j^2 + a_{j+1} a_j^2 - a_j a_{j-1}^2}{4\Delta x} \\
 &= \frac{(a_j - a_{j+1})^3}{12\Delta x} - \frac{a_j^3}{12\Delta x} + \frac{a_{j+1}^3}{12\Delta x} + \frac{a_{j+1} a_j^2 - a_j a_{j-1}^2}{4\Delta x} \\
 &= -\frac{\Delta x^2}{12} [D_+ (a)]_j^3 + \frac{1}{12} D_+ (a^3)_j + \frac{1}{4} D_+ (aS^{-1}a^2)_j.
 \end{aligned}$$

• We prove now Equation (3.85).

$$\begin{aligned}
 D_+ D_- (a)_j D_c \left(\frac{a^2}{2} \right)_j &= \left(\frac{a_{j+1} - 2a_j + a_{j-1}}{\Delta x^2} \right) \left(\frac{a_{j+1}^2 - a_{j-1}^2}{4\Delta x} \right) \\
 &= \frac{a_{j+1}^3 - a_{j+1}a_{j-1}^2 + a_{j-1}a_{j+1}^2 - a_{j-1}^3}{4\Delta x^3} + \frac{2a_j a_{j-1}^2 - 2a_j a_{j+1}^2}{4\Delta x^3} \\
 &= \frac{-a_{j+1}^3 - 3a_{j+1}a_{j-1}^2 + 3a_{j-1}a_{j+1}^2 + a_{j-1}^3}{12\Delta x^3} + \frac{a_{j+1}^3 - a_{j-1}^3}{3\Delta x^3} + \frac{a_{j+1}^3 - 3a_j a_{j+1}^2 + 3a_{j+1}a_j^2 - a_j^3}{6\Delta x^3} \\
 &\quad - \frac{a_{j+1}^3 - a_j^3}{6\Delta x^3} + \frac{a_j a_{j-1}^2 - a_{j+1}a_j^2}{2\Delta x^3} \\
 &= -\frac{2}{3} \left(\frac{a_{j+1} - a_{j-1}}{2\Delta x} \right)^3 + \frac{2}{3\Delta x^2} \frac{a_{j+1}^3 - a_{j-1}^3}{2\Delta x} + \frac{1}{6} \left(\frac{a_{j+1} - a_j}{\Delta x} \right)^3 - \frac{1}{6\Delta x^2} \frac{a_{j+1}^3 - a_j^3}{\Delta x} - \frac{1}{2\Delta x^2} D_+ (aS^{-1}a^2)_j \\
 &= -\frac{2}{3} [D_c (a)]_j^3 + \frac{2}{3\Delta x^2} D_c (a^3)_j + \frac{1}{6} [D_+ (a)]_j^3 - \frac{1}{6\Delta x^2} D_+ (a^3)_j - \frac{1}{2\Delta x^2} D_+ (aS^{-1}a^2)_j.
 \end{aligned}$$

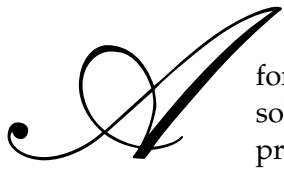
• To prove Equation (3.86), we develop in the same way.

$$\begin{aligned}
 a_j D_+ D_- (a)_j &= \frac{a_j a_{j+1} - 2a_j^2 + a_j a_{j-1}}{\Delta x^2} \\
 &= -\frac{a_{j+1}^2 - 2a_{j+1}a_j + a_j^2}{\Delta x^2} + \frac{-a_j a_{j+1} + a_{j+1}^2 - a_j^2 + a_j a_{j-1}}{\Delta x^2} \\
 &= -\frac{(a_{j+1} - a_j)^2}{\Delta x^2} + \frac{a_{j+1}^2 - a_j^2}{\Delta x^2} - \frac{a_{j+1}a_j - a_j a_{j-1}}{\Delta x^2} \\
 &= -[D_+ (a)]_j^2 + \frac{1}{\Delta x} D_+ (a^2)_j - \frac{1}{\Delta x} D_+ (aS^{-1}a)_j.
 \end{aligned}$$

□

ANNEXE C

Traveling wave solution for Fisher-KPP equation

 s explain in Subsection 5.8.1, proving the existence of a traveling wave for Fisher-KPP is hence enough to prove the existence of traveling wave solution of KdV⁻-Burgers equation. This annex is devoted to the existence proof of traveling wave for Fisher-KPP equation.

Proposition C.1. *There exist traveling waves with speed σ , solutions of Fisher-KPP equation connecting $(1, 0)^t$ to $(0, 0)^t$. These traveling waves decrease monotonically if $\sigma \geq 2$ and have an oscillatory behavior if $0 \leq \sigma < 2$.*

Remark C.2. *In most cases, the Fisher-KPP equation is used to model the evolution of a population and u stands for the density of population. Only the case $0 \leq u \leq 1$ has a biological meaning and the case of an oscillatory traveling wave is not relevant. This is why the proof for $\sigma < 2$ is often not detailed in literature. However, in the KdV-Burgers case, we do not know a priori the sign of the speed σ , we have thus developed a unified method.*

In the phase space, the Fisher-KPP EDO of the ansatz writes

$$\begin{cases} X' = Y, \\ Y' = -\sigma Y - X(1 - X), \end{cases} \quad (\text{C.1})$$

where $'$ always denotes the derivation with respect to ξ .

Proposition C.3. *The system (C.1) has two equilibrium points $(1, 0)^t$ and $(0, 0)^t$.*

- *The point $(1, 0)^t$ is a saddle point with the unstable eigenspace $\mathbb{E}_{(1,0)}^u = \text{Span} \left\{ \begin{pmatrix} 1 \\ \frac{-\sigma + \sqrt{\sigma^2 + 4}}{2} \end{pmatrix} \right\}$ and the stable eigenspace $\mathbb{E}_{(1,0)}^s = \text{Span} \left\{ \begin{pmatrix} 1 \\ \frac{-\sigma - \sqrt{\sigma^2 + 4}}{2} \end{pmatrix} \right\}$.*
- *If $\sigma \geq 2$, the point $(0, 0)^t$ is a stable node with stable eigenspace $\mathbb{E}_{(0,0)}^s = \text{Span} \left\{ \begin{pmatrix} 1 \\ \frac{-\sigma + \sqrt{\sigma^2 - 4}}{2} \end{pmatrix}, \begin{pmatrix} 1 \\ \frac{-\sigma - \sqrt{\sigma^2 - 4}}{2} \end{pmatrix} \right\}$.*
- *If $0 \leq \sigma < 2$, the point $(0, 0)^t$ is a stable spiral.*

Proof. • For the point $(1, 0)^t$, the Jacobian matrix is $\begin{pmatrix} 0 & 1 \\ 1 & -\sigma \end{pmatrix}$, with eigenvalues $\lambda_1 = \frac{-\sigma + \sqrt{\sigma^2 + 4}}{2} > 0$ and $\lambda_2 = \frac{-\sigma - \sqrt{\sigma^2 + 4}}{2} < 0$.

Annexe C.1. Location of the unstable manifold $\mathcal{W}_{(1,0)}^u$

- For the point $(0,0)^t$, the Jacobian matrix is $\begin{pmatrix} 0 & 1 \\ -1 & -\sigma \end{pmatrix}$ with eigenvalues $\lambda_1 = \frac{-\sigma + \sqrt{\sigma^2 - 4}}{2} < 0$ and $\lambda_2 = \frac{-\sigma - \sqrt{\sigma^2 - 4}}{2} < 0$ (if $\sigma \geq 2$) and $\lambda_1 = \frac{-\sigma + i\sqrt{4 - \sigma^2}}{2}$ and $\lambda_2 = \frac{-\sigma - i\sqrt{4 - \sigma^2}}{2}$ (if $0 \leq \sigma < 2$).

□

Since the point $(1,0)^t$ is a equilibrium point, it is not clear, *a priori*, if a solution can «take off» this point. This question is answered thanks to the unstable manifold theorem, Theorem 5.6.

Finding a traveling wave is eventually nothing else than proving that the unstable manifold of $(1,0)$ and the stable manifold of $(0,0)$ intersect each other. We have thus to understand what this unstable manifold $\mathcal{W}_{(1,0)}^u$ is. Since the unstable eigenspace $\mathbb{E}_{(1,0)}^u$ is a one-dimensional subspace, so does the unstable manifold $\mathcal{W}_{(1,0)}^u$ but we need to study it slightly more precisely and this is the aim of the following subsection.

C.1 Location of the unstable manifold $\mathcal{W}_{(1,0)}^u$

To study the location of the unstable manifold we use a specific function $\mathcal{L}$ whose main features are detailed as follows.

Proposition C.4. *The functional*

$$\begin{aligned} \mathcal{L} : \quad \left(-\infty, \frac{3}{2}\right) \times \mathbb{R} &\rightarrow \mathbb{R} \\ (X, Y) &\mapsto X^2 + Y^2 - \frac{2X^3}{3} \end{aligned} \tag{C.2}$$

is a Lyapunov function for $(0,0)$, i.e. it is decreasing along trajectories, $\mathcal{L}(0,0) = 0$ and $\mathcal{L}(X,Y) > 0$, for all $(X,Y)^t \neq (0,0)$.

Proof. Let $\xi \in \mathbb{R}$, system (C.1) gives

$$\begin{aligned} \frac{d}{d\xi} \mathcal{L}(X(\xi), Y(\xi)) &= 2X(\xi)Y(\xi) + 2Y(\xi)(-\sigma Y(\xi) - X(\xi)(1 - X(\xi))) - 2X^2(\xi)Y(\xi) \\ &= -2\sigma Y^2(\xi) \\ &\leq 0, \quad (\text{since } \sigma > 0). \end{aligned}$$

This function is then decreasing along trajectories.

It is straightforward to verify that $\mathcal{L}(X,Y) \geq 0$ for $(X,Y) \in (-\infty, \frac{3}{2}) \times \mathbb{R}$. □

Remark C.5. *The graph of $\mathcal{L}$ comprises two types of level sets $\Gamma_\lambda := \{(X,Y)^t, \text{ such that } \mathcal{L}(X,Y) = \lambda\}$*

- if $\lambda \geq \frac{1}{3}$, the level curve is composed only of one infinite branch
- if $\lambda < \frac{1}{3}$, the level curve is composed of a closed branch (in the half plane $\{X \leq 1\}$) and an infinite branch (in the half plane $\{X > 1\}$). These two branches do not intersect each other.

Figure C.1 illustrates the isolines of $\mathcal{L}$.

We essentially use the decreasing property of $\mathcal{L}$ and the following crucial observation : there exists a closed compact $\Omega \in (-\infty, \frac{3}{2}] \times \mathbb{R}$, containing the two equilibrium points $(0,0)^t$ and $(1,0)^t$ and the closed branches mentioned in Remark C.5. This observation comes from the following property.

Proposition C.6. Let $\Omega := \{(X, Y)^t, \text{ such that } \mathcal{L}(X, Y) \leq \mathcal{L}(1, 0) \text{ and } X \leq 1\}$.

- Then, Ω is compact and contains the two equilibrium points $(0, 0)^t$ and $(1, 0)^t$.
- All closed branches of isolines of $\mathcal{L}$ are included in Ω .
- Moreover, the unstable manifold of $(1, 0)^t$, $\mathcal{W}_{(1,0)}^u$, intersects Ω .

Remark C.7. We have represented the level sets of $\mathcal{L}$ and the domain Ω in Figure C.1.

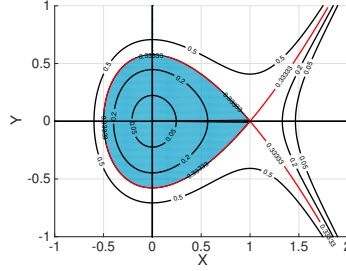


FIGURE C.1 – Level sets of $\mathcal{L}$ - the blue domain corresponds to the compact Ω and the red line to the level set going through $(1, 0)$

Proof of Proposition C.6. • The domain Ω is closed by definition. Moreover, if $(X, Y) \in \Omega$, then $X^2 + Y^2 - \frac{2X^3}{3} \leq \frac{1}{3}$ and $X \leq 1$ so

$$X^2 + Y^2 \leq \frac{1}{3} + \frac{2X^3}{3} \leq 1.$$

Thus, Ω is included in the ball of radius 1 and center $(0, 0)^t$ and then it is bounded. By Heine-Borel theorem, Ω is compact.

- Since $\mathcal{L}$ is a Lyapunov function, one has $\mathcal{L}(0, 0) = 0 < \mathcal{L}(1, 0)$ then $(0, 0)^t$ and $(1, 0)^t$ are included in Ω .
- By Remark C.5, the closed branches of isolines correspond to a level $\lambda < \frac{1}{3} = \mathcal{L}(1, 0)$ and $X \leq 1$. They then compose the domain Ω .
- The level set $\mathcal{L}(1, 0)$ (the red line in Figure C.1) is characterised by the equation

$$X^2 + Y^2 - \frac{2X^3}{3} = \frac{1}{3}$$

which corresponds to the two following curves

$$\begin{cases} Y_{top}(X) = \sqrt{\frac{2X^3}{3} - X^2 + \frac{1}{3}}, \\ Y_{bottom}(X) = -\sqrt{\frac{2X^3}{3} - X^2 + \frac{1}{3}}. \end{cases}$$

Both curves are not differentiable at $X = 1$ but own left and right derivatives. Indeed,

$$\frac{Y_{top}(X) - Y_{top}(1)}{X - 1} = \frac{\sqrt{\frac{2X^3}{3} - X^2 + \frac{1}{3}} - 0}{X - 1} = \frac{\sqrt{\frac{2(X-1)^3}{3} + (X-1)^2}}{X - 1} = \text{sgn}(X - 1) \sqrt{\frac{2(X-1)}{3} + 1}.$$

Likewise, one has

$$\frac{Y_{bottom}(X) - Y_{bottom}(1)}{X - 1} = -\text{sgn}(X - 1) \sqrt{\frac{2(X-1)}{3} + 1}.$$

Annexe C.1. Location of the unstable manifold $\mathcal{W}_{(1,0)}^u$

The right and left derivatives hence verify

$$\begin{cases} \partial_- Y_{top}(1) := \lim_{X \rightarrow 1^-} \frac{Y_{top}(X) - Y_{top}(1)}{X - 1} = -1, \\ \partial_+ Y_{top}(1) := \lim_{X \rightarrow 1^+} \frac{Y_{top}(X) - Y_{top}(1)}{X - 1} = 1, \end{cases} \quad \text{and} \quad \begin{cases} \partial_- Y_{bottom}(1) := \lim_{X \rightarrow 1^-} \frac{Y_{bottom}(X) - Y_{bottom}(1)}{X - 1} = 1, \\ \partial_+ Y_{bottom}(1) := \lim_{X \rightarrow 1^+} \frac{Y_{bottom}(X) - Y_{bottom}(1)}{X - 1} = -1. \end{cases}$$

Thus, the equations of the "tangents" to the level set $\mathcal{L}(1,0)$ at $(1,0)$ are

$$\begin{cases} y = -x + 1, \\ y = x - 1. \end{cases}$$

To verify if the unstable manifold $\mathcal{W}_{(1,0)}^u$ intersects Ω , we just have to ensure that the two scalar products between the unstable eigenvector and the two tangent vectors $\mathbf{t}_{top} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ and

$\mathbf{t}_{bottom} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ are both non negative or both non positive. It yields

$$\left\langle \begin{pmatrix} 1 \\ \frac{-\sigma + \sqrt{\sigma^2 + 4}}{2} \end{pmatrix}, \mathbf{t}_{top} \right\rangle = -1 + \frac{-\sigma + \sqrt{\sigma^2 + 4}}{2} \leq 0, \quad \text{car } \sigma \geq 0,$$

and

$$\left\langle \begin{pmatrix} 1 \\ \frac{-\sigma + \sqrt{\sigma^2 + 4}}{2} \end{pmatrix}, \mathbf{t}_{bottom} \right\rangle = -1 + \frac{\sigma - \sqrt{\sigma^2 + 4}}{2} \leq 0, \quad \text{car } \sigma \geq 0.$$

Therefore, proposition C.6 is true. $\square$

Figure C.2 illustrates the intersection between Ω and $\mathcal{W}_{(1,0)}^u$.

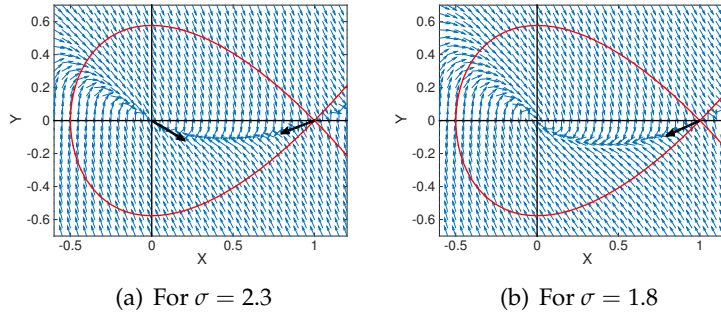


FIGURE C.2 – Vector field of the Fisher-KPP autonomous system and trapping region Ω - the black vectors correspond to unstable eigenvector for $(1,0)^t$ and stable eigenvector for $(0,0)^t$ which are connected via the traveling wave

Remark C.8. The Lyapunov function $\mathcal{L}$ of the Fisher-KPP equation defined in Proposition C.4 closely resembles the Hamiltonian $\mathcal{H}$ for (KdV-KS) equation (or more generally Equation (5.1)) defined in Proposition 5.30 in Chapter 5, in terms of both expressions (C.2) versus (5.22)) and Figures (Fig. C.1 versus 5.3). Indeed, a change of coordinates enables us to switch from one to another. By defining

$$\begin{cases} X^{fkpp}(\xi) := h(0)X(\sqrt{\Lambda}\xi) + 1, \\ Y^{fkpp}(\xi) := \sqrt{\Lambda}h(0)Y(\sqrt{\Lambda}\xi), \end{cases}$$

then, if $(X, Y)^t$ satisfies the limit system (5.21), the variables $(X^{fkpp}, Y^{fkpp})^t$ satisfies the Fisher-KPP system (C.1) with a null speed, i.e with $\sigma = 0$ and vice versa. It follows that trajectories of Fisher-KPP with a null speed are constants along level sets of $\mathcal{L}$ but only decreasing along these level sets with a non zero speed.

C.2 The proof of Proposition C.1

The existence of a traveling wave for Fisher-KPP equation follows the same guidelines as the one used for Equation (5.1). We choose $(X_0, Y_0) \in \mathcal{W}_{(1,0)}^u \cap \Omega$ (it is possible thanks to Proposition C.6) and consider the Cauchy problem

$$\begin{cases} X'(\xi) = Y(\xi), & \xi \in \mathbb{R}, \\ Y'(\xi) = -\sigma Y(\xi) - X(\xi)(1 - X(\xi)), & \xi \in \mathbb{R}, \\ (X(0), Y(0))^t = (X_0, Y_0)^t. \end{cases}$$

- **Proof of the existence of a local solution around $(1, 0)^t$.**

By Cauchy-Lipschitz theorem, there exists a unique maximal solution $(X, Y)^t$ defined locally around $(1, 0)^t$. Moreover, thanks to the unstable manifold theorem, it converges to $(1, 0)^t$ in $-\infty$. Thereafter, $(X, Y)^t$ will denote this solution.

- **Extension to $\mathbb{R}$ of the existence region of the solution.**

Proposition C.9. *The maximal solution from $(X_0, Y_0)^t$ is bounded. More precisely, it is trapped in Ω .*

Proof. Let $\xi > 0$ be in the domain of definition of the maximal solution $(X, Y)^t$. We know that $(X, Y)^t$ is defined locally around $(1, 0)^t$ then $(X(\xi), Y(\xi))^t \in (-\infty, \frac{3}{2}) \times \mathbb{R}$, the domain where $\mathcal{L}$ is a Lyapunov function.

Since $(X_0, Y_0)^t \in \Omega$, then $\mathcal{L}(X_0, Y_0) \leq \mathcal{L}(1, 0)$. However, by Proposition C.4, the function $\mathcal{L}$ is a Lyapunov function, it then decreases along trajectories. Therefore, it holds

$$\mathcal{L}(X(\xi), Y(\xi)) \leq \mathcal{L}(X_0, Y_0) \leq \mathcal{L}(1, 0) = \frac{1}{3}. \quad (\text{C.3})$$

Thanks to Remark C.5, $(X, Y)^t$ belongs then to a level curve composed of two branches which do not intersect each other. Since $(X, Y)^t$ is continuous, it can not jump from one branch to another. The choice of $(X_0, Y_0)^t$ imposes $X_0 \leq 1$ and so $(X, Y)^t$ belongs to the closed branch of the isoline :

$$X(\xi) \leq 1. \quad (\text{C.4})$$

Conditions (C.3) and (C.4) imply that $(X(\xi), Y(\xi))^t \in \Omega$. The maximal solution is then trapped in the compact Ω . $\square$

By the «Compact break» theorem 5.52, we conclude that the maximal solution $(X, Y)^t$ is global.

- **Convergence (with monotonicity or oscillations) towards $(0, 0)^t$.**

Corollary C.10. *The maximal solution $(X, Y)^t$ converges to $(0, 0)^t$.*

Proof. By the Poincaré-Bendixson theorem 5.54, we have to eliminate the periodic cycle case and the homoclinic case (when $(X, Y)^t$ goes back to $(1, 0)^t$). We have proved that the maximal solution stays in Ω included in the domain of definition of $\mathcal{L}$. Then $\mathcal{L}$ is decreasing along trajectories. Thanks the proof of Proposition C.4, we may even notice that

$$\frac{d}{d\xi} \mathcal{L}(X(\xi), Y(\xi)) < 0 \quad \text{if} \quad Y(\xi) \neq 0.$$

A trajectory can not be periodic (or converge to a periodic cycle) around $(0, 0)^t$ because if such a trajectory would exist, it would necessarily intersect the domain $\mathbb{R} \times \mathbb{R}^*$, where $\mathcal{L}$

Annexe C.2. The proof of Proposition C.1

is strictly decreasing. It could not reach its previous $\mathcal{L}$ -level in the future.

Since $\mathcal{W}_{(1,0)}^u$ is not tangent to the horizontal axis, and since we have chosen $(X_0, Y_0)^t \in \mathcal{W}_{(1,0)}^u \cap \Omega$, the functional $\mathcal{L}$ will be strictly decreasing around (X_0, Y_0) and will not reach its initial $\mathcal{L}$ -level in the future, which bans the homoclinic case. $\square$

Proposition C.11. *The traveling wave of speed σ is monotone if $\sigma > 2$. Otherwise, it oscillates around $(0, 0)^t$.*

Proof. If $\sigma > 2$, $(0, 0)^t$ is a stable node of the linearized system (cf Proposition C.3) whereas it is a stable spiral in the small speed case. $\square$

Remark C.12. *There exists an other (more well-known) proof of existence of a traveling wave for Fisher-KPP equation if $\sigma > 2$. This proof is based on the same guidelines as those before, but the compact Ω is replaced by a triangle trapping region. More precisely, let $\mathcal{T}$ be the triangle defined by the vertex $(1, 0)^t$, $(0, 0)^t$ and $(1, -\frac{\sigma}{2})^t$. Then the vector field defined by the Fisher-KPP autonomous system is pointing inwards $\mathcal{T}$. Indeed, the outside normal of the triangle $\mathcal{T}$ are $\mathbf{n}_{top} = (0, 1)^t$, $\mathbf{n}_{right} = (1, 0)^t$ and $\mathbf{n}_{inclined} = (-\frac{\sigma}{2}, -1)^t$, and all the scalar products with respect to the vector field are non positive :*

$$\left\langle \left(\begin{array}{c} Y \\ -\sigma Y - X(1-X) \end{array} \right)_{\substack{X \in [0,1] \\ Y=0}}, \mathbf{n}_{top} \right\rangle = -X(1-X) \leq 0, \text{ (since } X \in [0,1])$$

and

$$\left\langle \left(\begin{array}{c} Y \\ -\sigma Y - X(1-X) \end{array} \right)_{\substack{X=1 \\ Y \in [-\frac{\sigma}{2}, 0]}}, \mathbf{n}_{right} \right\rangle = Y \leq 0, \text{ (since } Y \in [-\frac{\sigma}{2}, 0])$$

and

$$\begin{aligned} \left\langle \left(\begin{array}{c} Y \\ -\sigma Y - X(1-X) \end{array} \right)_{\substack{X \in [0,1] \\ Y = -\frac{\sigma}{2}X}}, \mathbf{n}_{inclined} \right\rangle &= \frac{\sigma}{2}Y + X(1-X) \\ &= \left[-\frac{\sigma^2}{4} + 1 - X \right] X \leq 0, \text{ (since } \sigma > 2). \end{aligned}$$

Then the triangle $\mathcal{T}$ is a compact trapping region, which possesses one branch of $\mathcal{W}_{(0,1)}^u$ (the one directed by the eigenvector $(-1, \frac{\sigma - \sqrt{\sigma^2 + 4}}{2})^t$) and one branch of $\mathcal{W}_{(0,0)}^s$ (the one directed by the eigenvector $(1, \frac{-\sigma + \sqrt{\sigma^2 + 4}}{2})^t$).

The vector field as well as the trapping region $\mathcal{T}$ are represented in Figure C.3. We can see in the zoom figure that $\mathcal{T}$ is not a trapping region anymore as soon as $\sigma < 2$.

We prefer the previous method with Ω to the one with $\mathcal{T}$ because $\sigma = \frac{\lambda}{\sqrt{\mu}} \sqrt{-\frac{1}{c-u_-}}$ in our case. We do not know a priori the sign of $\sigma - 2$, so it is easier to have a unique method which covers all the cases.

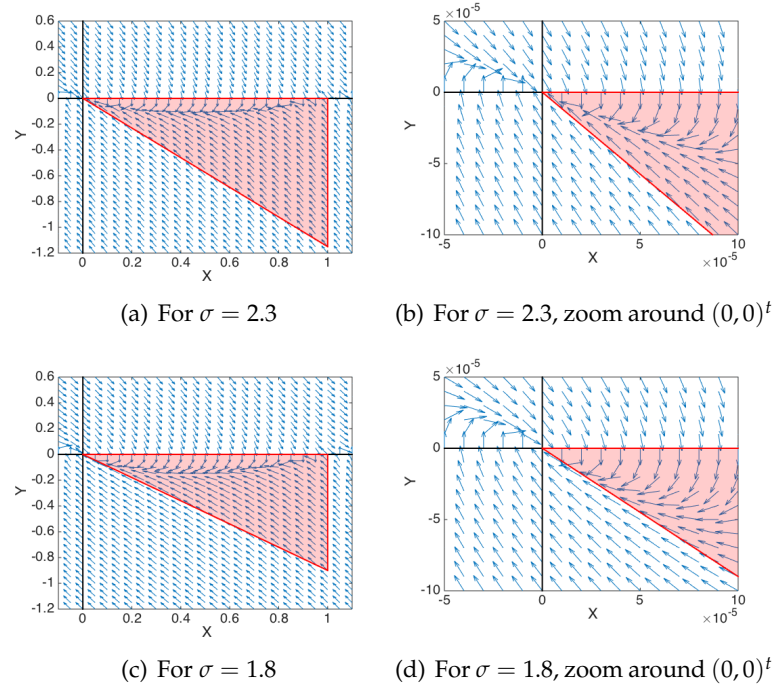


FIGURE C.3 – Vector field of the Fisher-KPP autonomous system and trapping triangle $\mathcal{T}$ - the black vectors correspond to unstable eigenvector for $(1,0)^t$ and stable eigenvector for $(0,0)^t$

Bibliographie

- [1] J. C. Alexander and R. Sachs. Linear Instability of Solitary Waves of a Boussinesq-type Equation : a computer assisted computation. *Nonlinear World*, 2(4) :471–507, 1995. → page(s) [197](#), [198](#)
- [2] B. Alvarez-Samaniego and D. Lannes. Large time existence for 3D water-waves and asymptotics. *Invent. math.*, 171 :485–541, 2008. → page(s) [5](#), [6](#), [7](#)
- [3] C. J. Amick. Regularity and uniqueness of solutions to the Boussinesq system of equations. *J. Differential Equations*, 54(2) :231–247, 1984. → page(s) [12](#)
- [4] C. T. Anh. On the Boussinesq/Full dispersion systems and Boussinesq/Boussinesq systems for internal waves. *Nonlinear Analysis : Theory, Methods & Applications*, 72(1) :409–429, 2010. → page(s) [11](#)
- [5] D. C. Antonopoulos, V. A. Dougalis, and D. E. Mitsotakis. Galerkin approximations of periodic solutions of Boussinesq systems. *Bulletin of the Greek Mathematical Society*, 57 :13–30, 2010. → page(s) [22](#)
- [6] D. C. Antonopoulos, V. A. Dougalis, and D. E. Mitsotakis. Numerical solution of Boussinesq systems of the Bona–Smith family. *Appl. Numer. Math*, 60(4) :314–336, 2010. → page(s) [22](#)
- [7] U. M. Ascher and R. I. McLachlan. On Symplectic and Multisymplectic Schemes for the KdV Equation. *J. Sci. Comput.*, 25(1) :83–104, 2005. → page(s) [19](#)
- [8] G. A. Baker, V. A. Dougalis, and O. A. Karakashian. Convergence of Galerkin Approximations for the Korteweg-de Vries Equation. *Math. Comp.*, 40(162) :419–433, 1983. → page(s) [19](#)
- [9] B. Barker. Numerical proof of stability of roll waves in the small-amplitude limit for inclined thin film flow. *J. Differential Equations*, 257(8) :2950–2983, 2014. → page(s) [9](#)
- [10] T. B. Benjamin, J. L. Bona, and J. J. Mahony. Model equations for long waves in nonlinear dispersive systems. *Philos. Trans. Royal Soc. London, Ser. A*, 272 :47–78, 1972. → page(s) [7](#), [8](#)
- [11] S. Benzoni-Gavage, D. Serre, and K. Zumbrun. Alternate Evans function and viscous shock waves. *SIAM J. Math. Anal*, 32(5) :929–962, 2001. → page(s) [194](#), [200](#)
- [12] A. L. Bertozzi, A. Münch, M. Shearer, and K. Zumbrun. Stability of compressive and undercompressive thin film travelling waves. *Euro. Jnl. of Applied Mathematics*, 12 :253–291, 2001. → page(s) [208](#)
- [13] J. L. Bona, H. Chen, O. Karakashian, and Y. Xing. Conservative, discontinuous Galerkin-methods for the generalized Korteweg-de Vries equation. *Math. Comp.*, 82(283) :1401–1432, 2013. → page(s) [20](#), [100](#)

BIBLIOGRAPHIE

- [14] J. L. Bona and M. Chen. A Boussinesq system for two-way propagation of nonlinear dispersive waves. *Phys. D*, 116(1) :191–224, 1998. → page(s) [22](#), [119](#), [148](#)
- [15] J. L. Bona and M. Chen. Singular Solutions of a Boussinesq System for Water Waves. *J. Math. Study*, 49(3) :205–220, 2016. → page(s) [22](#), [152](#), [153](#)
- [16] J. L. Bona, M. Chen, and J.-C. Saut. Boussinesq equations and other systems for small amplitude long waves in nonlinear dispersive media. I : Derivation and linear theory. *J. Nonlinear Sci.*, 12(4) :283–318, 2002. → page(s) [4](#), [207](#)
- [17] J. L. Bona, M. Chen, and J.-C. Saut. Boussinesq equations and other systems for small-amplitude long waves in nonlinear dispersive media : II. The nonlinear theory. *Nonlinearity*, 17(3) :925–952, 2004. → page(s) [11](#), [12](#)
- [18] J. L. Bona, T. Colin, and D. Lannes. Long Wave Approximations for Water Waves. *Arch. Ration. Mech. Anal.*, 178(3) :373–410, 2005. → page(s) [5](#), [6](#)
- [19] J. L. Bona, V. A. Dougalis, and D. E. Mitsotakis. Numerical solution of KdV–KdV systems of Boussinesq equations : I. The numerical scheme and generalized solitary waves. *Math. Comput. Simulation*, 74(2) :214–228, 2007. → page(s) [22](#)
- [20] J. L. Bona, V. A. Dougalis, and D. E. Mitsotakis. Numerical solution of Boussinesq systems of KdV–KdV type : II. Evolution of radiating solitary waves. *Nonlinearity*, 21(12) :28–25, 2008. → page(s) [22](#)
- [21] J. L. Bona and M. E. Schonbek. Travelling-wave solutions to Korteweg-de Vries-Burgers equations. *Proc. Roy. Soc. Edinburgh, Ser. A*, 101 :207–226, 1985. → page(s) [28](#), [30](#), [184](#), [185](#)
- [22] J. L. Bona and R. Smith. The Initial-Value Problem for the Korteweg-de Vries Equation. *Philos. Trans. Roy. Soc. London Ser. A*, 278(1287) :555–601, 1975. → page(s) [11](#)
- [23] J. L. Bona and R. Smith. A model for the two-way propagation of water waves in a channel. *Math. Proc. Cambridge Philos. Soc.*, 79(1) :167–182, 1976. → page(s) [12](#)
- [24] J. Bourgain. Fourier restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. *Geometric and Functional Anal.*, 3 :107–156, 209–262, 1993. → page(s) [10](#), [11](#), [72](#)
- [25] D. Bresch, F. Couderc, P. Noble, and J.-P. Vila. A generalization of the quantum Bohm identity : Hyperbolic CFL condition for Euler-Korteweg equations. *Comptes Rendus Mathématiques*, 354(1) :39–43, 2016. → page(s) [11](#), [117](#)
- [26] T. J. Bridges, G. Derks, and G. Gottwald. Stability and instability of solitary waves of the fifth-order KdV equation : a numerical framework. *Physica D*, 172 :190–216, 2002. → page(s) [197](#), [198](#)
- [27] C. Burtea. New long time existence results for a class of Boussinesq-type systems. *Journal de Mathématiques Pures et Appliquées*, 106(2) :203–236, 2016. → page(s) [12](#), [113](#), [116](#), [117](#), [119](#), [125](#)
- [28] C. Burtea and C. Courtès. Energy estimates for some numerical schemes on the *abcd* system. Manuscript en préparation, 2017. → page(s) [30](#)
- [29] F. Chazel. Influence of bottom topography on long water waves. *ESAIM Math. Model. Numer. Anal.*, 41(4) :771–799, 2007. → page(s) [5](#)
- [30] M. Chen. Exact Traveling-Wave Solutions to Bidirectional Wave Equations. *Internat. J. Theoret. Phys.*, 37(5) :1547–1567, 1998. → page(s) [119](#), [148](#), [149](#), [151](#)

-
- [31] M. Chen. Equations for bi-directional waves over an uneven bottom. *Math. Comput. Simulation*, 62(1) :3–9, 2003. → page(s) [5](#)
 - [32] D. S. Clark. Short proof of a discrete Grönwall inequality. *Discrete Applied Mathematics*, 16 :279–281, 1987. → page(s) [130](#)
 - [33] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao. Sharp global well-posedness for KdV and modified KdV on $\mathbb{R}$ and $\mathbb{T}$. *J. Amer. Math. Soc.*, 16 :705–749, 2003. → page(s) [11](#)
 - [34] C. Courtès. Convergence for PDEs with an arbitrary odd order spatial derivative term. accepted in Springer Proceedings in Mathematics and Statistics, 2017. → page(s) [29](#)
 - [35] C. Courtès, E. Franck, P. Helluy, and H. Oberlin. Study of physics-based preconditioning with high-order Galerkin discretization for hyperbolic wave problems. *Essaim proceedings and survey*, 55 :61–82, 2016. → page(s) [4](#)
 - [36] C. Courtès, F. Lagoutière, and F. Rousset. Numerical analysis with error estimates for the Korteweg-de Vries equation. preprint, 2017. → page(s) [29](#)
 - [37] W. Craig. An existence theory for water waves and the Boussinesq and Korteweg-de Vries scaling limits. *Comm. Partial Differential Equations*, 10(8) :787–1003, 1985. → page(s) [6](#)
 - [38] C. M. Dafermos. The second law of thermodynamics and stability. *Arch. Rational Mech. Anal.*, 70(2) :167–179, 1979. → page(s) [97](#)
 - [39] C. M. Dafermos. *Hyperbolic Conservation Laws in Continuum Physics*, volume 325. Springer Berlin Heidelberg, 3rd edition, 2010. → page(s) [21](#), [70](#), [71](#), [97](#)
 - [40] P. Daripa and R. K. Dash. A class of model equations for bi-directional propagation of capillary-gravity waves. *Internat. J. Engrg. Sci*, 41 :201–218, 2003. → page(s) [5](#)
 - [41] J. de Vries. *Elements of Topological Dynamics*, volume 257. Springer Netherlands, 1993. → page(s) [176](#)
 - [42] J.-P. Demailly. *Analyse numérique et équations différentielles*. EDP Sciences, 3e édition, 2006. → page(s) [38](#), [47](#)
 - [43] B. Després. Finite volume transport schemes. *Numer. Math.*, 108 :529–556, 2008. → page(s) [15](#)
 - [44] R. J. DiPerna. Uniqueness of Solutions to Hyperbolic Conservation Laws. *Indiana Univ. Math. J.*, 28(1) :137–188, 1979. → page(s) [70](#), [97](#)
 - [45] V. A. Dougalis and O. A. Karakashian. On Some High-Order Accurate Fully Discrete Galerkin Methods for the Korteweg-de Vries Equation. *Math. Comp.*, 45(172) :329–345, 1985. → page(s) [19](#)
 - [46] J. W. Evans. Nerve axon equations III : Stability of the nerve impulse. *Indiana Univ. Math. J*, 22 :577–593, 1972. → page(s) [194](#)
 - [47] B. Feng and R. Hu. A Survey On Homoclinic and Heteroclinic Orbits. *Appl. Math. E-Notes*, 3 :16–37, 2003. → page(s) [177](#)
 - [48] R. A. Fisher. The wave of advance of advantageous genes. *Ann. Eugenics*, 7 :355–369, 1937. → page(s) [186](#)

BIBLIOGRAPHIE

- [49] E. Godlewski and P.-A. Raviart. *Hyperbolic systems of conservation laws*, volume 3/4. Mathématiques & Applications (Paris) [Mathematics and Applications] Ellipses, Paris, 1991. → page(s) [67](#)
- [50] I. S. Greig and J. L. Morris. A Hopscotch Method for the Korteweg-de Vries Equation. *J. Comp. Phys*, 20 :64–80, 1976. → page(s) [18](#)
- [51] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, chapter Averaging and Perturbation from a Geometric Viewpoint, pages 166–226. Springer New York, 1983. → page(s) [175](#)
- [52] M. Haragus and G. Iooss. *Local bifurcations, center manifolds, and normal forms in infinite dimensional dynamical systems*. Springer EDP Sciences, 2011. → page(s) [160](#), [165](#), [166](#)
- [53] D. Henry. *Geometric Theory of Semilinear Parabolic Equations*, volume 840 of *Lecture Note in Mathematics*. Springer. New-York, 1981. → page(s) [193](#)
- [54] H. Holden, K. H. Karlsen, N. H. Risebro, and T. Tao. Operator splitting for the KdV equation. *Math. Comp.*, 80(274) :821–846, 2011. → page(s) [19](#)
- [55] H. Holden, U. Koley, and N. H. Risebro. Convergence of a fully discrete finite difference scheme for the Korteweg-de Vries equation. *IMA J. Numer. Anal.*, 35(3) :1047–1077, 2015. → page(s) [18](#), [68](#)
- [56] H. Holden, C. Lubich, and H. Risebro. Operator Splitting for partial differential equations with Burgers non linearity. *Math. Comp*, 82(281) :173–185, 2013. → page(s) [19](#), [68](#)
- [57] J. K. Hunter and J.-M. Vanden-Broeck. Solitary and periodic gravity-capillary waves of finite amplitude. *J. Fluid Mech*, 134 :205–219, 1983. → page(s) [7](#)
- [58] M. A. Johnson, P. Noble, L. M. Rodrigues, and K. Zumbrun. Spectral stability of periodic wave trains of the Korteweg-de Vries/Kuramoto-Sivashinsky equation in the Korteweg-de Vries limit. *Trans. Amer. Math. Soc*, 367 :2159–2212, 2015. → page(s) [8](#)
- [59] R. S. Johnson. A non-linear equation incorporating damping and dispersion. *J. Fluid Mech*, 42(1) :49–60, 1970. → page(s) [28](#)
- [60] B. B. Kadomtsev and V. I. Petviashvili. On the stability of solitary waves in weakly dispersing media. *Sov. Phys. Doklady*, 15(6) :539–541, 1970. → page(s) [6](#)
- [61] H. Kanazawa, T. Matsuo, and T. Yaguchi. A conservative compact finite difference scheme for the KdV equation. *JSIAM Letters*, 4 :5–8, 2012. → page(s) [19](#)
- [62] T. Kano and T. Nishida. Une justification mathématique pour l’équation de Korteweg-de Vries approchant des ondes longues de surface de l’eau. *Proc. Japan Acad*, 61(10) :345–348, 1985. → page(s) [6](#)
- [63] T. Kato. On the Korteweg-de Vries equation. *Manuscripta Math*, 28(1–3) :89–99, 1979. → page(s) [9](#)
- [64] C. E. Kenig, G. Ponce, and L. Vega. Well-posedness of the initial value problem for the Korteweg-de Vries equation. *J. Amer. Math. Soc.*, 4(2) :323–347, 1991. → page(s) [10](#), [11](#), [72](#), [73](#)
- [65] C. E. Kenig, G. Ponce, and L. Vega. The Cauchy problem for the Korteweg-de Vries equation in Sobolev spaces of negative indices. *Duke Math. J.*, 77(1) :1–21, 1993. → page(s) [72](#)

-
- [66] C. E. Kenig, G. Ponce, and L. Vega. A bilinear estimate with applications to the KdV equation. *Journal A. M. S.*, 9(2) :573–603, 1996. → page(s) [10](#)
 - [67] U. Koley. Finite difference schemes for the Korteweg-de Vries-Kawahara equation. *International Journal of Numerical analysis and modeling*, 13(3) :344–367, 2016. → page(s) [68](#)
 - [68] A. Kolmogorov, I. Petrovskii, and N. Piskunov. A study of the diffusion equation with increase in the amount of substance, and its application to a biological problem. In V.M. Tikhomirov editor, *Selected Works of A.N. Kolmogorov I*, Springer Netherlands, pages 242–270, 1991. Translated by V.M. Volosov from *Bull. Moscow Univ., Math. Mech.* 1 :1–26, 1937. → page(s) [186](#)
 - [69] D. J. Korteweg and G. de Vries. On the change of form of long waves advancing in a rectangular canal and on a new type of long stationary waves. *Philos. Mag.*, 5(39) :422–443, 1895. → page(s) [6](#)
 - [70] Y. Kuramoto and T. Tsuzuki. On the Formation of Dissipative Structures in Reaction-Diffusion Systems. *Progr. Theoret. Phys.*, 54(3), 1975. → page(s) [26](#)
 - [71] Y. A. Kuznetsov. *Elements of Applied Bifurcation Theory*. 2nd edition, Springer New York, 1998. → page(s) [165](#), [175](#), [177](#), [178](#)
 - [72] P. Lafitte-Godillon. *Stabilité des profils de chocs dans les systèmes de lois de conservation*. PhD thesis, École Normale Supérieure de Lyon, 2001. → page(s) [200](#), [203](#)
 - [73] F. Lagoutière. *Modélisation mathématique et résolution numérique de problèmes de fluides compressibles à plusieurs constituants*. PhD thesis, Université Pierre et Marie Curie, 2000. → page(s) [183](#)
 - [74] D. Lannes. *The water waves problem*, volume 188 of *Math. Surveys Monogr.* American Mathematical Society, 2013. → page(s) [5](#), [6](#)
 - [75] D. Lannes and J.-C. Saut. Weakly transverse Boussinesq systems and the Kadomtsev-Petviashvili approximation. *Nonlinearity*, 19 :2853–2875, 2006. → page(s) [4](#), [6](#)
 - [76] P. D. Lax and R. D. Richtmyer. Survey of the stability of linear difference equations. *Comm. Pure Appl. Math.*, 9 :267–293, 1956. → page(s) [16](#)
 - [77] R. Levy and M. Shearer. The motion of a thin liquid film driven by surfactant and gravity. *SIAM J. Appl. Math.*, 66(5) :1588–1609, 2006. → page(s) [28](#)
 - [78] J. Li and M. R. Visbal. High-order Compact Schemes for Nonlinear Dispersive Waves. *J. Sci. Comput.*, 6(1) :1–23, 2006. → page(s) [19](#)
 - [79] F. Linares, D. Pilod, and J. C. Saut. Well-Posedness of strongly dispersive two-dimensional surface wave Boussinesq systems. *SIAM J. Math. Anal.*, 44(6) :4195–4221, 2012. → page(s) [11](#)
 - [80] F. Linares and G. Ponce. *Introduction to nonlinear dispersive equations*. Springer Verlag, 2nd edition, 2015. → page(s) [11](#), [72](#)
 - [81] Y. Mammeri. Comparaison entre modèles d’ondes de surface en dimension 2. *ESAIM : M2AN*, 41(3) :513–542, 2007. → page(s) [8](#)
 - [82] M. B. A. Mansour. Traveling wave solutions of the Burgers-KdV equation with a fourth order term. *Reports on Mathematical Physics*, 63(2) :153–161, 2009. → page(s) [28](#)

BIBLIOGRAPHIE

- [83] Q. Mengzhao. Difference Schemes for the Dispersive Equation. *Computing*, 31(3) :261–267, 1983. → page(s) [38](#)
- [84] M. S. Mock. On Fourth-Order Dissipation and Single Conservation Laws. *Communications on Pure and Applied Mathematics*, 29(4) :383–388, 1976. → page(s) [27](#), [28](#)
- [85] J. D. Murray. *Mathematical Biology : I. An Introduction*. Springer, Third edition, 1993. → page(s) [187](#)
- [86] P. Noble and L. M. Rodrigues. Whitam’s Modulation Equations and Stability of Periodic Wave Solutions of the Korteweg-de Vries-Kuramoto-Sivashinsky Equation. *Indiana Univ. Math. J*, 62(3) :753–783, 2013. → page(s) [9](#)
- [87] R. L. Pego, P. Smereka, and M. I. Weinstein. Oscillatory instability of traveling waves for a KdV-Burgers equation. *Physica D*, 67 :45–65, 1993. → page(s) [204](#), [208](#)
- [88] R. L. Pego and M. I. Weinstein. Eigenvalues and Instabilities of Solitary Waves. *Phil. Trans. R. Soc. Lond. A*, 340(1656) :47–94, 1992. → page(s) [193](#), [194](#)
- [89] D. H. Peregrine. Calculations of the development of an undular bore. *J. Fluid. Mech.*, 25(02) :321–330, 1966. → page(s) [7](#), [22](#)
- [90] V. 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. → page(s) [67](#)
- [91] J.-C. Saut and R. Temam. Remarks on the Korteweg-de Vries equation. *Israel Journal of Mathematics*, 24(1) :78–87, 1976. → page(s) [9](#)
- [92] S. Schecter. Adjoint equation and Melnikov function. Lecture notes, 2013. → page(s) [176](#), [177](#)
- [93] M. E. Schonbek. Existence of solutions for the Boussinesq system of equations. *J. Differential Equations*, 42(3) :325–352, 1981. → page(s) [12](#)
- [94] D. Serre. *Systems of Conservation Laws I : Hyperbolicity, Entropies, Shock Waves*. Cambridge University Press, 1999. → page(s) [28](#)
- [95] G. I. Sivashinsky. Nonlinear analysis of hydrodynamic instability in laminar flames – I. Derivation of basic equations. *Acta Astronautica*, 4(11–12) :1177–1206, 1977. → page(s) [26](#)
- [96] A. Sjöberg. On the Korteweg-de Vries equation : existence and uniqueness. *J. Math. Anal. Appl.*, 29 :569–579, 1970. → page(s) [18](#)
- [97] E. Tadmor. Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems. *Acta Numerica*, 12 :451–512, 2003. → page(s) [117](#)
- [98] T. R. Taha and M. J. Ablowitz. Analytical and Numerical Aspects of Certain Nonlinear Evolution Equation. III Numerical, Korteweg-de Vries Equation. *J. Comp. Phys*, 55 :231–253, 1984. → page(s) [18](#)
- [99] T. Tao. *Nonlinear Dispersive Equations : Local and Global Analysis*. CBMS Regional Conference Series in Mathematics. American Mathematical Society, 2006. → page(s) [11](#), [73](#)
- [100] G. Teschl. *Ordinary Differential Equations and Dynamical Systems*, volume 140 of *Graduate Studies in Mathematics*. American Mathematical Society, 2012. → page(s) [160](#), [183](#)

- [101] A. C. Vliegenthart. On finite-difference methods for the Korteweg-de Vries equation. *Journal of Engineering Mathematics*, 5(2) :137–155, 1971. → page(s) [17](#), [18](#)
- [102] H. A. Win. Model equation of surface waves of viscous fluid down an inclined plane. *J. Math. Kyoto Univ. (JMKYAZ)*, 33(3) :803–824, 1993. → page(s) [8](#), [9](#)
- [103] R. Winther. A conservative finite element method for the Korteweg-de Vries equation. *Math. Comp.*, 34(149) :23–43, 1980. → page(s) [18](#)
- [104] J. Yu and J. Kevorkian. Nonlinear evolution of small disturbances into roll waves in an inclined open channel. *J. Fluid Mech*, 243 :575–594, 1992. → page(s) [9](#)
- [105] J. Yu and Y. Yan. Evolution of small periodic disturbances into roll waves in channel flow with internal dissipation. *Stud. Appl. Math*, 111(1) :1–27, 2003. → page(s) [9](#)
- [106] N. J. Zabusky and M. D. Kruskal. Interaction of "solitons" in a collisionless plasma and the recurrence of initial states. *Phys. Rev. Lett.*, 15(6) :240–243, 1965. → page(s) [17](#)

Titre : Analyse numérique de systèmes hyperboliques-dispersifs

Mots Clefs : équations aux dérivées partielles, équation de Korteweg-de Vries, différences finies, estimations d'erreur, convergence numérique, ondes progressives

Résumé : Le but de cette thèse est d'étudier certaines équations aux dérivées partielles hyperboliques-dispersives. Une part importante est consacrée à l'analyse numérique et plus particulièrement à la convergence de schémas aux différences finies pour l'équation de Korteweg-de Vries et les systèmes $abcd$ de Boussinesq. L'étude numérique suit les étapes classiques de consistance et de stabilité. Nous transposons au niveau discret la propriété de stabilité fort-faible des lois de conservations hyperboliques. Nous déterminons l'ordre de convergence des schémas et le quantifions en fonction de la régularité de Sobolev de la donnée initiale. Si nécessaire, nous régularisons la donnée initiale afin de toujours assurer les estimations de consistance. Une étape d'optimisation est alors nécessaire entre cette régularisation et l'ordre de convergence du schéma.

Une seconde partie est consacrée à l'existence d'ondes progressives pour l'équation de Korteweg-de Vries-Kuramoto-Sivashinsky. Par des méthodes classiques de systèmes dynamiques : système augmenté, fonction de Lyapunov, intégrale de Melnikov, par exemple, nous démontrons l'existence d'ondes oscillantes de petite amplitude.

Title : Numerical analysis of hyperbolic-dispersive systems

Keys words : partial differential equations, Korteweg-de Vries equation, finite differences, error estimates, numerical convergence, traveling waves

Abstract : The aim of this thesis is to study some hyperbolic-dispersive partial differential equations. A significant part is devoted to the numerical analysis and more precisely to the convergence of some finite difference schemes for the Korteweg-de Vries equation and $abcd$ systems of Boussinesq. The numerical study follows the classical steps of consistency and stability. The main idea is to transpose at the discrete level the weak-strong stability property for hyperbolic conservation laws. We determine the convergence rate and we quantify it according to the Sobolev regularity of the initial datum. If necessary, we regularize the initial datum for the consistency estimates to be always valid. An optimization step is thus necessary between this regularization and the convergence rate of the scheme.

A second part is devoted to the existence of traveling waves for the Korteweg-de Vries-Kuramoto-Sivashinsky equation. By classical methods of dynamical systems : extended systems, Lyapunov function, Melnikov integral, for instance, we prove the existence of oscillating small amplitude traveling waves.