Finite difference methods for hyperbolic problems with boundaries: stability and multiscale analysis

Benjamin Boutin

COLLEGE DOCTORAL BRETAGNE
MATHS, TELECOMS
INFORMATIQUE, SIGNAL
SYSTEMES, ELECTRONIQUE

Université de Rennes

HABILITATION À DIRIGER DES RECHERCHES

UNIVERSITÉ DE RENNES

ÉCOLE DOCTORALE MATISSE N° 601
Mathématiques, télécommunications, informatique, signal, systèmes, électronique
Spécialité : *Mathématiques*

Benjamin BOUTIN

Finite difference methods for hyperbolic problems with boundaries: stability and multiscale analysis

Unité de recherche : IRMAR – UMR CNRS 6625

Thèse présentée et soutenue à Rennes le 7 juillet 2023

devant le jury composé de

Claire CHAINAIS-HILLAIRET	Professeure, Université de Lille	rapportrice
Bruno DESPRÉS	Professeur, Sorbonne Université	rapporteur
Matthias EHRHARDT	Professor, Bergische Universität Wuppertal	rapporteur
Pauline LAFITTE	Professeure, CentraleSupélec	examinatrice
David LANNES	Directeur de recherche, CNRS, Université de Bordeaux	examinateur
Miguel RODRIGUES	Professeur, Université de Rennes	examinateur
Nicolas SEGUIN	Directeur de recherche, INRIA, Université Côte d'Azur, Montpellier	examinateur

après avis des rapporteurs : Claire CHAINAIS-HILLAIRET, Bruno DESPRÉS, Matthias EHRHARDT.

Remerciements

En tout premier lieu, mes remerciements s'adressent aux rapporteurs de ce manuscrit ainsi qu'aux membres du jury de soutenance. Je leur suis reconnaissant de l'intérêt qu'ils portent à mon travail et leur participation à ce jury m'honore.

Je remercie très chaleureusement tous mes collaborateurs, à commencer par ceux qui, à Paris, m'ont introduit à la recherche mathématique : Christophe Chalons, Frédéric Coquel, Edwige Godlewski, Frédéric Lagoutière et Philippe LeFloch. Quel meilleur souvenir que celui des exposés enthousiastes de Pierre-Arnaud Raviart ? Je remercie également Thierry Goudon et Pauline Lafitte pour leur excellent accueil lors de mon année à Lille. Mes explorations mathématiques ont ensuite été parsemées d'expériences singulières et variées, à l'image des paysages bretons : des rencontres, des retrouvailles, parfois des déceptions mais le plus souvent des petites et des grandes joies dans la curiosité et dans le partage. Merci à Christophe Berthon, Jean-François Coulombel, Nicolas Crouseilles, Denis Michel, Nicolas Raymond et Nicolas Seguin pour nos escapades mathématiques et les travaux stimulants menés au cours des dernières années !

Dans le métier d'enseignant-chercheur, un temps important est directement ou indirectement consacré aux étudiants. Aussi j'ai une pensée pour l'ensemble des étudiants de licence et de master dont j'ai un peu accompagné la route et qui ont incidemment accompagné la mienne. Je remercie plus spécialement Mégane Bournissou avec qui j'ai beaucoup partagé, ainsi que ceux avec lesquels j'ai le plus travaillé : l'infatiguable Thi Hoai Thuong Nguyen, le créatif et serein Pierre Le Barbenchon. Sans nos collaborations et la présence de Nicolas dans leur encadrement, ce travail d'habilitation n'existerait probablement pas.

Je tiens à remercier les directeurs de l'IRMAR et de l'UFR, Mihai et Karel, et tous les collègues responsables de formations, en particulier Jean-Marie et Jean-Christophe. Je remercie aussi l'ensemble des collègues BIATSS qui nous accompagnent au quotidien : Marie-Aude, Annie et Véronique, nos secrétaires en or ; et ceux qui facilitent les aspects matériels et techniques : Eric, Dominique, Maryse et Olivier. Merci aussi à Florian pour la gestion efficace des missions.

Enfin, je remercie tous les collègues et amis de l'IRMAR, de l'UFR et de l'ENS pour les pauses partagées autour de la machine à café, les échanges dans les couloirs, les visites impromptues ou les autres moments hors les murs, plus particulièrement les concerts au 102 ou sous le pommier.

À mes parents et à Mathieu,
c'est un immense plaisir de vivre ce moment en votre présence.

PUBLICATIONS

International refereed journals

[A/BCL+08b] B. Boutin, C. Chalons, F. Lagoutière, and P. G. LeFloch. Convergent and conservative schemes for nonclassical solutions based on kinetic relations. I. *Interfaces Free Bound.* 10 3:399–421, 2008.

[A/BCR10] B. Boutin, C. Chalons, and P.-A. Raviart. Existence result for the coupling problem of two scalar conservation laws with Riemann initial data. *Math. Models Methods Appl. Sci.* 20 10:1859–1898, 2010.

[A/BCL11] B. Boutin, F. Coquel, and P. G. LeFloch. Coupling techniques for nonlinear hyperbolic equations. I. Self-similar diffusion for thin interfaces. *Proc. Roy. Soc. Edinburgh Sect. A*, 141 5:921–956, 2011.

[A/BCL13] B. Boutin, F. Coquel, and P. G. LeFloch. Coupling techniques for nonlinear hyperbolic equations. III. The well-balanced approximation of thick interfaces. *SIAM J. Numer. Anal.* 51 2:1108–1133, 2013.

[A/BBT15] C. Berthon, B. Boutin, and R. Turpault. Shock profiles for the shallow-water Exner model. *Adv. Appl. Math. Mech.* 7 3:267–294, 2015.

[A/BCL15] B. Boutin, F. Coquel, and P. G. LeFloch. Coupling techniques for nonlinear hyperbolic equations. IV. Well-balanced schemes for scalar multi-dimensional and multi-component laws. *Math. Comp.* 84 294:1663–1702, 2015.

[A/MBR16] D. Michel, B. Boutin, and P. Ruelle. The accuracy of biochemical interactions is ensured by endothermic stepwise kinetics. *Prog. Biophys. Mol. Biol.* 121 1:35–44, 2016.

[A/BC17] B. Boutin and J.-F. Coulombel. Stability of finite difference schemes for hyperbolic initial boundary value problems: numerical boundary layers. *Numer. Math. Theory Methods Appl.* 10 3:489–519, 2017.

[A/BR17] B. Boutin and N. Raymond. Some remarks about flows of Hilbert-Schmidt operators. *J. Evol. Equ.* 17 2:805–826, 2017.

[A/BNS20] B. Boutin, T. H. T. Nguyen, and N. Seguin. A stiffly stable semi-discrete scheme for the characteristic linear hyperbolic relaxation with boundary. *ESAIM Math. Model. Numer. Anal.* 54 5:1569–1596, 2020.

[A/FJL+20] G. Flouriot, C. Jehanno, Y. Le Page, P. Le Goff, B. Boutin, and D. Michel. The basal level of gene expression associated with chromatin loosening shapes Waddington landscapes and controls cell differentiation. *J. Mol. Biol.* 432 7:2253–2270, 2020.

[A/BCL21] B. Boutin, F. Coquel, and P. G. LeFloch. Coupling techniques for nonlinear hyperbolic equations. II. Resonant interfaces with internal structure. *Netw. Heterog. Media*, 16 2:283–315, 2021.

[A/BNS+21] B. Boutin, T. H. T. Nguyen, A. Sylla, S. Tran-Tien, and J.-F. Coulombel. High order numerical schemes for transport equations on bounded domains. *ESAIM: ProcS*, 70:84–106, 2021.

[A/BLS23] B. Boutin, P. Le Barbenchon, and N. Seguin. On the stability of totally upwind schemes for the hyperbolic initial boundary value problem. *IMA Journal of Numerical Analysis*, 2023. (In press, doi:10.1093/imanum/drad040).

Manuscript

[PhD/Bou09] B. Boutin. *Étude mathématique et numérique d'équations hyperboliques non-linéaires : couplage de modèles et chocs non classiques.* Université Pierre et Marie Curie - Paris VI, 2009.

Conference proceedings

[C/ABC+08] A. Ambroso, B. Boutin, F. Coquel, E. Godlewski, and P. G. LeFloch. Coupling two scalar conservation laws via Dafermos' self-similar regularization. In, *Numerical Mathematics and Advanced Applications*, pages 209–216. Springer Berlin Heidelberg, 2008.

[C/BCL+08a] B. Boutin, C. Chalons, F. Lagoutière, and P. G. LeFloch. A sharp interface and fully conservative scheme for computing nonclassical shocks. In *Numerical Mathematics and Advanced Applications*, pages 217–224. Springer Berlin, 2008.

[C/BCG08] B. Boutin, F. Coquel, and E. Godlewski. Dafermos regularization for interface coupling of conservation laws. In S. Benzoni-Gavage and D. Serre, editors, *Hyperbolic Problems: Theory, Numerics, Applications*, pages 567–575. Springer Berlin, 2008.

[C/BBF+09] L. Boudin, B. Boutin, B. Fornet, T. Goudon, P. Lafitte, F. Lagoutière, and B. Merlet. Fluid-particles flows: a thin spray model with energy exchanges. In *CEMRACS 2008—Modelling and Numerical Simulation of Complex Fluids*, volume 28 of *ESAIM Proc.* Pages 195–210. EDP Sci., Les Ulis, 2009.

[C/BBC+11] M. Billaud Friess, B. Boutin, F. Caetano, G. Faccanoni, S. Kokh, F. Lagoutière, and L. Navoret. A second order anti-diffusive lagrange-remap scheme for two-component flows. In *CEMRACS'10 Research Achievements: Numerical Modeling of Fusion*, volume 32 of *ESAIM Proc.* Pages 149–162. EDP Sci., Les Ulis, 2011.

[C/BDH+11] B. Boutin, E. Deriaz, P. Hoch, and P. Navaro. Extension of ALE methodology to unstructured conical meshes. In *CEMRACS'10 Research Achievements: Numerical Modeling of Fusion*, volume 32 of *ESAIM Proc.* Pages 31–55. EDP Sci., Les Ulis, 2011.

[C/BCL12] B. Boutin, F. Coquel, and P. G. LeFloch. Coupling techniques for nonlinear hyperbolic equations. In F. Ancona, A. Bressan, P. Marcati, and A. Marson, editors, *HYP2012 Padova*, volume 8 of *Applied Mathematics*, pages 349–356, Italy. American Institute of Mathematical Sciences, 2012.

Submitted preprints

[P/ABC] M. Anandan, B. Boutin, and N. Crouseilles. High order asymptotic preserving scheme for linear kinetic equations with diffusive scaling.

[P/BCC+] B. Boutin, A. Crestetto, N. Crouseilles, and J. Massot. Modified Lawson methods for Vlasov equations.

[P/BLS] B. Boutin, P. Le Barbenchon, and N. Seguin. Stability of finite difference schemes for the hyperbolic initial boundary value problem by winding number computations.

[P/BNS] B. Boutin, T. H. T. Nguyen, and N. Seguin. A stiffly stable fully discrete scheme for the damped wave equation using discrete transparent boundary condition.

The above papers and preprints are sorted and referenced by type, then by ascending dates. Throughout the manuscript, the references to these publications use the corresponding reference keys: [A/rticles], [C/onferences], [P/reprints]. *The other references correspond to the global bibliography found at the end of the manuscript.*

Almost all (pre)publications are on the open archive:
https://cv.hal.science/benjamin-boutin.

Contents

INTRODUCTION (FR)

Panorama de mes travaux

L'étude des effets induits par la présence d'un *bord dans des problèmes d'évolution continus ou discrets* est au cœur de mes travaux de recherche, ceci spécifiquement dans le contexte des systèmes d'équations aux dérivées partielles hyperboliques linéaires, de lois de conservation hyperboliques non-linéaires, ou encore de leurs approximations numériques. Dans ces différentes situations, diverses échelles sont susceptibles d'intervenir à travers les phénomènes de *viscosité*, de *relaxation* ou de *discrétisation*. Ces échelles sont présentes parfois pour des raisons inhérentes à la théorie sous-jacente. C'est le cas pour les solutions faibles entropiques en tant que limites évanescentes d'approximations paraboliques, ou pour les modèles de relaxation faisant intervenir une limite singulière dans des termes d'ordre inférieur. Pour ce qui concerne le cas de méthodes numériques de type volumes finis ou différences finies, les échelles en jeu sont alors directement liées aux paramètres de discrétisation et, parfois simultanément, aux autres échelles concomitamment envisagées. Les interactions entre ces différentes échelles et le bord du domaine sont susceptibles d'engendrer des effets parasites inattendus. Ceux-ci se manifestent typiquement à travers l'apparition de couches limites, nuisant parfois sévèrement aux propriétés de stabilité dans l'asymptotique, et plus souvent dégradant la qualité de l'approximation.

Travaux liés à la thèse de Doctorat et miscellanées

Les travaux décrits ci-après ne seront pas développés précisément dans la suite du manuscrit. Ils concernent en grande partie des recherches directement liées à mes *travaux de thèse* [PhD/Bou09] portant sur les lois de conservation hyperboliques non-linéaires. En premier lieu, la série de papiers [A/BCL11 ; A/BCL13 ; A/BCL15 ; A/BCL21], les actes de congrés [C/ABC$^+$08 ; C/BCG08 ; C/BCL12], ainsi que la publication [A/BCR10] portent sur l'étude du couplage de tels modèles à travers une interface spatiale fixée. Plus précisément, le couplage envisagé présente un caractère *non-conservatif*. Le point de vue diffère ainsi assez fondamentalement du cadre plus standard des lois de conservations à flux discontinus [BV06 ; AKR11 ; And15]. La motivation est au contraire de pouvoir capturer les solutions qui sont entropiques en dehors des interfaces mais possiblement continues à leur traversée malgré la discontinuité du flux. Le point de vue retenu est alors celui du recollement de deux demi-problèmes aux limites. Une première approche emploie les traces admissibles de Dubois et LeFloch [DL88], faisant suite aux travaux de Bardos, Leroux et Nédélec [BLN79]. Une seconde approche s'appuie sur le procédé de régularisation visqueuse à la Dafermos [Daf73] via des estimations d'interaction d'ondes non-linéaires inspirées des travaux de LeFloch et Tzavaras [LT99]. La dernière approche envisagée consiste en une modélisation par interface épaissie et en la mise en place de stratégies numériques inspirées des travaux de Greenberg et Leroux [GL96]

de façon à permettre la préservation d'états stationnaires prescrits, en l'occurence continus à l'interface. L'analyse de convergence de ces méthodes, traitées dans le cas multidimensionnel, requiert les techniques de solutions mesures entropiques de DiPerna [DiP85]. Dans tous ces travaux, les résultats d'existence de solutions sont obtenus. L'unicité n'est pas systématiquement aquise et demeure in fine une question ouverte dans le cas d'interfaces minces. On peut cependant noter que les divers procédés d'approximation envisagés permettent tous de dégager un principe de sélection au moins partielle des solutions, tout du moins en comparaison du cadre le plus général du couplage par des traces admissibles.

Un travail connexe [A/BCL$^+$08b ; C/BCL$^+$08a] concerne le développement d'une nouvelle stratégie numérique permettant le calcul de solutions non-classiques de lois de conservation scalaires. Ces solutions correspondent à des discontinuités sous-compressives liées à une approximation d'ordre supérieur en limite de petite diffusion-dispersion. Alternativement, ces solutions sont caratérisées par une relation cinétique décrivant la dynamique des discontinuités non-classiques [BL02] entre deux phases. La difficulté de l'approximation de ces solutions repose sur le fait que les méthodes usuelles, de façon à être stables, introduisent une diffusion numérique qui nuit à la capture des solutions non-classiques. Une procédure de reconstruction locale permet d'imposer la relation cinétique convenable et de supprimer toute diffusion numérique pour l'approximation numérique des chocs non-classiques.

Dans [A/BBT15], nous nous intéressons à un modèle d'écoulement gravitaire de Saint-Venant-Exner faisant intervenir des produits non-conservatifs. Ces modèles décrivent des écoulements à surface libre, en eaux peu profondes, sur un fond affecté par des effets sédimentaires[1] (dépôt et érosion) dont l'évolution est régie par des lois de comportement plus ou moins empiriques. De façon générale, la définition d'un cadre mathématique adapté aux produits non-conservatifs a été entreprise dès les travaux de Volpert [Vol67] et complétée plus récemment par ceux de Dal Maso, LeFloch et Murat [DLM95]. Les profils de chocs non-conservatifs sont déterminés et comparés avec ceux calculés par différents schémas numériques récents de la littérature.

Les publications [C/BBF$^+$09 ; C/BBC$^+$11 ; C/BDH$^+$11] correspondent à des rapports de projets de recherche effectués au cours de sessions d'été du CEMRACS[2].

Le travail [P/BCC$^+$] soumis plus récemment concerne l'utilisation de méthodes d'intégration en temps de type Lawson adaptées à la résolution de modèles cinétiques de type Vlasov et [P/ABC] concerne la mise en place de méthodes d'ordre élevé, préservant l'asymptotique de diffusion pour des modèles cinétiques linéaires.

Contenu de ce manuscrit d'habilitation

Dans sa majeure partie (Chapitres 1 à 3), les sujets développés dans ce manuscrit concernent plusieurs aspects complémentaires de l'analyse numérique de schémas de différences finies dédiés à l'approximation de solutions de problèmes hyperboliques linéaires en présence de bords.

[1] Travail effectué avec le support du GdR EGRIN 3485 "Modélisation & simulations numériques Ecoulements Gravitaires et RIsques Naturels", désormais GdR MathGeoPhy

[2] Le CEMRACS est un évènement scientifique de la SMAI organisé au CIRM l'été sur 6 semaines, permettant à des jeunes chercheurs de travailler sur des projets de recherche, ceci après une première semaine d'école d'été.

Ces aspects vont de la détermination de propriétés de stabilité vis-à-vis de la condition de bord numérique (Chapitres 1 et 3), à leur utilisation en vue de déterminer des développements asymptotiques des solutions numériques, dont l'utilisation pour une analyse optimale de convergence est précieuse (Chapitre 2). Un autre point de vue, présent à plusieurs reprises dans le manuscrit est celui de l'étude spectrale d'opérateurs en dimension infinie, soit du type Toeplitz ou quasi-Toeplitz (Chapitres 1 et 2), soit du type Hilbert-Schmidt (Chapitre 4). Les Chapitres 4 et 5 sont essentiellement indépendants du reste du manuscrit et emploient des aspects géométriques pour l'asymptotique en temps grand de systèmes dynamiques. Des perspectives de recherches sont présentées à la fin des Chapitres 1, 2 et 3. Une très syntétique présentation des chapitres est la suivante :

- Une introduction à la théorie générale de stabilité pour le problème discret en domaine borné débute le Chapitre 1, introduisant en particulier les outils classiques usuels de cette étude et ouvrant la voie à la présentation des publications [A/BLS23 ; P/BLS] toutes deux issues de la thèse de P. Le Barbenchon [Le 23]. Ces travaux portent sur la mise en place et la justification de méthodes numériques efficaces permettant d'apprécier la validité de la condition de Kreiss-Lopatinskii uniforme dans le cas de schémas de différences finies avec bord.

- Le Chapitre 2 aborde les aspects de consistance au bord, afin de permettre une étude de convergence améliorée pour des schémas assez généraux. La méthode repose sur la construction de développements asymptotiques des solutions numériques, valables dans le domaine intérieur de calcul ainsi qu'au voisinage du bord. L'identification des couches limites numériques est alors centrale. Il est également question d'autres phénomènes multi-échelles qui peuvent être identifiés par une méthodologie analogue. Il s'agit de travaux reliés à la publication [A/BC17] en collaboration avec J.-F. Coulombel, ainsi qu'au travail [A/BNS+21] issu d'un encadrement de projet supporté par l'ANR NABUCO[3] durant le CEMRACS 2019.

- Dans le Chapitre 3, l'étude porte sur la présence conjointe d'effets de termes de relaxation et d'un bord, ainsi que dans un deuxième temps d'effets liés à la discrétisation numérique. L'étude se limite au cas linéaire sous la condition habituelle de Kreiss-Lopatinskii, le terme de relaxation étant soumis aux propriétés usuelles de dissipativité. Dans le cas continu, les développements de couches limites de relaxation et le caractère uniformément bien posé du problème, par rapport au paramètre ϵ, sont connus dans la littérature, en entièrement caractérisés par une condition appelée « Stiff Kreiss Condition »(SKC). Dans le papier [A/BNS20] issu de la thèse [Ngu20] de T. H. T. Nguyen nous démontrons le caractère uniformément stable d'un schéma semi-discrétisé en espace. Ce schéma est constitué sur la base de techniques de sommation par parties discrètes (SBP) utiles pour les méthodes d'énergies et de transformée de Laplace ou en Z selon le cadre (semi-discret ou discret respectivement). Le résultat d'uniforme stabilité du schéma n'est néanmoins obtenu que sur un sous-ensemble strict de la condition SKC.

[3] ANR-17-CE40-0025 Numerical Boundaries and Coupling.

Par une technique de conditions de bord discrète transparente, nous proposons dans [P/BNS] un nouveau schéma discret stable uniformément dans les paramètres de relaxation et de discrétisation, sous la seule condition SKC.

- Le Chapitre 4 porte sur la formalisation et l'analyse de systèmes dynamiques en dimension infinie, à valeurs opérateurs de Hilbert-Schmidt, de la forme de double-crochets. Ce travail correspond à la publication [A/BR17]. Il est en lien d'une part avec la méthode itérative QR pour le calcul du spectre d'une matrice de taille finie, et d'autre part avec les aspects géométriques spécifiques aux flots de crochet. Des résultats de convergence sont obtenus pour la classe de flots envisagée.

- Pour terminer, le travail [A/FJL+20] est décrit dans le Chapitre 5. Il s'agit d'une recherche interdisciplinaire menée avec des biologistes de l'« Institut de Recherche en Santé, Environnement et Travail » (IRSET – Université de Rennes). Les mécanismes de multistabilité liés à la différentiation cellulaire de l'hématopoïèse sont mis en évidence par le calcul de paysages de Waddington obtenus par résolution numérique d'une équation de Fokker-Planck.

INTRODUCTION (EN)

Overview of my work

The study of the effects induced by the presence of a *boundary in continuous or discrete evolution problems* lies at the heart of my research works, more specifically in the context of systems of linear hyperbolic partial differential equations, non-linear hyperbolic conservation laws, or their numerical approximations. In these different situations, various scales are likely to be present through phenomena such as *viscosity, relaxation, discretization.* These scales are sometimes present for reasons that are inherent to the underlying theory. This is the case for entropy weak solutions, being evanescent limits of higher order viscosity approximations, or also for relaxation models in which a singular limit is considered for lower order terms. With regard to the case of numerical methods such as finite volumes or finite differences, the involved scales are then directly related to the discretization parameters and, sometimes simultaneously, to the other scales of the model concomitantly considered. The interactions between these different scales and the boundary of the domain are likely to generate unexpected parasitic effects. These typically manifest themselves through the appearance of boundary layers, sometimes severely impairing the stability properties in the asymptotic process, and more often at least degrading the quality of the approximation.

Work related to the PhD thesis and miscellaneous

The work described below will not be developed in detail in the rest of the manuscript. They largely concern research directly related to my *thesis work* [PhD/Bou09] on nonlinear hyperbolic conservation laws. Firstly, the series of papers [A/BCL11; A/BCL13; A/BCL15; A/BCL21], the conference proceedings [C/ABC$^+$08; C/BCG08; C/BCL12] , as well as the publication [A/BCR10]relate to the study of the coupling of such models through a fixed spatial interface. More precisely, the considered coupling has a *non-conservative* character. The point of view thus differs quite fundamentally from the more standard framework of conservation laws with discontinuous fluxes [BV06; AKR11; And15]. The motivation is on the contrary to be able to capture the solutions which are entropic outside the interfaces but possibly continuous at their crossings despite the discontinuity of the flux. The point of view retained is then that of the gluing together of two boundary half-problems. A first approach uses the admissible traces of Dubois and LeFloch [DL88], following the work of Bardos, Leroux, and Nédélec [BLN79]. A second approach relies on the viscous regularization process a la Dafermos [Daf73] via nonlinear wave interaction estimates inspired by the work of LeFloch and Tzavaras [LT99]. The last approach considered consists of a modeling by thickened interface and the implementation of numerical strategies inspired by the work of Greenberg and Leroux [GL96] in order to allow the preservation of prescribed stationary states, in this case continuous at interface. The convergence analysis of these methods, treated in the multidimensional case,

requires the entropic measurement solution techniques of DiPerna [DiP85]. In all these works, the existence results of solutions are obtained. Uniqueness is not systematically acquired and ultimately remains an open question in the case of thin interfaces. It can however be noted that the various approximation methods considered all make it possible to identify a principle of at least partial selection of the solutions, at least in comparison with the most general framework of coupling by admissible traces.

A related work [A/BCL$^+$08b; C/BCL$^+$08a] concerns the development of a new numerical strategy allowing the computation of non-classical solutions of scalar conservation laws. These solutions correspond to undercompressive discontinuities related to a higher order approximation in the small diffusion-dispersion limit. Alternatively, these solutions are characterized by a kinetic relation describing the dynamics of non-classical discontinuities [BL02] between two phases. The difficulty in approximating these solutions lies in the fact that the usual numerical methods, in order to be stable, introduce a numerical diffusion which harms the capture of non-classical solutions. A local reconstruction procedure makes it possible to impose the appropriate kinetic relation and to suppress any numerical diffusion for the numerical approximation of non-classical shocks.

In [A/BBT15], we are interested in a gravity flow model of Saint-Venant-Exner involving non-conservative products. These models describe free surface flows, in shallow waters, on a bottom affected by sedimentary effects[4] (deposition and erosion) whose evolution is governed by more or less empirical behaviour laws. In general, the definition of a mathematical framework adapted to non-conservative products has been undertaken since the work of Volpert [Vol67] and completed more recently by those of Dal Maso, LeFloch, and Murat [DLM95]. The non-conservative shock profiles are determined and compared with those calculated by various recent numerical schemes from the literature.

The publications [C/BBF$^+$09; C/BBC$^+$11; C/BDH$^+$11] correspond to reports of research projects carried out during summer sessions of CEMRACS[5].

The work [P/BCC$^+$] submitted more recently concerns the use of modified Lawson-type time integration methods suitable for solving Vlasov-type kinetic models and [P/ABC] concerns the development of high order methods that preserve the diffusion asymptotic for linear kinetic models.

Contents of this habilitation manuscript

In its major part (Chapters 1 to 3), the subjects developed in this manuscript concern several complementary aspects of the numerical analysis of finite difference schemes dedicated to the approximation of solutions of linear hyperbolic problems in the presence of boundaries. These aspects range from the determination of stability properties with respect to the numerical boundary condition (Chapters 1 and 3), to their use in order to determine asymptotic expansions of numerical solutions, whose use for an optimal convergence analysis is valuable

[4]This work has been partially supported by GdR EGRIN 3485 "Modeling & numerical simulations Gravitary Flows and Natural RIsks", now GdR MathGeoPhy

[5]CEMRACS is a scientific event of SMAI organized at CIRM in the summer during 6 weeks, allowing young researchers to work on research projects, after a first week of summer school.

(Chapter 2). Another point of view, present several times in the manuscript, is that of the spectral study of operators in infinite dimension, either of the Toeplitz or quasi-Toeplitz type (Chapters 1 and 2), or of the Hilbert-Schmidt type (Chapter 4). Chapters 4 and 5 are essentially independent from the rest of the manuscript and employ geometric aspects for the large-time asymptotic of dynamical systems. Research perspectives are presented at the end of Chapters 1, 2 and 3. A very synthetic presentation of the chapters is as follows:

- An introduction to the general theory of stability for the discrete problem in a bounded domain begins Chapter 1, introducing in particular the usual classical tools of this study and opening the way to the presentation of the publications [A/BLS23; P/BLS] both from the thesis of P. Le Barbenchon [Le 23]. This work focuses on the establishment and justification of effective numerical methods for assessing the validity of the uniform Kreiss-Lopatinskii condition in the case of finite difference schemes with a boundary.

- Chapter 2 deals with boundary consistency aspects, in order to allow an improved convergence study for fairly general schemes. The method is based on the construction of asymptotic expansions of the numerical solutions, valid in the inner computational domain as well as in the vicinity of the boundary. The identification of discrete boundary layers is then central. It also discusses other multi-scale phenomena that can be identified from an analogous methodology. These are works related to the publication [A/BC17] in collaboration with J.-F. Coulombel, as well as to the work [A/BNS$^+$21] resulting from a project supervision during CEMRACS 2019, supported by the funding project ANR NABUCO[6].

- In Chapter 3, the study focuses on the joint presence of effects of relaxation terms and of a boundary, as well as in a second time of effects related to numerical discretization. The study is limited to the linear case under the usual Kreiss-Lopatinskii condition, the relaxation term being subject to the usual dissipative properties. In the continuous case, the expansions of relaxation boundary layers and the uniformly well-posed character of the problem, with respect to the parameter ϵ, are known in the literature, fully characterized by a condition called "Stiff Kreiss Condition"(SKC). In the paper [A/BNS20] resulting from the thesis [Ngu20] of T. H. T. Nguyen we demonstrate the uniformly stable character of a semi-discretized scheme in space. The scheme is made up on the basis of techniques of discrete summation by parts (SBP), useful for the energy method, and of Laplace or Z transform according to the framework (semi-discrete or discrete respectively). The result of uniform stability of the scheme is nevertheless only obtained on a strict subset of the SKC condition.
  By a technique of discrete transparent boundary condition, we propose in [P/BNS] a new fully discrete scheme, stable uniformly in the relaxation and discretization parameters, under the only condition SKC.

- Chapter 4 deals with the formalization and analysis of dynamical systems in infinite dimension, with values in the set of Hilbert-Schmidt operators, and having a double-

[6] ANR-17-CE40-0025 Numerical Boundaries and Coupling.

brackets structure. This work corresponds to the publication [A/BR17]. It is concerned the one hand with to the iterative QR method for computing the spectrum of a matrix in finite dimensional spaces, and on the other hand with the geometrical aspects specific to bracket flows. Convergence results are obtained for the considered class of flows.

- Finally, the work [A/FJL+20] is described in Chapter 5. This is an interdisciplinary research carried out with biologists from the "Institute for Research in Health, Environment and Work" (IRSET – University of Rennes). The mechanisms of multistability linked to the cellular differentiation of haematopoiesis are highlighted by the calculation of Waddington landscapes obtained by numerical resolution of a Fokker-Planck equation.

CHAPTER 1

Discrete initial boundary value problems

The present chapter is intended to discuss the stability theories for initial boundary value problems (IBVP), on one hand for first order linear hyperbolic partial differential evolution equations, and on the other hand for linear finite difference schemes. The well-posedness or the stability properties of the corresponding IBVPs are essential and the main issue is to fully characterize these properties by means of concise (e.g. algebraic) conditions. More detailed presentations and developments of the general theory can be found for example in the books by Benzoni-Gavage and Serre [BS07] for the continuous problem, and in the book by Gustafsson, Kreiss, and Oliger [GKO13] and the lecture notes by Coulombel [Cou13] for discrete schemes. We restrict the following presentation to the main lines of these two theories. After that, we will present an overview of some contributions to the numerical study of Kreiss-Lopatinskii determinants, with applications to the strong (GKS-) stability properties for the discrete IBVP with commonly used boundary conditions. The presented results are mainly based on the publications [A/BLS23; P/BLS] recalled hereunder, related to the PhD work of Pierre Le Barbenchon within the years 2020–2023. Most of the illustrations can be reproduced by using the Python library "boundaryscheme" [LN23] (see doi:10.5281/zenodo.7773741) developed by P. Le Barbenchon.

[A/BLS23] B. Boutin, P. Le Barbenchon, and N. Seguin. On the stability of totally upwind schemes for the hyperbolic initial boundary value problem. *IMA Journal of Numerical Analysis*, 2023. (In press, doi:10.1093/imanum/drad040).

[P/BLS] B. Boutin, P. Le Barbenchon, and N. Seguin. Stability of finite difference schemes for the hyperbolic initial boundary value problem by winding number computations.

1.1 Well-posedness theory for linear hyperbolic problems

Main lines The study of first order linear hyperbolic IBVPs is based on two theories. The first one is the theory of Friedrichs [FL67] specifically adapted to symmetric systems with maximal dissipative boundary conditions. From integration by parts and using the dissipativity property at the boundary, some energy estimates are available so as to deduce then the required a priori estimates. These estimates encompass the interior norm of the solution as well as its trace at the boundary. The second theory is from Kreiss [Kre70]. It handles with more general problems (lack of symmetry, of symmetrizability, of maximal dissipativity at the boundary) and is based on the construction of frequency-dependent dissipative symmetrizers also known as Kreiss symmetrizers. As a counterpart, the a priori estimates first consist in resolvent estimates and concern zero initial data only. A more technical part for closing the well-posedness theory towards semigroup estimates then rely on a causality argument, duality formulation, and the Gårding and Leray method. It is due to Rauch [Rau72] and we omit here any discussion in that direction.

Notations and setting For convenience and tightness in the presentation, the physical geometry under consideration is here mostly restricted to the simple case of a straight half-space, namely $x = (y, x_d) \in \mathbb{R}^d_+ := \mathbb{R}^{d-1} \times (0, +\infty)$ and the unknown is $u(x,t) \in \mathbb{R}^N$. The evolution operator and the boundary conditions are linear with constant coefficients, namely the *continuous IBVP* has the form

$$\begin{aligned} \mathsf{L}u &= F, && (x,t) \in \mathbb{R}^d_+ \times \mathbb{R}_+, \\ \mathsf{B}u_{|x_d=0} &= g, && (y,t) \in \mathbb{R}^{d-1} \times \mathbb{R}_+, \\ u_{|t=0} &= f, && x \in \mathbb{R}^d_+. \end{aligned} \tag{1.1}$$

where the differential operator L is given by

$$\mathsf{L} = \partial_t + \sum_{j=1}^{d} \mathsf{A}_j \partial_{x_j}, \tag{1.2}$$

$\mathsf{B} \in \mathcal{M}_{m,N}(\mathbb{R})$ being a full rank matrix: $\mathsf{rank}\,\mathsf{B} = m \leq N$ and, for $1 \leq j \leq d$, $\mathsf{A}_j \in \mathcal{M}_N(\mathbb{R})$ being real valued matrices. The data are F, g and f, respectively an interior source term, the boundary data and the initial data, in L^2 spaces of their variables.

Hyperbolicity A first important necessary condition for the well-posedness is related to the time-hyperbolicity of the operator L. In the sequel, we simply talk about *hyperbolicity* since the time is clearly the principal direction for evolution problems. The hyperbolicity property is addressed to the Cauchy problem only (without space boundaries). It is fully characterized from the Fourier space symbol function $\mathcal{A}$ defined for frequencies $\xi \in \mathbb{R}^d$ by $\mathcal{A}(\xi) = \sum_{j=1}^d i\xi_j \mathsf{A}_j$. More precisely the well-posedness of the Cauchy problem associated to the operator L in appropriate function spaces is then equivalent to the uniform power boundedness of the family of matrices $\mathcal{A}(\xi)$ for $\xi \in \mathbb{R}^d$ or equivalently (by an appropriate normalization)

for ξ such that $|\xi| = 1$. Among the various subfamilies of hyperbolic problems, let us now recall the principal ones. The operator L is said to be

- *weakly hyperbolic* if for all $\xi \in \mathbb{R}^d$, $|\xi| = 1$, the matrix $\mathcal{A}(\xi)$ has spectrum inside $i\mathbb{R}$;
- *strongly hyperbolic* or shortly **hyperbolic**, if in addition the matrices $\mathcal{A}(\xi)$ are uniformly diagonalizable;
- *semi-strictly hyperbolic* if in addition the eigenvalues of $\mathcal{A}(\xi)$ have constant multiplicities;
- *strictly hyperbolic* if in addition the eigenvalues of $\mathcal{A}(\xi)$ have multiplicity one;
- *symmetric hyperbolic* if $\mathcal{A}(\xi)$ is skew-hermitian;
- *symmetrizable hyperbolic* if there exists a hermitian positive definite matrix S such that $S\mathcal{A}(\xi)$ is skew-hermitian.

It has to be noticed that the weakest notion of hyperbolicity above is related to a stability property in the Hadamard sense only, namely with possible loss of derivatives in semigroup estimates. It is not as robust as are the others and in particular does not directly enable the possible treatment of supplementary lower order perturbation terms, useful to cover the quasi-linear or the nonlinear case. On the contrary Strang [Str67] has shown that strongly hyperbolic problems (with no loss of derivatives) are stable by zeroth-order perturbations. The same issue occurs when dealing with the IBVP and requires an adapted form of stability with respect to the boundary data.

Kreiss-Lopatinskii conditions In the *one-dimensional* case $d = 1$, many of the previously mentioned hyperbolicity notions coincide, due to the scalar form of the frequency parameter ξ. The strong, semi-strict and symmetrizable hyperbolicity are then equivalent to the requirement that the matrix $\mathsf{A} = \mathsf{A}_1$ has real eigenvalues and a complete set of eigenvectors. This is true if A is real symmetric. If the boundary is *non-characteristic* in the sense that $\mathsf{Ker}\,\mathsf{A} = \{0\}$, then we may decompose the full space $\mathbb{R}^N$ according to the characteristic fields of A into

$$\mathbb{R}^N = E_+(\mathsf{A}) \oplus E_-(\mathsf{A}). \tag{1.3}$$

Here the space $E_+(\mathsf{A})$ (respectively $E_-(\mathsf{A})$) is defined from the eigenprojection of A associated to the $p_+ = \dim E_+(\mathsf{A})$ positive (respectively p_- negative) eigenvalues of A, with multiplicities. They correspond physically to rightgoing (respectively leftgoing) scalar waves in the time-evolution problem. The algebraic solving of the boundary equation $\mathsf{B}u_{|x=0} = g$ then corresponds, after decomposing $u_{|x=0} = u_+ + u_-$ matching (1.3), to the solving of the equation $\mathsf{B}u_+ = g - \mathsf{B}u_-$. Therefore, inverting the matrix B acting from $E_+(\mathsf{A})$ to $\mathbb{R}^m$ is mandatory. To that aim, two complementary properties are required: $\mathsf{rank}\,\mathsf{B} \geq p_+$ for the uniqueness; $\mathbb{R}^m \subset \mathsf{B}E_+(\mathsf{A})$ for the existence. Finally the convenient structural property is the following necessary condition:

One-dimensional Kreiss-Lopatinskii condition:

$$\mathbb{R}^N = \mathsf{Ker}\,\mathsf{B} \oplus E_+(\mathsf{A}).$$

Under the strong hyperbolicity and the Kreiss-Lopatinskii condition, the continuous one-dimensional IBVP admits a full well-posedness setting.

The treatment of *multidimensional* problems is done somehow similarly through modal Laplace-Fourier analysis, based first on the fundamental *dispersion relation*

$$\det \mathcal{L}(\tau, \xi) = 0, \tag{1.4}$$

where we set $\mathcal{L}(\tau, \xi) := \tau \mathsf{Id} + \mathcal{A}(\xi)$. Here $\tau \in \mathbb{C}$ is related to the time-Laplace dual parameter and $\xi \in \mathbb{R}^d$ to the space-Fourier dual variable. The zero set of the dispersion relation gathers the frequencies (τ, ξ) associated to nonzero modal solutions, that is of the form $u(t, x) = \mathsf{e}^{\tau t + ix\cdot\xi}\varphi$ for some $\varphi \in \mathbb{C}^N \setminus \{0\}$. The solving of the boundary equations is done by using the tangential part of the differential operator L along the space boundary $\mathbb{R}^{d-1} \times \{0\}$. A partial space-Fourier transform is helpful to algebraize the tangential variable $y \in \mathbb{R}^{d-1}$ with dual frequency variable $\eta \in \mathbb{R}^{d-1}$. The symbol for the normal problem at the boundary reads

$$\mathcal{G}(\tau, \eta) = -\mathsf{A}_d^{-1}(\tau \mathsf{Id} + \mathcal{A}_0(\eta)) \tag{1.5}$$

where we set $\mathcal{A}_0(\eta) = \sum_{j=1}^{d-1} i\eta_j \mathsf{A}_j$. The solutions of the form $u(t, x) = \mathsf{e}^{\tau t + iy\cdot\eta}\varphi(x_d)$ are then associated to functions $\varphi(x_d) = \mathsf{exp}(x_d\, \mathcal{G}(\tau, \eta))\varphi(0)$ for $\varphi(0) \in \mathbb{C}^N$. Similarly to the one-dimensional case, the boundary condition $\mathsf{B}u_{|x_d=0} = g$ is intended to prescribe in a unique way the convenient value for $\varphi(0) \in \mathbb{C}^N$ so that the solution φ takes value in $L^2(\mathbb{R}_+)$ and depends continuously on the boundary data g in appropriate function spaces. An important structural result comes from the hyperbolicity property itself (thus from the Cauchy well-posedness). It consists in a separation property for the positive and negative eigenspaces of the matrices $\mathcal{G}(\tau, \eta)$. The result is due to Hersh and appears as a natural generalization of the previous decomposition (1.3) to frequency-parameterized cases.

Lemma 1 (Separation [Her63])**.** *Assume the operator* L *to be hyperbolic and the boundary to be non-characteristic in the sense that* $\mathsf{Ker}\, \mathsf{A}_d \neq \{0\}$*. Then* $\mathbb{C}^N = E_+(\mathsf{A}_d) \oplus E_-(\mathsf{A}_d)$*. Moreover for any* $(\tau, \eta) \in \mathbb{C} \times \mathbb{R}^{d-1}$ *with* $\mathsf{Re}\, \tau > 0$*:*

$$\mathbb{C}^N = \mathbb{E}^{\mathsf{s}}(\tau, \eta) \oplus \mathbb{E}^{\mathsf{u}}(\tau, \eta),$$

with in addition $\dim \mathbb{E}^{\mathsf{s}}(\tau, \eta) = \dim E_+(\mathsf{A}_d)$ *and* $\dim \mathbb{E}^{\mathsf{u}}(\tau, \eta) = \dim E_-(\mathsf{A}_d)$*.*

Here the stable space $\mathbb{E}^{\mathsf{s}}(\tau, \eta)$ (respectively the unstable space $\mathbb{E}^{\mathsf{u}}(\tau, \eta)$) is defined as the sum of the eigenspaces of $\mathcal{G}(\tau, \eta)$ associated to eigenvalues with negative (respectively positive) real parts. These spaces depend on the frequency parameter as positively homogeneous functions with degree 0, thus it is interesting to introduce the set of normalized frequency parameters $\Sigma := \{(\tau, \eta) \in \mathbb{C} \times \mathbb{R}^{d-1},\ \mathsf{Re}\, \tau > 0,\ |\tau|^2 + |\eta|^2 = 1\}$ and its boundary $\Sigma_0 := \{(\tau, \eta) \in \mathbb{C} \times \mathbb{R}^{d-1},\ \mathsf{Re}\, \tau = 0,\ |\tau|^2 + |\eta|^2 = 1\}$. To preclude the existence of time-unstable solutions with finite space energies, the following condition is required:

Kreiss-Lopatinskii condition:

$$\forall (\tau, \eta) \in \Sigma, \quad \mathsf{Ker}\, \mathsf{B} \cap \mathbb{E}^{\mathsf{s}}(\tau, \eta) = \{0\}.$$

The condition admits several equivalent formulations (nonzero determinant, resolvent inequality) but it is important to observe that the stability for the IBVP is generally obtained only under the following reinforced condition, where $|\cdot|$ denotes here any finite-dimensional norm:

Uniform Kreiss-Lopatinskii Condition (UKLC):

$$\mathsf{rank}\,\mathsf{B} = \dim E_+(\mathsf{A}_d)$$
$$\exists C>0,\ \forall(\tau,\eta)\in\Sigma,\ \forall\varphi\in\mathbb{E}^{\mathsf{s}}(\tau,\eta),\ |\mathsf{B}\varphi|\geq C|\varphi|.$$

For many hyperbolic problems such as strictly or semi-strictly hyperbolic it is known that the spaces $\mathbb{E}^{\mathsf{s}}(\tau,\eta)$ and $\mathbb{E}^{\mathsf{u}}(\tau,\eta)$ depend analytically on the parameters $(\tau,\eta)\in\Sigma$ and admit a continuous extension to Σ_0 (see Métivier [Mét04]). In particular the UKLC can be formulated by means of a determinant constructed from an orthonormal basis $\{e_1,e_2,\cdots,e_r\}$ of $\mathbb{E}^{\mathsf{s}}(\tau,\eta)$ as $\Delta(\tau,\eta):=\mathsf{det}\,(\mathsf{B}e_1(\tau,\eta),\mathsf{B}e_2(\tau,\eta),\cdots,\mathsf{B}e_r(\tau,\eta))$.

Reformulation by the UKLC determinant:

$$\exists\delta>0,\ \forall(\tau,\eta)\in\overline{\Sigma},\ |\Delta(\tau,\eta)|\geq\delta.$$

A priori estimates and general results Equipped with these tools, the general theory for the well-posedness is based on several other notions: discrete block structure condition, Kreiss symmetrizers, continuous extension of the formulation, UKLC resolvent reformulation and finally energy estimates adapted to that reformulation. We do not detail these aspects here and refer the interested reader to the book by [BS07] or to the work by Métivier [Mét17].

Let us sketch some first steps, that will be important in the forthcoming similar stability analysis for discrete IBVPs. Even if the existence of a solution with traces along the boundary is an important part of the problem, assume it and let us only discuss the useful a priori estimates (actually they participate to prove the existence by a duality argument). The time-Laplace transformed version of (1.1) is considered for zero initial data f and without source term F. Under the UKLC, from the Laplace-Parseval identity and the invertibility of B on $\mathbb{E}^{\mathsf{s}}$, the following estimate follows for all $\gamma>0$:

$$\int_0^\infty \mathsf{e}^{-2\gamma t}|u(t,\eta,0)|^2\mathsf{d}t \quad\lesssim\quad \int_0^\infty \mathsf{e}^{-2\gamma t}|\mathsf{B}u(t,\eta,0)|^2\mathsf{d}t. \tag{1.6}$$

Here and everywhere after, identities of the form $X\lesssim Y$ mean that there exists a constant $C>0$ independent of X, Y and the other parameters in the formula (including γ here), such that $X\leq CY$. In the following lines, we also make use of the following **notations of norms** to shorten and to alleviate the reading of exponential weights:

$$\|u(t,\cdot)\|^2_{\mathbb{R}^d_+}:=\int_{\mathbb{R}^d_+}|u(t,x)|^2\mathsf{d}x,\qquad \|u(t,\cdot,0)\|^2_{\mathbb{R}^{d-1}}:=\int_{\mathbb{R}^{d-1}}|u(t,y,0)|^2\mathsf{d}y,$$
$$\|u\|^2_\gamma:=\int_0^{+\infty}\mathsf{e}^{-2\gamma t}\|u(t,\cdot)\|^2_{\mathbb{R}^d_+}\mathsf{d}t,\qquad |u|^2_\gamma:=\int_0^{+\infty}\mathsf{e}^{-2\gamma t}\|u(t,\cdot,0)\|^2_{\mathbb{R}^{d-1}}\mathsf{d}t.$$

The previous estimate (1.6) on traces thus also reads $|u|^2_\gamma\lesssim|\mathsf{B}u|^2_\gamma$. The general theorem to have in mind afterwards is the following one. The estimates also encompass interior norms, source terms and even non-zero initial data.

Theorem 2 (Benzoni-Gavage and Serre [BS07]). *Let us consider a strictly or a semi-strictly hyperbolic system (1.1) with a non-characteristic boundary condition.*
The Uniform Kreiss-Lopatinskii Condition is then equivalent to the strong stability property

$$\gamma\|u\|_\gamma^2 + |u|_\gamma^2 \quad \lesssim \quad |\mathsf{B}u|_\gamma^2 + \frac{1}{\gamma}\|\mathsf{L}u\|_\gamma^2. \tag{1.7}$$

If satisfied, then the PDE is strongly well-posed in L^2: for any data f, g, F in L^2 spaces, there exists a unique solution u in L^2 space:

$$\sup_{t>0}\left(\mathrm{e}^{-2\gamma t}\|u(t,\cdot)\|^2\right) + \gamma\|u\|_\gamma^2 + |u|_\gamma^2 \quad \lesssim \quad \|f\|^2 + |g|_\gamma^2 + \frac{1}{\gamma}\|F\|_\gamma^2. \tag{1.8}$$

1.2 Stability theory for linear finite difference schemes

The stability theory for finite difference methods dates back first to 1928 and the work by Courant, Friedrichs, and Lewy [CFL28] where the famous eponym condition is introduced, based on the necessary inclusion property of the theoretical dependency domain into the numerical one. Later on, from Crank and Nicolson [CN47] and Charney, Fjörtoft, and von Neumann [CFvN50], the helpful Fourier-spectral condition facilitates the stability analysis for linear methods set over domains without boundaries (periodic or infinite).

Notations and setting In order to highlight the similarities with the well-posedness theory previously discussed for PDEs, we introduce now a quite abstract notation and consider *discrete scalar IBVPs* on the quarter plane put under the form

$$\begin{aligned} (\mathsf{L}u)_j^n &= F_j^n, && (j,n) \in \mathbb{N}\times\mathbb{N}, \\ (\mathsf{B}u)_j^n &= g^n, && (j,n) \in \{0\}\times\mathbb{N}, \\ (\mathsf{I}u)_j^n &= f_j, && (j,n) \in \mathbb{N}\times\{0\}. \end{aligned} \tag{1.9}$$

The unknown $u = (u_j^n)$ could be vector-valued but we restrict here to scalar values only. The finite difference operator L is given by a finite constant coefficients linear combination of powers of the time-translation $\mathsf{T} : u^n \mapsto u^{n+1}$ and of the space-shift $\mathsf{S} : u_j \mapsto u_{j+1}$. Actually the operators S in defined on the whole set of sequences indexed by $\mathbb{Z}$ when we discuss the problem without boundaries (or by periodicity on periodic in space domains). The interior scheme is

$$\mathsf{L} = \sum_{\sigma=0}^{k}\sum_{\ell=-r}^{p} a_{\sigma,\ell}\mathsf{T}^\sigma\mathsf{S}^\ell. \tag{1.10}$$

The operator $\mathsf{I} = (\mathsf{Id}, T, \dots, T^{k-1})$ is used to define the k first time steps from the multistep initial data f. Actually, for a scheme having r left points, the space-shift S is defined from only the indices $\{-r, -r+1, \dots, -1\} \cup \mathbb{N}$ in the following natural way: $\mathsf{S}(u_{-r}, u_{-r+1}, \dots) := (u_{-r+1}, u_{-r+2}, \dots)$ and $\mathsf{S}^{-1}(u_{-r}, u_{-r+1}, \dots) := (0, u_{-r}, , u_{-r+1}, \dots)$. The boundary scheme operator $\mathsf{B} = (\mathsf{B}_1, \dots, \mathsf{B}_r)^T$ has to provide the r ghost cells required from L at each time step.

We consider here simply a time-independent local form, with for $1 \leq q \leq r$:

$$\mathsf{B}_q = \sum_{\ell=-r}^{m} b_{q,\ell} \mathsf{S}^\ell, \tag{1.11}$$

Both the full Cauchy problem associated to (1.9) (without boundary) and the IBVP have to be solvable, what may require sometimes complicated algebraic properties in the present general setting, but is quite straightforward for a given usual scheme. The corresponding solvability assumption is that, being given some data f, g and F, the existence and uniqueness of a solution u to (1.9) is guaranteed.

As an example, let us consider the Beam-Warming scheme for solving (with positive velocity $a > 0$) the advection problem $\partial_t u + a\partial_x u = 0$ together with discrete Dirichlet boundary conditions $u^n_{-2} = u^n_{-1} = g^n$. Setting $\lambda = a\Delta t/\Delta x$, the scheme corresponds to the operators

$$\begin{aligned}
\mathsf{L} &= \frac{1}{\Delta t}\left(\mathsf{T} - \tfrac{\lambda(\lambda-1)}{2}\mathsf{S}^{-2} - \lambda(2-\lambda)\mathsf{S}^{-1} - \tfrac{(2-\lambda)(1-\lambda)}{2}\mathsf{S}^0\right), \\
\mathsf{B} &= (\mathsf{S}^{-2}, \mathsf{S}^{-1})^T.
\end{aligned}$$

Fourier Von Neumann symbolic analysis In between the most general case (1.9) and the previous simple example, we introduce for future purpose the explicit linear multistep methods based on the method-of-lines. These family of schemes will again be used later in the Chapter 2. The space discretization is first done and, after that, a scalar time-integrator is used:

$$\frac{1}{\Delta t}\sum_{\sigma=0}^{k} \alpha_\sigma u_j^{n+\sigma} + \frac{1}{\Delta x}\sum_{\sigma=0}^{k-1} \beta_\sigma \sum_{\ell=-r}^{p} a_l u_{j+\ell}^{n+\sigma} = 0, \tag{1.12}$$

or

$$\mathsf{L} = \frac{1}{\Delta t}\sum_{\sigma=0}^{k} \alpha_\sigma \mathsf{T}^\sigma + \frac{1}{\Delta x}\sum_{\sigma=0}^{k-1} \beta_\sigma \mathsf{T}^\sigma \sum_{\ell=-r}^{p} a_l \mathsf{S}^\ell.$$

Formally, the *discrete dispersion relation* is obtained by considering solutions of the form $u_j^n = z^n \kappa^j$, or by replacing symbolically T by z and S by κ. In the scalar-unknown setting, it reads

$$\mathcal{P}_{\mathrm{LMM}}(z,\kappa) := \frac{1}{\Delta t}\rho_1(z) + \frac{1}{\Delta x}\rho_0(z)\mathcal{A}(\kappa) = 0. \tag{1.13}$$

Here we denote the Dahlquist polynomial of the time linear recurrence $\rho_1(z) = \sum_{\sigma=0}^{k} \alpha_\sigma z^\sigma$ and $\rho_0(z) = \sum_{\sigma=0}^{k-1} \beta_\sigma z^\sigma$, and the Fourier symbol for the space discrete operator $\mathcal{A}(\kappa) = \sum_{\ell=-r}^{p} a_\ell \kappa^\ell$. With these notations, the consistency property with $\partial_t u + a\partial_x u = 0$ is related to tangency properties at $(z,\kappa) = (1,1)$: $\rho_1(1) = \mathcal{A}(1) = 0$, $\rho_1'(1) = \rho_0(1) = 1$ and $\mathcal{A}'(1) = a$. In addition, the underlying Cauchy ℓ^2-stability for both the periodic case or the infinite ($j \in \mathbb{Z}$) domain is then fully determined by the so-called *root condition*.

Root condition:

$$\forall \kappa \in \mathbb{S}^1,\ \forall z \in \mathbb{C},\ \left[\rho_1(z) + \tfrac{\Delta t}{\Delta x}\rho_0(z)\mathcal{A}(\kappa) = 0 \Rightarrow |z| < 1 \text{ or } |z| = 1 \text{ is simple root}\right].$$

This condition follows from the power boundedness property of the companion matrices

associated to the involved time recurrence. The stability domains of classical time integration method are known (see Hairer and Wanner [HW96]) so that the stability of a given space discretization is most of the time either possible for Δt sufficiently small, or simply impossible.

Always for the scalar multistep scheme (1.12) with initial data f and source term $F = 0$ (in the sense of (1.9) but now for $j \in \mathbb{Z}$), the root condition is equivalent to the validity of the following stability estimate:

$$\sup_{n\geq 0} \sum_{j\in\mathbb{Z}} |u_j^n|^2 \Delta x \quad \lesssim \quad \sum_{j\in\mathbb{Z}} |f_j|^2 \Delta x.$$

Normal mode analysis and strong stability In order to introduce now the boundary problem, the following norms are used, that depend on $\gamma > 0$ and on Δt and Δx in a consistent way, being compared to the continuous case:

$$\|u\|_\gamma^2 := \sum_{n\geq 0} \sum_{j=-r}^{+\infty} \mathsf{e}^{-2\gamma n\Delta t} |u_j^n|^2 \Delta x \Delta t, \quad |u|_\gamma^2 := \sum_{n\geq 0} \sum_{j=-r}^{p} \mathsf{e}^{-2\gamma n\Delta t} |u_j^n|^2 \Delta t.$$

Similarly to the continuous case and (1.6), the first step is to estimate the boundary values of the discrete solution from the set of boundary data:

$$|u|_\gamma^2 \quad \lesssim \quad |\mathsf{B}u|_\gamma^2 := \sum_{n\geq 0} \sum_{q=1}^{r} \mathsf{e}^{-2\gamma n\Delta t} |(\mathsf{B}_q u)_0^n|^2 \Delta t, \tag{1.14}$$

The two seminal papers for the boundary stability theory for finite difference schemes are the ones by Kreiss [Kre68] and Gustafsson, Kreiss, and Sundström [GKS72]. The definition to consider is the following.

Definition 3 (Gustafsson, Kreiss, and Sundström [GKS72])**.** *The numerical scheme* (1.9) *is* strongly (GKS-) stable *if any solution u_j^n with zero initial data $f = 0$ satisfies the following estimate, independent of $\gamma > 0$ and $\Delta t \in (0, 1]$:*

$$\frac{\gamma}{1+\gamma\Delta t} \|u\|_\gamma^2 + |u|_\gamma^2 \quad \lesssim \quad |\mathsf{B}u|_\gamma^2 + \frac{1+\gamma\Delta t}{\gamma} \|\mathsf{L}u\|_\gamma^2. \tag{1.15}$$

A few years before these results, from the *Moscou-Novossibirsk* research group has emerged in 1956 the Babenko-Gelfand procedure for analysing the boundary stability for difference equations, and directly from there the necessary condition from Godunov and Rjabenkii [GR63]. This is exactly the discrete counterpart of the Kreiss-Lopatinskii *necessary condition* and makes use of an adapted stable space $\mathbb{E}_{\mathsf{s}}(z)$ for the discrete setting. Actually, under the Cauchy ℓ^2-stability of the scheme, the following separation result is available.

Lemma 4 (Root splitting)**.** *Assume the scheme* (1.12) *to be Cauchy ℓ^2-stable. Then for all $z \in \mathcal{U}$, the relation dispersion $\mathcal{P}(z, \cdot)$ has exactly r roots in $\mathbb{D} := \{\kappa,\ |\kappa| < 1\}$ and p roots in $\mathcal{U} := \{\kappa,\ |\kappa| > 1\}$ (with multiplicities).*

From this result, this is possible to solve the z-transformed version of the boundary problem with zero initial data f (and zero source terms F here for convenience). We recall that the

z-transform is defined for $|z| > 1$ and a sequence $(u^n)_{n\geq 0}$ with $u^0 = 0$, by the following series transform $\hat{u}(z) := \sum_{n\geq 0} z^{-n} u^n$. From there the solution to (1.9)-(1.10) satisfies:

$$\begin{aligned} &\sum_{\ell=-r}^{p} \left(\sum_{\sigma=0}^{k} a_{\sigma,\ell} z^{\sigma} \right) \hat{u}_{j+\ell}(z) = 0, \quad j \geq 0, \\ &(\mathsf{B}\hat{u}(z))_0 = \hat{g}(z). \end{aligned} \tag{1.16}$$

The homogeneous space linear recurrence relation for $(\hat{u}_j(z))_{j\geq 0}$ has order $r + p$ (under a non-degeneracy for the extreme terms that is known as non-characteristic property in the literature) and therefore the linear space $\mathbb{E}(z)$ of their solutions has dimension $r + p$. In detail the space $\mathbb{E}_{\mathsf{s}}(z)$, defined as the linear subspace of $\mathbb{E}(z)$ with zero limit at infinity (so in $\ell^2(\{-r, \dots, -1\} \cup \mathbb{N})$), has dimension r exactly. Similarly, the linear subspace $\mathbb{E}_{\mathsf{u}}(z)$ of solutions with zero limit at $-\infty$ has dimension p. Finally the previous root splitting result leads to the following decomposition:

$$\forall z \in \mathcal{U}, \quad \mathbb{E}(z) = \mathbb{E}_{\mathsf{s}}(z) \oplus \mathbb{E}_{\mathsf{u}}(z).$$

We can now state hereafter an adapted statement with notations similar to the PDE case. Let us observe that there is a slight abuse in the notation since we denote $\mathsf{Ker}\,\mathsf{B}$ instead of $\mathsf{Ker}\,(\mathsf{B}\,\cdot)_0$.

Proposition 5 (Godunov-Rjabenkii necessary condition)**.** *If there exists a complex value z with $|z| > 1$, such that $\mathsf{Ker}\,\mathsf{B} \cap \mathbb{E}_{\mathsf{s}}(z) \neq \{0\}$, then the discrete IBVP (1.9) is not strongly stable.*

In short, this condition excludes the existence of a solution to the homogeneous version of (1.16) with value in $\ell^2(\mathbb{N})$. As for the PDE case, the Godunov-Rjabenkii condition is actually no sufficient to synthesize back the estimate (1.14). To that aim a reinforced uniform version of the inequality is required:

Theorem 6 (Gustafsson, Kreiss, and Sundström [GKS72])**.** *Under natural solvability properties of the scheme, the strong stability of (1.9) is equivalent to the following uniform version of the Godunov-Rjabenkii condition*

$$\exists C > 0,\ \forall z \in \mathcal{U},\ \forall u \in \mathbb{E}_{\mathsf{s}}(z),\ |(\mathsf{B}u)_0| \geq C|u|. \tag{1.17}$$

In the literature the previous condition is simply known as UKLC to mimic the continuous case. Here the involved norms are actually finite-dimensional, due to the structure of $\mathbb{E}_{\mathsf{s}}(z)$ discussed previously. Even more, the stable space $\mathbb{E}_{\mathsf{s}}(z)$ is a holomorphic fiber bundle over $\mathcal{U}$ and in many comfortable situations such as scalar problems, it can be continuously extended to the unit circle, see [GKS72; Cou13]. Again, resolvent inequalities, block structures and Kreiss symmetrizers are involved in the general theory.

Some usefull results Let us now quickly state two results, that will appear as fundamental tools in the rest of Chapter 1 as well as in Chapter 2.

Lemma 7 (Strong stability of Dirichlet boundary condition)**.** *Let us consider a scalar scheme of the form* (1.9)*, with natural solvability conditions, together with the Dirichlet boundary condition* B*. The Cauchy* ℓ^2*-stability of the scheme is then sufficient to guarantee its strong stability.*

Initially obtained by Goldberg and Tadmor [GT81] for both dissipative and unitary schemes, this result is extended to the general case by Coulombel [Cou13]. An important point is that the inflowing or outflowing nature of the underlying characteristic at the boundary has no incidence on the strong stability for the Dirichlet boundary condition. Of course, we cannot overestimate the incidence of such a choice in terms of the quality of the approximation. Consistency issues are still to discuss (to follow in Chapter 2). A result similar to the previous Lemma is proved by Goldberg [Gol77] for dissipative schemes with extrapolation boundary conditions.

The second result is the discrete counterpart of the method by Rauch [Rau72], with the aim of including nonzero initial data for a full estimate of the discrete solution with respect to the data. It is obtained by J.-F. Coulombel by using adapted Leray-Gårding multipliers in order to have a discrete time integration method for the discrete energy, for multidimensional systems. The statement is quite similar to Theorem 2.

Theorem 8 (Full estimate – Coulombel [Cou15; Cou20])**.** *Consider the scalar scheme* (1.12) *assumed to be Cauchy* ℓ^2*-stable, non-degenerate and with simple roots in the relation dispersion* $z \mapsto \mathcal{P}(z,\kappa)$ *for any* $\kappa \in \mathbb{S}^1$*. The strong stability is equivalent to the UKLC. If satisfied, then for any data* f*,* g *and* F *the solution satisfies, independently of* $\gamma > 0$ *and* $\Delta t \in (0,1]$*:*

$$\sup_{n>0}\left(\mathrm{e}^{-2\gamma n\Delta t}\|u^n\|^2\right) + \frac{\gamma}{1+\gamma\Delta t}\|u\|_\gamma^2 + |u|_\gamma^2 \quad \lesssim \quad \|f\|^2 + |\mathsf{B}u|_\gamma^2 + \frac{1+\gamma\Delta t}{\gamma}\|\mathsf{L}u\|_\gamma^2. \tag{1.18}$$

1.3 Kreiss-Lopatinskii determinants for finite difference schemes

Intrinsic Kreiss-Lopatinksii determinant A crucial question for proving the strong stability of a given discrete IBVP is to detect the zeros of the Kreiss-Lopatinskii determinant. It corresponds to testing the invertibility of B on $\mathbb{E}_{\mathsf{s}}$ leading to the uniform estimate (1.17). Very often in the literature, this is done case by case and thus available only for very special discrete schemes and boundary conditions where the algebra is helpful, i.e. of low degree or reducible. In a more general framework, the Kreiss-Lopatinskii determinant is defined directly from the stable vector bundle $\mathbb{E}_{\mathsf{s}}$ and the boundary operator B. It naturally inherits structural properties as holomorphicity and continuity. However, its concrete computation depends on the considered basis for $\mathbb{E}_{\mathsf{s}}$. Due to the non-trivial monodromy group for the set of roots of the dispersion relation, as functions on z, no holomorphic representation of the individual vector in that basis do exist in general. The previous discussion nevertheless guarantees that the basis as a whole is a holomorphic vector bundle on $\mathcal{U}$.

In the work [A/BLS23], we propose a basis-independent version of the Kreiss-Lopatinskii determinant, which we call the *intrinsic Kreiss-Lopatinskii determinant*. The idea is very simple and enables the use of complex analysis tools to look numerically for the zeros of Δ by complex winding number computations. One can mention the work by Thuné [Thu86] who also developed a numerical method to check the strong stability.

Definition 9 (Intrinsic Kreiss-Lopatinskii determinant)**.** *Let us denote, for any value $z \in \overline{\mathcal{U}}$ a basis $\mathsf{E}(z) = (\mathsf{e}_1(z), \dots, \mathsf{e}_r(z))$ of the r-dimensional vector space $\mathbb{E}_{\mathsf{s}}(z)$, made of sequences indexed by $\{-r, \dots, -1\} \cup \mathbb{N}$. The intrinsic Kreiss-Lopatinskii determinant is the following function:*

$$\Delta(z) = \frac{\det\left(\mathsf{B}\mathsf{E}(z)\right)_0}{\det\left(\Pi_r \mathsf{E}(z)\right)_0}, \tag{1.19}$$

where $\Pi_r := (\mathsf{Id}, \mathsf{S}, \dots, \mathsf{S}^{r-1})^T$.

To get the quite complicated notation, let us look at an example. Assume that $r = 2$, exactly two roots $\kappa_1 \neq \kappa_2$ are in $\mathbb{D}$, and the linear space $\mathbb{E}_{\mathsf{s}}(z)$ is then spanned by $\mathsf{e}_1(z) = (\kappa_1^j)_{j \geq -2}$, $\mathsf{e}_2(z) = (\kappa_2^j)_{j \geq -2}$. Consider the boundary condition $u_{-2} = 0$ and $u_{-1} = u_0$ represented by $\mathsf{B} = (\mathsf{S}^{-2}, \mathsf{Id} - \mathsf{S}^{-1})^T$. In that case, we obtain the formula:

$$(\Pi_2 \mathsf{E}(z))_0 = \begin{pmatrix} 1 & 1 \\ \kappa_1 & \kappa_2 \end{pmatrix}, \qquad (\mathsf{B}\mathsf{E}(z))_0 = \begin{pmatrix} \kappa_1^{-2} & \kappa_2^{-2} \\ 1 - \kappa_1^{-1} & 1 - \kappa_2^{-1} \end{pmatrix}.$$

In some favorable situations, the intrinsic determinant can be computed directly from the coefficients of the scheme.

Theorem 10 (Theorem 13 in [A/BLS23])**.** *Let $a > 0$. Consider a scalar one-step explicit finite difference scheme of the form*

$$\begin{aligned} u_j^{n+1} &= \sum_{\ell=-r}^{p} a_k u_{j+\ell}^n, & j \geq r \\ u_j^n &= \sum_{k=0}^{m-1} b_{j,k} u_k^n + g_j^n, & 0 \leq j \leq r-1, \end{aligned} \tag{1.20}$$

with $p = 0$ (totally upwind), consistent and Cauchy ℓ^2-stable. Then the intrinsic Kreiss-Lopatinskii determinant reads

$$\Delta(z) = (-1)^{r(r-m)} \left(\frac{a_{-r}}{a_0 - z}\right)^{m-r} \det C(z), \tag{1.21}$$

where $C(z)$ is a matrix depending on $(a_\ell)_{-r \leq \ell \leq p}$, on B, and polynomially on z.

Actually, the matrix C can be deduced from the boundary matrix B by an algorithmic gaussian elimination procedure directly related to the linear recurrence relation induced by the interior scheme.

The previous results incidentally gives an alternative direct proof for the holomorphicity and the continuity properties of Δ over $\overline{\mathcal{U}}$ as a result of the property $|a_0| < 1$. The following corollary is central in the forthcoming applications and enables a purely geometrical observation

for determining the strong stability of the considered scheme from the integer value of the complex index $\mathsf{Ind}_{\Delta(\mathbb{S}^1)}(0)$:

Corollary 11 (Corollary 15 in [A/BLS23]). *Let the assumptions of the Theorem 10 hold and assume moreover the absence of neutral instabilities, meaning that* $0 \notin \Delta(\mathbb{S}^1)$. *The equation* $\Delta(z) = 0$ *has then exactly* $r - \mathsf{Ind}_{\Delta(\mathbb{S}^1)}(0)$ *zeros in* $\mathcal{U}$.

In a second work [P/BLS], the result is extended in particular to the case $p \geq 1$. In that case, no such explicit formulation for the Kreiss-Lopatinskii determinant is available but a slight reformulation of the boundary condition is then useful to reduce as much as possible the form of Δ and get the important bounds and properties. Finally, this is proved that the proposed formulation is able to transfer the holomorphicity and continuous extension from $\mathbb{E}_{\mathsf{s}}$ to Δ, whatever is the choice of the basis $\mathsf{E}(z)$.

Theorem 12 ([P/BLS]). *Let* $a > 0$. *Consider a one-step explicit finite difference scheme of the form* (1.20), *consistent and Cauchy* ℓ^2*-stable. The intrinsic Kreiss-Lopatinskii determinant* Δ *is holomorphic on* $\mathcal{U}$ *and continuous on* $\overline{\mathcal{U}}$.

Strategy for winding number computations Applying the residue theorem to the function Δ on the circle $\mathbb{S}^1$, a numerical strategy is available to put the previous discussion in action. The improvement compared to the existing literature is that the method does not use an arbitrary dimension truncation of (quasi-)Toeplitz matrices plus the computation of the spectral radius of such a large matrix. As a counterpart, the numerical experiments are exposed to some technical conditioning difficulties due to large amplitude variations in the modulus of Δ. The winding number computation is then improved by an adaptive mesh refinement strategy based on the work by Zapata and Martín [ZM13] and García Zapata and Díaz Martín [GD12].

1.4 Numerical experiments

Inverse Lax-Wendroff boundary condition Tan and Shu [TS10] develop the inverse Lax-Wendroff (ILW) procedure so as to improve the order of consistency of an inflow boundary for transport-like problems. The method uses the PDE itself so as to transform the space derivatives that appear in extrapolation at the boundary into time derivatives. Then the available physical boundary data can be used in order to define artificial boundary conditions. For example, for the one-dimensional advection problem $\partial_t u + a\partial_x u = 0$, the following relation holds, for $k \in \mathbb{N}^*$,

$$\frac{\partial^k u}{\partial x^k} = \frac{(-1)^k}{a^k} \frac{\partial^k u}{\partial t^k}.$$

So from Taylor expansion up to order d, the ghost points at the inflow boundary $u(t,0) = g(t)$ are chosen to be:

$$u_j^n = \sum_{k=0}^{d-1} \frac{(j\Delta x)^k}{k!} \frac{\partial^k u}{\partial x^k}(n\Delta t, 0) = \sum_{k=0}^{d-1} \frac{(j\Delta x)^k}{k!} (-1)^k \frac{g^{(k)}(n\Delta t)}{a^k}, \quad -r \leq j \leq -1.$$

The method is particularly efficient and benefits directly the stability features available for Dirichlet boundary conditions (Lemma 7): any Cauchy ℓ^2-stable scheme is convenient.

Simplified ILW by extrapolation procedure Motivated by the difficulty to extend the ILW method to multidimensional situations, where defining the inverse Lax-Wendroff procedure requires hard procedures to mix normal and tangential derivatives along the boundary, Vilar and Shu [VS15] proposed a simplified version of the ILW method denoted here "Sk_dILWd". In that method, only the first $k_d - 1$ derivatives of g are substituted in the expansion, next terms for orders from k_d to d are defined from the interior points by an extrapolation procedure at the same order d. At the end, the formula (for the one-dimensional case) takes the form

$$u_j^n = \sum_{k=0}^{k_d-1} \frac{(-j\Delta x)^k}{k!} \frac{g^{(k)}(n\Delta t)}{a^k} + \sum_{k=k_d}^{d-1} \frac{j^k}{k!} \sum_{s=0}^{d-1} p_{k,s}^{(d)} u_s^n, \quad -r \le j \le -1, \tag{1.22}$$

where the coefficients $p_{k,s}^{(d)}$ correspond to the extrapolation procedure.

With our method, we are able now to draw the Kreiss-Lopatinskii curve $\Delta(\mathbb{S}^1)$ for the corresponding discrete IBVP. Now the result by [GT81] (Lemma 7) does not apply anymore and the strong stability of the half-problem has to be investigated again: the CFL parameter λ has to be chosen so as to get the full stability property (even for a Cauchy dissipative ℓ^2-scheme). To that aim, we compute the curve $\Delta(\mathbb{S}^1)$ for several values of the CFL parameter λ (see Figure 1.1). The Beam-Warming scheme (see [Des09]) is known to be Cauchy ℓ^2-stable exactly for $\lambda \in (0, 2]$. The instability for $\lambda = 1.4$ and the stability for $\lambda = 1.6$ are thus proved.

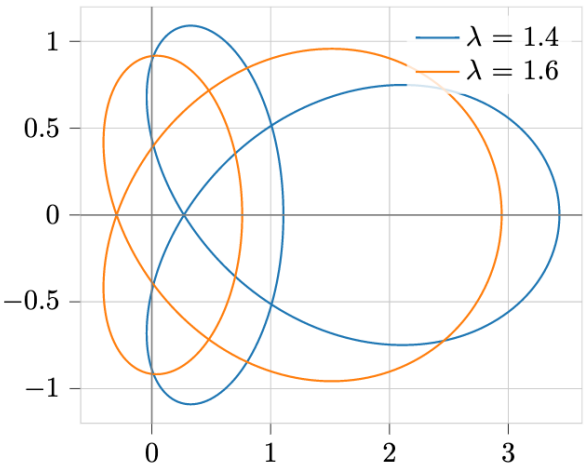

Figure 1.1: Set $\Delta(\mathbb{S}^1)$ for the Beam-Warming scheme with S2ILW3.

Misalignment of grids The ILW and SILW procedures can be very easily extended to the case where a misalignment between physical boundaries and grid points occur. Consider

the boundary condition $u(t, x_\sigma) = g(t)$ now prescribed at a point $x_\sigma = \sigma\Delta x$. The value of σ is thought as possibly changing when time evolves (for moving boundaries $x_\sigma(t)$) or along another space direction missing here (for multidimensional setting with curved or oblique boundaries, with $x_\sigma(y)$). The aim is to provide strong stability results that are independent of the value of σ, at least in a known set of values.

In that case, the adapted SILW procedure depends on σ and reads:

$$u_j^n = \sum_{k=0}^{k_d-1} \frac{(-(j+\sigma)\Delta x)^k}{k!} \frac{g^{(k)}(n\Delta t)}{a^k} + \sum_{k=k_d}^{d-1} \frac{(j+\sigma)^k}{k!} \sum_{s=0}^{d-1} p_{k,s}^{(d)} u_s^n, \quad -r \le j \le -1.$$

For the Beam-Warming scheme, we explore simultaneously the set of the parameters λ and $\sigma \in [-0.5, 0.5]$ for which the strong stability occurs (see Figure 1.2). The main advantage of the method is now clear: the computational cost to draw these maps is reduced and does not depend on high-dimensional spectrum approximations for matrices, but rather on the number of points used to compute $\mathsf{Ind}_{\Delta(\mathbb{S}^1)}(0)$. It has also to be observed that exceptional points with $0 \in \Delta(\mathbb{S}^1)$ only happens quite rarely, along transition curves only.

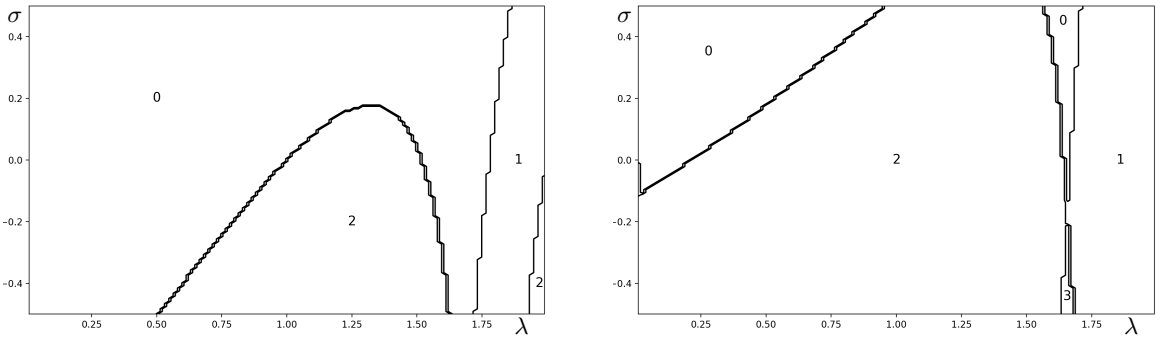

Figure 1.2: Number of boundary instabilities for the Beam-Warming scheme with S2ILW3 (left) and with S1ILW3 (right).

Simplified ILW by reconstruction procedure In the literature an alternative procedure to simplify ILW boundary conditions by simplifying high order terms is proposed by Dakin, Després, and Jaouen [DDJ18]. We skip here many details to not rise unnecessarily the technicality. The boundary condition, denoted $\mathcal{R}^{d,k_d}$, has the form

$$(u_{-r}^n, \dots, u_{-1}^n)^T = \mathcal{Y}_- \mathcal{Y}_+^{-1} (u_0^n, \dots, u_{d-k_d-2}^n)^T + G^n,$$

with a convenient reconstruction matrix $\mathcal{Y}_- \mathcal{Y}_+^{-1}$ (depending on σ) and G^n depending also on the boundary data g. The Figure 1.3 represents the number of boundary instabilities computed for the fifth order Lax-Wendroff method [LM06] with such reconstruction procedures at the boundary.

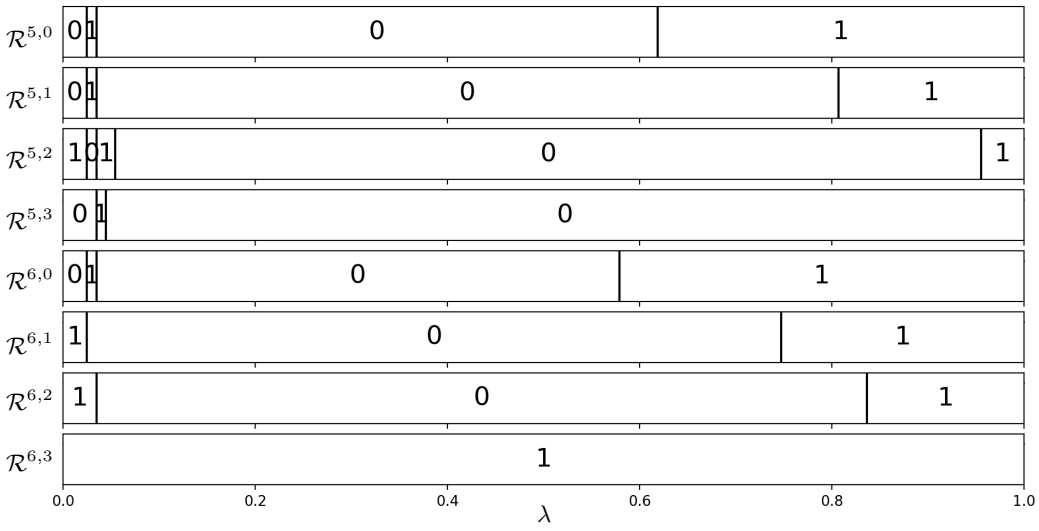

Figure 1.3: Number of boundary instabilities for the fifth order Lax-Wendroff scheme with different reconstruction boundary conditions for $\lambda \in]0, 1]$ and $\sigma = 0.4$.

1.5 Perspectives

(1A) Porting of the technique to other problems. The proposed technique is particularly powerful for high order (so large stencils) schemes for which the complicated underlying algebra generally precludes the stability proof by hand. A first natural research direction is to grasp and surely adapt the previously established determinantal technique to *other discrete boundary conditions* handling with *more general PDEs, higher dimension problems or even schemes for kinetic problems*. For example, recently Li, Shu, and Zhang [LSZ17] studied an inverse Lax-Wendroff technique adapted to central schemes for solving diffusion equations; Al Hassanieh, Banks, Henshaw, and Schwendeman [ABH+22] defined local compatibility boundary conditions (CBC) for multidimensional wave equations. A second porting of the technique is for drawing stability maps such as Figure 1.2 in the case of parameterized discrete transmission conditions between several numerical schemes handled with domain decomposition techniques. The application of interest is the interfacial coupling of hyperbolic and/or dispersive models in the framework of water wave models with artificial transmission conditions, first in the linear setting or in a linearized analysis.

(1B) Spectral theory, Wiener-Hopf factorizations and Shur complements. The theory of stability for the discrete problems may be analysed through several frameworks. For example, finite difference schemes with constant coefficients and no boundary (periodic or doubly-infinite domains) are well-known through circulant Toeplitz matrices (i.e. Laurent operators). The Fourier-diagonalization-multiplier traduces the uniform boundedness of the associated discrete semi-group. This is the most classical situation, where several aspects coincide: matrix spectrum, resolvent inequalities, uniform power boundedness.

In the presence of boundaries, the situation slightly differs since we generally face then

non-selfadjoint operators. General (pseudo-)spectral analysis for large matrices is then covered by the Kreiss Matrix Theorem. The literature and the results on the topic is large (Nevanlinna [LN91; Nev01], Spijker and collaborators [BDS02; BS00; BS02; Spi17], Szehr [Sze14] and the book by Trefethen and Embree [TE05]).

On the other side, the case with Dirichlet boundary conditions on the one-dimensional half-space leads to Toeplitz operators (and their finite sections), whose spectral theory is quite well understood. More general boundary conditions are studied through their quasi-Toeplitz form: the asymptotic of their spectrum and pseudospectrum (resolvent inequalities) for large dimensions is handled by Reichel and Trefethen [RT92] and Beam and Warming [BW93] (see also the older work by Schmidt and Spitzer [SS60] concerning the asymptotic spectrum of Toeplitz matrices). Some of these aspects have been explored in the PhD work of P. Le Barbenchon (numerical exploration of bulges of the pseudospectrum of large Toeplitz matrices and their relation with generalized eigenvalues of the boundary problem).

Actually the semi-infinite matrix obtained from another boundary condition defines a Fredholm operator that is a compact perturbation of the original Toeplitz matrix. Therefore their spectra only differ from a finite set of discrete points. The Kreiss-Lopatinskii determinant is in some sense the reduced finite-dimensional determinant of that perturbation, computed from reducing first the symbolic form by means of Wiener-Hopf factorization (the Hersh separation lemma).

A perspective is to clarify these spectral aspects with the help of spectral theory and harmonic functions theory.

CHAPTER 2

Multiscale expansions for discrete boundaries

The present chapter is focused on multiscale expansions describing the solutions to finite difference schemes, and on their uses to analyze the convergence properties and understand some stability features. The typical decomposition of the discrete solution is achieved from modes emanating from the boundary and usual Fourier interior modes. In a first section, we introduce some general ideas partially related to the previous chapters. Then, from the strong stability property and a consistency analysis adapted to the asymptotic expansion, we obtain some improved convergence results. The considered analysis enables a new descriptive proof of semigroup estimates for the discrete mixed initial boundary value problem, that are close to be optimal, in comparison with the continuous case. Some various possible extensions of these techniques are finally discussed. The presentation is mainly based on the publications [A/BC17; A/BNS$^+$21] recalled hereunder.

[A/BC17] B. Boutin and J.-F. Coulombel. Stability of finite difference schemes for hyperbolic initial boundary value problems: numerical boundary layers. *Numer. Math. Theory Methods Appl.* 10 3:489–519, 2017.

[A/BNS$^+$21] B. Boutin, T. H. T. Nguyen, A. Sylla, S. Tran-Tien, and J.-F. Coulombel. High order numerical schemes for transport equations on bounded domains. *ESAIM: ProcS*, 70:84–106, 2021.

2.1 Multiscale expansions

> Even if an equation is nondispersive, any discrete model of it will be dispersive.
>
> Trefethen [Tre82]

Finite difference schemes can be interpreted as discrete non-local operators that filter informations in concordance with a given continuous problem, in the limit of small discretization parameters Δx and Δt. The filtering effect is due typically to the underlying smoothness of the approximated objects and/or to damping properties. However, in some situations, some parts of the discrete solution may refuse to obey the expected limiting problem, or obey another limiting process. This is true especially when the the approximated solution is lacking in smoothness, due e.g. to non-linearities, or when impurities affect the scheme, e.g. inhomogeneities or boundaries. These last specific situations can actually be considered itself as a lack of smoothness in the coefficients of the filter itself.

The above quoted citation from Trefethen reminds us about the dispersive nature of difference schemes, and one should have in mind that a scheme does not only solve the single partial differential equation (PDE) it is designed for. In fact, it solves a full set of PDEs, among which the expected model has somehow a dominant presence due to the previously mentioned filtering effect. The other hidden PDEs are unexpected but can be awakened for example by the presence of inhomogeneities or boundaries. Thus the paradigm hereafter is to consider the discrete solution to a linear numerical scheme with or without boundaries as being the approximate superposition of solutions to several PDEs involving various time and space scales and more or less independent structural properties. This is a natural continuation of the common Fourier symbolic analysis for the Cauchy problem [Str04], but with an enriched basis of representation.

The dispersive analysis of numerical methods is clearly not new and has received many interests from the 1980's. It follows the group velocity wave analysis from Brillouin [Bri60] and some important contributions with the use of dispersive features for discrete schemes are done in particular by Vichnevetsky [Vic81a; Vic81b], by Trefethen [Tre82; Tre84], by Higdon [Hig86a; Hig86b], by Michelson [Mic87]. The paper by Trefethen [Tre84] actually makes a rich connection in between the strong (GKS-) instability and the presence of what he calls "rightgoing steady-state solutions". In particular the amplitude of instability factors are related to kind of instabilities present in the scheme.

For non-linear schemes with boundaries, the study of boundary layers and of their stability has been studied for example by Chainais-Hillairet and Grenier [CG01] for the Lax-Friedrichs scheme (see also Godillon [God03] for the more general study of discrete shock profiles in that case), and by Ye [Ye04] for non-characteristic relaxation models.

From now on, we consider only finite difference schemes for solving the one-dimensional scalar constant coefficient transport problem

$$\partial_t u + a\partial_x u = 0, \qquad x > 0,\ t > 0, \tag{2.1}$$

but many ideas are still available for higher order constant coefficients initial boundary value PDEs.

Normal modes, conjugated schemes and equivalent equations To begin with, we consider *normal modes* of the form $u_j^n = z^n \kappa^j$ with $(z, \kappa) \in \mathbb{C}^2$ as the typical solution involved by the roots of the dispersion relation $\mathcal{P}(z, \kappa)$. The space structure and the time dynamic of these particular solutions are given by some simple considerations. For example, a mode (z, κ) is said to be *stationary* if $z = 1$ and more generally *unitary* if $|z| = 1$. It is said to be *damped* if $|z| < 1$ and *unstable* if $|z| > 1$. Concerning the space structure in dimension one, the mode is *left localized* if $|\kappa| < 1$, *right localized* if $|\kappa| > 1$, and finally *extended* if $|\kappa| = 1$. This last case precisely correspond to the classical Fourier modes for the problem without boundaries. When there is no ambiguity, for example when dealing with the half problem set on $\mathbb{N}$, *localized* means simply *left localized*.

The linear analysis allows to easily consider a *modulated* version of the previous solution, that is a solution of the form $u_j^n = v(t^n, x_j)\underline{z}^n \underline{\kappa}^j$ where v is a smooth profile function and $\mathcal{P}(\underline{z}, \underline{\kappa}) = 0$. Then $v_j^n = v(t^n, x_j)$ solves then approximately a conjugated scheme shifted in (z, κ), having then to be precise the dispersion relation

$$\underline{\mathcal{P}}(z, \kappa) := \mathcal{P}(\underline{z}z, \underline{\kappa}\kappa) = 0.$$

In particular, for (z, κ) close to $(1, 1)$, the tangency properties of the zero set of $\underline{\mathcal{P}}$ unveil the *modulated PDE* satisfied by the profile function v at the lowest order. The consistency property of the scheme with a given PDE (such as (2.1) here) thus reduces to these tangency properties, in the spirit of matching expansions. The first leading equivalent equation at the scale Δt and Δx reads, after some Taylor expansions:

$$\Delta t \frac{\partial \underline{\mathcal{P}}}{\partial z}(1, 1)\partial_t v + \Delta x \frac{\partial \underline{\mathcal{P}}}{\partial \kappa}(1, 1)\partial_x v = 0.$$

According to the possible vanishing in the sequence of derivatives of $\underline{\mathcal{P}}$ at $(1, 1)$, it may be necessary to increase the order of the corresponding Taylor expansions so as to find out the dominant term in the expansion and finally a higher order modulated PDE. In view of having in the end the good scaling in the discretization parameters Δt and Δx, this is then necessary to change accordingly the time and/or space scale for the modulated profile. For example, assuming

$$\frac{\partial \underline{\mathcal{P}}}{\partial \kappa}(1, 1) = 0 \text{ and } \frac{\partial \underline{\mathcal{P}}}{\partial z}(1, 1) \neq 0,$$

then, this is natural to consider the solution to the conjugated scheme v_j^n to have the form $v_j^n = v(n\Delta t, j\Delta x^{1/2})$ so that the profile function v solves at low order the PDE problem:

$$\Delta t \frac{\partial \underline{\mathcal{P}}}{\partial z}(1, 1)\partial_t v + \frac{\Delta x}{2} \frac{\partial^2 \underline{\mathcal{P}}}{\partial \kappa^2}(1, 1)\partial_x^2 v = 0.$$

In the end, one understands that along the zero set of the dispersion relation $\mathcal{P}$ of the numerical scheme can be found a set of PDEs, each one being consistent with the corresponding

conjugated schemes. All of them have various stability and scaling properties. Of course, in view of a given final expected application, not any of these numerous modes and scales have to be considered. Let us give hereafter some quick guidelines.

- Handling with the problem without boundaries, the Cauchy ℓ^2-stability requires that $0 \notin \mathcal{P}(\mathcal{U}, \mathbb{S}^1)$. For smooth initial data (both in time and space) only modes for $|\kappa| = 1$ (Fourier analysis) and $|z| \leq 1$ are relevant in the convergence analysis. Of course, damped modes for $|z| < 1$ cannot damage seriously the convergence results for positive times. For linear multistep methods, the relevant analysis can make use of the so-called one-step underlying method [EN88]. However, for weakly stable multistep method, because of the presence of unitary non-trivial modes $|z| = 1$ and $|\kappa| = 1$, an appropriate choice for the initial conditions is required so as to guarantee the expected convergence result without spurious time oscillations in the solutions.

- Handling with one boundary at the left of the domain, the Fourier analysis has to be supplemented by considering in addition possible localized unitary modes ($|z| = 1$ and $|\kappa| < 1$) and extended unitary modes ($|z| = 1$ and $|\kappa| = 1$). Of course, among these modes the stationary ones ($z = 1$) play a central role for the convergence analysis with smooth in time solutions. However more general situations may also require to deal with the full set of unitary modes.

For the rest of the chapter, the presentation is reduced to *explicit linear multistep schemes* constructed by an *method-of-lines*, discretizing first in space. They were already met in the previous chapter but for convenience we recall hereafter the form for the half-problem set over $j \in \mathbb{N}$ with discrete boundary condition from an finite difference operator B.

$$\begin{aligned} (\mathsf{L}u)_j^n := \frac{1}{\Delta t} \sum_{\sigma=0}^{k} \alpha_\sigma \, u_j^{n+\sigma} + \sum_{\sigma=0}^{k-1} \beta_\sigma \, \frac{1}{\Delta x} \sum_{\ell=-r}^{p} a_\ell \, u_{j+\ell}^{n+\sigma} = 0, \qquad & (n,j) \in \mathbb{N} \times \mathbb{N}, \\ (\mathsf{B}u)_j^n = g_j^n, \qquad & 0 \leq j \leq r-1, \ n \geq k, \\ u_j^\sigma = f_j^\sigma, \qquad & j \in \mathbb{N}, \ 0 \leq \sigma \leq k-1. \end{aligned} \tag{2.2}$$

We recall that the corresponding dispersion relation for the interior scheme is

$$\mathcal{P}(z, \kappa) := \frac{1}{\Delta t} \rho_1(z) + \frac{1}{\Delta x} \rho_0(z) \mathcal{A}(\kappa) = 0, \tag{2.3}$$

with the notations already introduced around (1.13).

2.2 Discrete boundary layers

The mostly observed parasitic phenomenon when solving discrete initial boundary value problems consists in spurious oscillations located close to the boundaries of the computational domain. Actually, this phenomenon is present for almost any scheme, sometimes at imperceptible scales. Their careful analysis and description is particularly interesting and leads to

improved convergence estimates by adapting the consistency analysis through appropriate asymptotic expansions. Strong stability properties have here to be already available so as to go easily from boundary consistency errors towards convergence errors. This is the case for example when dealing with the Dirichlet discrete boundary conditions, known from Goldberg and Tadmor [GT81] to lead to a strongly stable scheme when associated to any scalar scheme that exhibits Cauchy ℓ^2-stability for the interior domain.

Localized stationary modes The case of semi-discrete or of method-of-lines multistep schemes is more convenient to cover because the algebra related to the discrete space operator can then be separated from the other aspects. Let us consider the fully discrete case (2.2) and stationary modes only. They admit the modal form $u_j^n = z^n \varphi_j$ with $z = 1$ and where $(\varphi_j)_j$ is some appropriate complex valued sequence. The set of stationary solutions is entirely spanned by the geometrical sequences corresponding to the roots of the dispersion relation for $z = 1$:

$$\mathcal{P}(1, \kappa) = 0, \tag{2.4}$$

with $\kappa \in \mathbb{C}$. Among them, we distinguish the *localized* modes (in ℓ^2) spanned by the *interior* roots only (meaning $|\kappa| < 1$), from the *extended* modes also involving the unitary roots (meaning $|\kappa| = 1$).

From now on, we restrict our attention to space operators satisfying the following structural assumption that consists in the *absence of extended stationary mode*:

$$\forall \kappa \in \mathbb{S}^1 \setminus \{1\},\ \mathcal{A}(\kappa) \neq 0. \tag{2.5}$$

As a typical example, a one-step scheme based on Euler forward time integrator that is dissipative in the sense of Kreiss satisfies the previous assumption, namely:

$$|1 - \tfrac{\Delta t}{\Delta x} \mathcal{A}(e^{i\theta})| \leq 1 - c\eta^{2m}, \quad |\eta| \leq \pi.$$

Other non-dissipative schemes, such as the Lax-Friedrichs scheme also satisfy the required assumption.

Adapted consistency analysis The classical convergence error denoted by $e_{n,j}^{\mathsf{cl}} = u(t^n, x_j) - u_j^n$ is estimated through the main stability property and estimates for the truncation error $\varepsilon_{n,j}^{\mathsf{cl}} = \mathsf{L}u(t^n, x_j)$ and for the initial error $e_{\sigma,j}^{\mathsf{cl}}$ for $0 \leq \sigma \leq k-1$. Now, considered as an improved comparison object, we slightly modify the process by adding a new whole asymptotic expansion $u_{n,j}^{\mathsf{app}}$ at dominant order:

$$e_{n,j} = u_{n,j}^{\mathsf{app}} - u_j^n.$$

The method consists in constructing the appropriate boundary layer expansion so that $e_{n,j}$ solves the discrete IBVP:

$$\begin{aligned} (\mathsf{L}e)_{n,j} &= \varepsilon_{n,j}, & (n,j) &\in \mathbb{N}\times\mathbb{N} \\ (\mathsf{B}e)_{n,j} &= \eta_{n,j}, & 0 &\le j \le r-1,\ n \ge k \\ e_{\sigma,j} &= 0, & 0 &\le \sigma \le k-1,\ j \ge 0 \end{aligned}$$

with a zero initial data, small boundary truncation error η and small interior truncation error ε, thus forcing the boundary inconsistency to reduce. These smallness requirements motivate the appropriate expansion for having the form

$$u^{\mathsf{app}}_{n,j} = u_{\mathsf{int}}(t^n, j\Delta x) + u_{\mathsf{bl}}(t^n, j) \tag{2.6}$$

where u_{int} solves the interior equation and u_{bl} corrects the boundary truncation error. It has to not modify severely the interior truncation error $\varepsilon^{\mathsf{cl}}$ and thus to correspond to well-chosen localized modes only. The main step to choose appropriately the boundary layer term is the boundary equation that, for convenience, we present here in the semi-discrete case only:

$$(\mathsf{B}u_{\mathsf{bl}}(t))_j = -(\mathsf{B}u_{\mathsf{int}}(t))_j, \quad 0 \le j \le r-1. \tag{2.7}$$

and $u_{\mathsf{bl}}(t) \in \mathbb{E}_{\mathsf{ss}}(1)$ the strictly stable space associated to $z=1$, involving only interior roots $|\kappa|<1$, while u_{int} is directly related to the usual consistency mode $(z,\kappa)=(1,1)$. According to the sign of a, $\kappa=1$ corresponds to a mode in the continuous extension $\mathbb{E}_{\mathsf{s}}(1)$ of the stable space. The following lemma is a convenient reinterpreted version of the Hersh Lemma from [A/BC17].

Lemma 13 (Construction of $\mathbb{E}_{\mathsf{ss}}(1)$)**.** *Let $a \neq 0$ and consider a scheme of the form (2.2) Cauchy ℓ^2-stable and without extended stationary mode. Then $\mathcal{A}$ admits exactly*

- *$\kappa = 1$ as simple root, moreover $(1^j) \in \mathbb{E}_{\mathsf{s}}(1)$ iff $a>0$,*
- *r interior roots $|\kappa|<1$ and $p-1$ exterior roots $|\kappa|>1$ (with mult.), if $a<0$,*
- *$r-1$ interior roots $|\kappa|<1$ and p exterior roots $|\kappa|>1$ (with mult.), if $a>0$.*

It has to be observed that for outgoing problem $a<0$, then $\mathbb{E}_{\mathsf{s}}(1) = \mathbb{E}_{\mathsf{ss}}(1)$ while for the incoming problem $a>0$, then $\mathbb{E}_{\mathsf{s}}(1) = \mathbb{E}_{\mathsf{ss}}(1) \cup \{1\}$. In any case, $\dim \mathbb{E}_{\mathsf{s}}(1) = r$ and the equation (2.7) is always solvable from the invertibility of B on $\mathbb{E}_{\mathsf{s}}(1)$ (UKLC).

Extrapolation boundary conditions The most convenient discrete boundary condition to program is the Dirichlet boundary condition, even if rarely consistent. It consists in setting the required ghost values directly from some prescribed exterior data. A standard improvement of this choice, especially promoted for outflowing problems, consists in considering extrapolatory discrete boundary conditions. It is also known as discrete Neumann boundary conditions. The stability properties of these methods have been studied by Goldberg [Gol77]

for dissipative or unitary schemes, or by Goldberg and Tadmor [GT81] for Cauchy ℓ^2-stable schemes with Dirichlet boundary conditions.

Let us fix some notations and begin by introducing the following right sided discrete difference operator $(\mathsf{D}_+ v)_j = v_{j+1} - v_j$. The previously mentioned boundary conditions take the following form:

$$(\mathsf{D}_+^m u)_\ell = g_\ell, \quad 0 \leq \ell \leq r-1. \tag{2.8}$$

where $(g_\ell)_{0\leq\ell\leq r-1}$ is the discrete boundary data. For example, when solving physically outflowing problems where no data is thus available from the continuous setting, the most easy choice consists in prescribing homogeneous discrete boundary conditions at some given order m:

- $m = 0$ (homogeneous Dirichlet): $u_0 = \ldots = u_{r-1} = 0$;
- $m = 1$ (homogeneous Neumann): $u_0 = \ldots = u_{r-1} = u_r$;
- $m = 2$ (second-order homogeneous Neumann): $u_1 - u_0 = \ldots = u_r - u_{r-1} = u_{r+1} - u_r$.

Smooth functions naturally benefits the following consistency error at the boundary:

Lemma 14 (Consistency of extrapolation)**.** *Let $m \in \mathbb{N}$. There exists $C > 0$ such that for any $v \in H^{m+1}(\mathbb{R})$ and any $0 \leq \ell \leq r-1$: $|(\mathsf{D}_+^m v)_\ell| \leq C\Delta x^m \|v^{(m)}\|_{H^1(\mathbb{R})}$, where $v_j = \frac{1}{\Delta x}\int_{C_j} v(y)\,dy$ for $j \in \mathbb{Z}$.*

Discrete derivatives of confluent Vandermonde matrices The main step for identifying the boundary layer terms lies in the solving of the boundary equation (2.7). The Dirichlet boundary conditions (2.8), for $m = 0$, is solved, for example when the involved roots $\kappa_1, \ldots, \kappa_r$ of $\mathcal{P}(1, \kappa)$ have multiplicity one, by inverting the classical Vandermonde matrix:

$$\begin{pmatrix} 1 & 1 & \ldots & 1 \\ \kappa_1 & \kappa_2 & \ldots & \kappa_r \\ \vdots & \vdots & & \vdots \\ \kappa_1^{r-1} & \kappa_2^{r-1} & \ldots & \kappa_r^{r-1} \end{pmatrix}.$$

Now for roots $\kappa_1, \ldots, \kappa_r$ of $\mathcal{P}(1, \kappa)$ with possible multiplicities, this requires to invert the confluent Vandermonde matrix:

$$V = \begin{pmatrix} \kappa_0^{(1)} & \kappa_0^{(2)} & \ldots & \kappa_0^{(r)} \\ \kappa_1^{(1)} & \kappa_1^{(2)} & \ldots & \kappa_1^{(r)} \\ \vdots & \vdots & & \vdots \\ \kappa_{r-1}^{(1)} & \kappa_{r-1}^{(2)} & \ldots & \kappa_{r-1}^{(r)} \end{pmatrix},$$

where the sequences $(\kappa_j^{(\ell)})_j$ for $1 \leq \ell \leq r$ correspond to a relabeling of the sequences $j(j-1)\ldots(j-\sigma+1)\kappa^{j-\sigma}$ for $0 \leq \sigma \leq \mu - 1$ considered for any interior root κ of $\mathcal{P}(1, \kappa)$ with multiplicity μ. The solving of Neumann extrapolation boundary condition at order m requires the inversion of the following *discrete derivative at order m of the confluent Vandermonde*

matrix:

$$\mathsf{D}^m V := \begin{pmatrix} (\mathsf{D}_+^m \kappa^{(1)})_0 & (\mathsf{D}_+^m \kappa^{(2)})_0 & \dots & (\mathsf{D}_+^m \kappa^{(r)})_0 \\ (\mathsf{D}_+^m \kappa^{(1)})_1 & (\mathsf{D}_+^m \kappa^{(2)})_0 & \dots & (\mathsf{D}_+^m \kappa^{(r)})_1 \\ \vdots & \vdots & & \vdots \\ (\mathsf{D}_+^m \kappa^{(1)})_{r-1} & (\mathsf{D}_+^m \kappa^{(2)})_{r-1} & \dots & (\mathsf{D}_+^m \kappa^{(r)})_{r-1} \end{pmatrix}.$$

Incidentally, all these matrices are invertibles. For the last case, a direct proof is available in [A/BNS+21] using the divided difference algebra and the specific Leibniz formula (see Popoviciu [Pop40] and de Boor [dBoo05]).

Asymptotic expansions for outflow boundary conditions From the adapted consistency analysis adapted to the constructed expansion, using the scale localization property of the boundary terms and comparing different discrete trace formulations.

Theorem 15 (Outflow Dirichlet [A/BC17])**.** *Let us consider a consistent Cauchy ℓ^2-stable scalar scheme of the form (2.2) with homogeneous boundary condition (2.8) at order $m = 0$. Assume that there is no extended stationary mode. For smooth initially compatible data $f \in H^2(\mathbb{R}_+^*)$, the discrete solution u_j^n has the asymptotic expansion*

$$u_j^n = u_{\mathsf{int}}(t^n, x_j) + u_{\mathsf{bl}}(t^n, j) + h.o.t,$$

where

- *u_{int} is the exact solution to $\partial_t u + a\partial_x u = 0$ $(a < 0)$ with initial data f,*
- *$u_{\mathsf{bl}}(t^n, \cdot) \in \mathbb{E}_{\mathsf{s}}(1)$ solves $u_{\mathsf{bl}}(t^n, \cdot) = -u_{\mathsf{int}}(t^n, 0)$.*

Theorem 16 (High order outflow [A/BNS+21])**.** *Let us consider a consistent order k Cauchy ℓ^2-stable scalar scheme of the form (2.2) with homogeneous boundary conditions (2.8) at order $m < k$. Assume that there is no extended stationary mode. For smooth initially compatible data $f \in H^{k+1}(\mathbb{R}_+^*)$, the discrete solution u_j^n has the asymptotic expansion*

$$u_j^n = u_{\mathsf{int}}(t^n, j\Delta x) + \Delta x^m u_{\mathsf{bl}}(t^n, j) + h.o.t. \tag{2.9}$$

where

- *u_{int} is the exact solution to $\partial_t u + a\partial_x u = 0$ $(a < 0)$ with initial data f,*
- *$u_{\mathsf{bl}}(t^n, \cdot) \in \mathbb{E}_{\mathsf{s}}(1)$ solves $\mathsf{D}_+^m u_{\mathsf{bl}}(t^n, \cdot) = -\Delta x^{-m} \mathsf{D}_+^m u_{\mathsf{int}}(t^n, \cdot)$.*

In the previous above expansions, due to the explicit form of the the boundary layer, the following estimate is available:

$$\sup_{n \in \mathbb{N}} \sup_{0 \le j \le r-1} |u_{\mathsf{bl}}(t^n, j)| \lesssim \Delta x^{1/2} \|f\|_{H^{m+1}}.$$

In a previous work by Coulombel and Lagoutière [CL20], the transport equation on a onedimensional bounded domain is studied. The numerical scheme under consideration is

supposed to be ℓ^2-stable with a given consistency order k. The inflowing boundary is handled with homogeneous Dirichlet boundary condition and the outflowing with extrapolation at order $m \leq k$. By a new energy method, the authors obtain a quantified convergence result in the norm $\ell_t^\infty \ell_x^2$ as $O(\Delta x^{\min(k,m)})$ and then deduce an estimate in the norm $\ell_t^\infty \ell_x^\infty$ as $O(\Delta x^{\min(k,m)-1/2})$. The purpose of the work in [A/BNS+21] is twofold. First, we extend the result and analysis to the case of non-homogeneous incoming data, by means of the inverse Lax-Wendroff procedure proposed by Tan and Shu [TS10] considered at the convenient order. The global order of the scheme is that way not destroyed. The second point is to make use of the previously discussed boundary layer expansions in order to improve the convergence estimate by a factor $\Delta x^{1/2}$, this result being supported by the effective simulations.

Theorem 17 (Improved convergence estimates [A/BNS+21])**.** *Consider the assumptions of Theorem 16 (with $m < k$). Then the discrete solution u_j^n satisfies the following convergence estimates in the $\ell_t^\infty \ell_x^2$ norm:*

$$\sup_{0\leq n\leq N_T} \left(\sum_{j\geq 0} \Delta x |u_j^n - u_{\mathsf{int}}(t^n, x_j)|^2 \right)^{1/2} \leq C(T) \Delta x^{m+1/2} \|f\|^2_{H^{k+1}}, \tag{2.10}$$

and in the $\ell_t^\infty \ell_x^\infty$-norm:

$$\sup_{0\leq n\leq N_T} \sup_{j\geq 0} |u_j^n - u_{\mathsf{int}}(t^n, x_j)|^2 \leq C(T) \Delta x^{m} \|f\|^2_{H^{k+1}}. \tag{2.11}$$

Some numerical illustrations are represented on Figures 2.1 and 2.2. They correspond to several schemes of order $k = 1$ (upwind), $k = 2$ (Lax-Wendroff), $k = 3$ (O3), $k = 5$ (Lax-Wendroff 5), considered with the inflow high order Inverse Lax-Wendroff boundary condition at order k and extrapolation boundary condition at order m. The Figure 2.1 is for the case $m = k - 2$. The numerical results are in agreement with the estimates in the previous theorem. The upwind scheme has no outflow numerical boundary condition and still of order 1. The Lax-Wendroff scheme does not converge in the ℓ_x^∞ norm but converges at order almost $1/2$ in ℓ_x^2. The third order scheme has order almost 1 in the ℓ_x^∞ norm but almost $3/2$ in ℓ_x^2. The fifth order scheme has order almost 3 in the ℓ_x^∞ norm but almost $7/2$ in ℓ_x^2. With the same computation in the case $m = k$, the Figure 2.1 supports the validity of no loss in the order for the norms but actually our method with the asymptotic boundary layer expansion is not able to prove the result.

Application to discrete semigroup estimates Using an additional corrective term $\Delta x\ u_{\mathsf{bl},1}$ to handle with the time variations of the trace $u_{\mathsf{int}}(t^n, 0)$, the asymptotic expansion present in the Theorem 15 is slightly improved so as to get the following semigroup estimate.

Corollary 18 (Semigroup-like estimate [A/BC17])**.** *Under the assumptions of Theorem 15, the numerical solution satisfies the close to optimal $\ell_t^\infty \ell_x^2$-semigroup estimate:*

$$\sup_{0\leq n\leq N_T} \sum_{j\geq 0} \Delta x |u_j^n|^2 \leq C(\|f\|^2_{L^2} + \Delta x\, T^3 \|f\|^2_{H^2}), \tag{2.12}$$

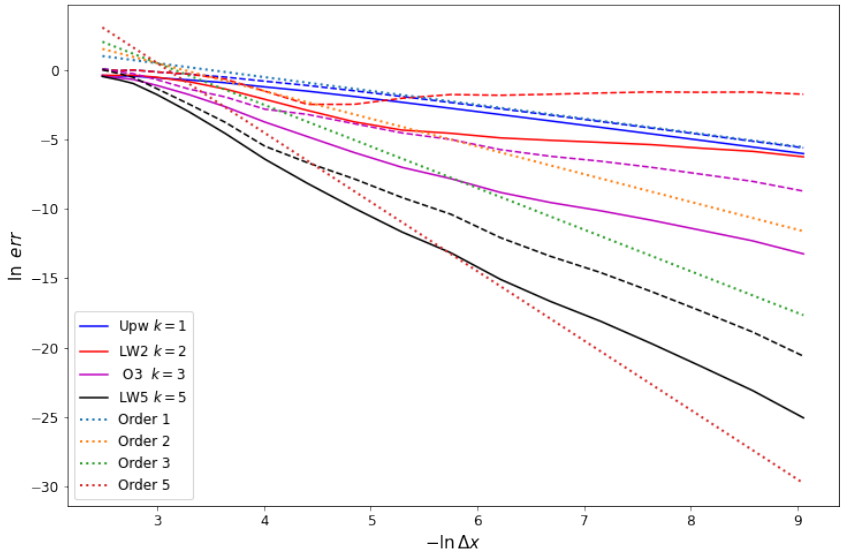

Figure 2.1: Convergence curves ℓ^2_x (solid) / ℓ^∞_x (dashed) with $m = k - 2$.

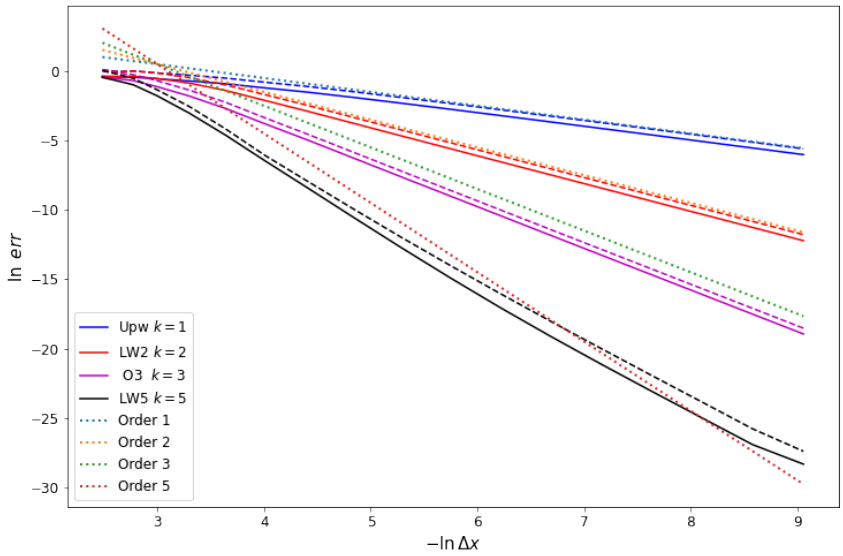

Figure 2.2: Convergence curves ℓ^2_x (solid) / ℓ^∞_x (dashed) with $m = k$.

where the constant $C > 0$ only depends on the scheme.

Let us comment briefly the interest of this result. In the continuous PDE case, obtaining semigroup stability estimates from the hyperbolicity and under the UKLC condition is not easy in any case. The symmetrizable hyperbolic systems with dissipative boundary condition is covered since the seminal work by Friedrichs and Lax [FL67]. The methodology has been extended by Rauch [Rau72] to any strictly hyperbolic system, later by Audiard [Aud11] to semi-strictly hyperbolic systems and recently by Métivier [Mét17] for a very general class of systems (having the *block structure* property). The same difficulty holds for discrete numerical schemes, that is: how to assemble the Cauchy stability to the strong (GKS) stability in order to obtain a full estimate of the form (1.18) with non zero initial data. The argument by Wu

[Wu95], restricted however to the scalar or to the onedimensional one-step schemes, consist in constructing from the Cauchy ℓ^2-stability an auxiliary strictly dissipative boundary condition. Then the result by Goldberg and Tadmor [GT81] enables a superposition argument making use of the Dirichlet boundary conditions. Several more recent works treat with this topic to extend the result, always by designing auxiliary dissipative boundary conditions in the spirit of energy methods: Coulombel and Gloria [CG11], Coulombel [Cou15; Cou20].

Application to the interfacial coupling of discrete schemes The previous discussion deal only with boundary problems, but with homogeneous (constant coefficients) interior schemes. This excludes somehow the treatment of nonlinearities or of non-homogeneous schemes. The specific case of piecewise homogeneous schemes (onedimensional domain decomposition) is however possibly analyzed by the same above ideas. In the literature, this situation has already been discussed earlier by Vichnevetsky [Vic81b], Giles and Thompkins [GT83] and Giles and Thompkins [GT85], and more recently Trefethen and Chapman [TC04] with the construction of approximate pseudomodes for twisted Toeplitz matrices.

The previous asymptotic boundary layer analysis, supported by the discrete trace estimates for strongly stable boundary schemes, appear to be a convenient tool to prove mathematically the convergence of discrete coupling procedure between various models, even for high order situations. Incidentally, the perspective (1A) is here a point of interest to activate the required stability estimates.

2.3 Propagative and glancing wavepackets

The previous last discussions are restricted to the case of finite difference schemes without extended stationary modes, thus satisfying the assumption (2.5). When extended stationary modes are present, new discrete parasitic phenomenon may occur, sometimes in an unacceptable way, but sometimes only reducing the order of convergence of the overall scheme. This is the case only if the damping properties of the scheme is not strong enough to preclude the presence of neutrally stable modes. Diffusive schemes (in the sense of Lax) do not enter this framework but when solving purely dispersive phenomenon with high accuracy or with discrete preservation of energies, it appears crucial to handle with this kind of considerations, let us mention for example the works [KN20; BCN20], where discrete transparent boundary conditions are computed in that aim or in [BGK+22], where perfect matched layers are used for mixed hyperbolic-dispersive equations. Actually, only a few situations are clearly understood at the theoretical level and asymptotic expansions based on discrete boundary layer and/or propagative wave-packets is usefull for analyzing the general situation.

Once again, only *stationary* modes are considered here, associated to the amplification factor $z = 1$, but many ideas and results can likely be naturally extended to other unitary modes $|z| = 1$.

Propagative wavepackets The ansatz for an appropriate expansion in the more general case has the form of a WKB asymptotic expansion with several terms

$$u_{n,j}^{\mathsf{app}} = u_{\mathsf{int}}(t^n, x_j) + \sum_{\kappa \in \mathsf{K}_+} {u_{\mathsf{int}}}^{(\kappa)}(t^n, x_j)\kappa^j + u_{\mathsf{bl}}(t^n, j) \tag{2.13}$$

where K_+ denotes the set of $\kappa \in \mathbb{S}^1$ satisfying $\mathcal{A}(\kappa) = 0$ and having positive (real) group velocity $\kappa\mathcal{A}'(\kappa) > 0$. The associated profile ${u_{\mathsf{int}}}^{(\kappa)}$ then solves the incoming transport equation $\partial_t u + \kappa\mathcal{A}'(\kappa)\partial_x u = 0$ with zero initial data and the "appropriate" incoming boundary value obtained from the coupling of waves concerned by $\mathbb{E}_{\mathsf{s}}(1)$. The procedure is then entirely similar to the solving of the boundary problem (2.7).

Glancing modes As already discussed at the beginning of this chapter, a particular situation appears when the group velocity $\kappa\mathcal{A}'(\kappa)$ is zero. Different scales in space then have to be considered so as to correctly define the profile equation for the asymptotic expansion.

Lemma 19. *Assume the scheme (2.2) to be consistant and ℓ^2-stable. Let $\kappa \in \mathbb{S}^1$ be a root of $\mathcal{A}$ of order $g \in \mathbb{N}^\star$ exactly, meaning: $\mathcal{A}(\kappa) = \mathcal{A}'(\kappa) = \ldots = \mathcal{A}^{(g-1)}(\kappa) = 0$, and $\mathcal{A}^{(g)}(\kappa) \neq 0$. Then the following property holds, according to the remainder in the division of g by 4:*

$$\begin{cases} \mathsf{Re}\left(\kappa^g\, \mathcal{A}^{(g)}(\kappa)\right) > 0 & \text{if } g \bmod 4 = 0, \\ \mathsf{Im}\left(\kappa^g\, \mathcal{A}^{(g)}(\kappa)\right) = 0 & \text{if } g \bmod 4 = 1, \\ \mathsf{Re}\left(\kappa^g\, \mathcal{A}^{(g)}(\kappa)\right) < 0 & \text{if } g \bmod 4 = 2, \\ \mathsf{Im}\left(\kappa^g\, \mathcal{A}^{(g)}(\kappa)\right) = 0 & \text{if } g \bmod 4 = 3. \end{cases}$$

As for the very low order equivalent PDEs, the ℓ^2-stability of the scheme precludes the ill-posedness for the equivalent equation of the conjugated scheme around $(1, \kappa)$, that is

$$\frac{\partial v}{\partial t} + \frac{1}{g!}\kappa^g\, \mathcal{A}^{(g)}(\kappa)\frac{\partial^g v}{\partial y^g} = 0. \tag{2.14}$$

The is the main idea behind the previous lemma. With the same notations, the corresponding expected WKB asymptotic expansion involves now a term of the form $v_\kappa(t^n, j\Delta x^{1/g})\kappa^j$ where v_κ solves the previous PDE problem (see Figure 2.4 for examples of such effective modulated envelope modes). For solving the similar boundary problem (2.8), an invertibility result is required that reduces actually to a dimension argument for Vandermonde-like matrices. The result is the following, that corresponds actually to the decomposition of the space $\mathbb{E}_{\mathsf{s}}(1)$ from the different multiplicities occuring for $|\kappa| = 1$, in the spirit of Hersh lemma [Her63].

Proposition 20 (Extension of the counting lemma to general stationary modes)**.** *Assume the scheme (2.2) to be consistant and ℓ^2-stable and consider the roots of $\mathcal{A}$.*
Let $\mathsf{K}_+ = \{\kappa \in \mathbb{S}^1, /\mathcal{A}(\kappa) = 0,\ \kappa\mathcal{A}'(\kappa) > 0\}$, $\mathsf{K}_- = \{\kappa \in \mathbb{S}^1, /\mathcal{A}(\kappa) = 0,\ \kappa\mathcal{A}'(\kappa) < 0\}$ and $n_+^{(1)} := n_+$, $n_-^{(1)} := n_-$ the cumulated multiplicities of the corresponding simple roots.
For any $g \in \mathbb{N}^\star$, consider also the integer $n^{(g)} \in g\mathbb{N}$ equal to the sum of the multiplicities of the roots having multiplicity $g \in \mathbb{N}^\star$ exactly. For odd values of $g \in \mathbb{N}^\star$, let us separate the (possibly

empty) set $\mathsf{K}^{(g)}$ *of the roots having multiplicity* g *into two disjoint sets, namely* $\mathsf{K}_+^{(g)}$ *and* $\mathsf{K}_-^{(g)}$*, according to the sign of the quantity* $\mathsf{Re}\left(\kappa^g\,\mathcal{A}^{(g)}(\kappa)\right)$*. Let us denote, in accordance with this partition,* $n_+^{(g)} \in g\mathbb{N}$ *and* $n_-^{(g)} \in g\mathbb{N}$ *their cumulated multiplicities so that* $n^{(g)} = n_+^{(g)} + n_-^{(g)}$*. Then one has:*

$$r = n_{\mathbb{D}} + \sum_{g\,\text{even}} \tfrac{1}{2} n^{(g)} + \sum_{g \bmod 4 = 1} \left(\tfrac{g+1}{2g} n_+^{(g)} + \tfrac{g-1}{2g} n_-^{(g)}\right) + \sum_{g \bmod 4 = 3} \left(\tfrac{g-1}{2g} n_+^{(g)} + \tfrac{g+1}{2g} n_-^{(g)}\right),$$
$$p = n_{\mathcal{U}} + \sum_{g\,\text{even}} \tfrac{1}{2} n^{(g)} + \sum_{g \bmod 4 = 1} \left(\tfrac{g-1}{2g} n_+^{(g)} + \tfrac{g+1}{2g} n_-^{(g)}\right) + \sum_{g \bmod 4 = 3} \left(\tfrac{g+1}{2g} n_+^{(g)} + \tfrac{g-1}{2g} n_-^{(g)}\right).$$

Actually, the existence of schemes with high order glancing scheme may seem to be hypothetical since many tangential and stability constraints are present. We give hereafter an existence result for such schemes at a given glancing order, with glancing modes of order g for the value $\kappa = -1$ to keep easily real coefficients. This result is up to now, not thought as being of a real importance for the applications but enables however some simple illustrations for the possible glancing effect at discrete boundaries.

Proposition 21 (Existence and uniqueness of glancing schemes with minimal width)**.** *For any integer* $g \geq 2$*, there exists a finite difference scheme satisfying the following properties:*

- *consistency with* $\partial_t u + a\partial_x u = 0$ *with* $a < 0$*,*
- *stability in the sense of the strong property* $\mathsf{Re}\ \mathcal{A}(\mathbb{S}^1) \cap \{\mathsf{Re}\, z > 0\} = \emptyset$*,*
- *existence of a* g*-glancing mode.*

Moreover, there is uniqueness of such a scheme when minimizing the width of the stencil $r+p+1$*. The concerned scheme is described as follows. We define* $r = \frac{g-1}{2}$ *for* g *odd or* $r = \frac{g}{2}$ *for* g *even, and then set* $p = g + 1 - r$ *so that the stencil has width* $g + 2$*. The corresponding scheme is then*

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + \frac{a}{2^g \Delta x}\left[-u_{j-r}^n + \sum_{\ell=1}^{g} \left(\binom{g}{\ell - 1} - \binom{g}{\ell}\right) u_{j-r+\ell}^n + u_{j-r+g+1}^n\right] = 0, \quad j \geq r. \tag{2.16}$$

This result is proved by means of elementary polynomial algebra. From the consistency properties, the space operator may be factorized under the following form

$$\mathcal{A}(\kappa) = \frac{a}{2^g}(\kappa + 1)^g (\kappa - 1)\kappa^q,$$

where q is an appropriate value, uniquely defined from the Cauchy ℓ^2-stability property and local analysis close to the particular frequencies $\kappa = 1$ and $\kappa = -1$. Rather that a detailed proof, we prefer here to give for the interested reader the list of the very first schemes for the Euler forward time discretization. The Figure 2.3 represents the amplification factor of the schemes with parameter g and q. The green curves corresponds to stable schemes and the red curves to unstable schemes.

- Scheme $g = 2$, with $q = -1$, $r = 1$, $p = 2$:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + \frac{a}{4\Delta x}[-u_{j-1} - u_j + u_{j+1} + u_{j+2}] = 0, \quad j \geq 1$$

- Scheme $g = 3$, with $q = -1$, $r = 1$, $p = 3$:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + \frac{a}{8\Delta x}[-u_{j-1} - 2u_j + 2u_{j+2} + u_{j+3}] = 0, \quad j \geq 1$$

- Scheme $g = 4$, with $q = -2$, $r = 2$, $p = 3$:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + \frac{a}{16\Delta x}[-u_{j-2} - 3u_{j-1} - 2u_j + 2u_{j+1} + 3u_{j+2} + u_{j+3}] = 0, \quad j \geq 2$$

- Scheme $g = 5$, with $q = -2$, $r = 2$, $p = 4$:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + \frac{a}{32\Delta x}[-u_{j-2} - 4u_{j-1} - 5u_j + 5u_{j+2} + 4u_{j+3} + u_{j+4}] = 0, \quad j \geq 2$$

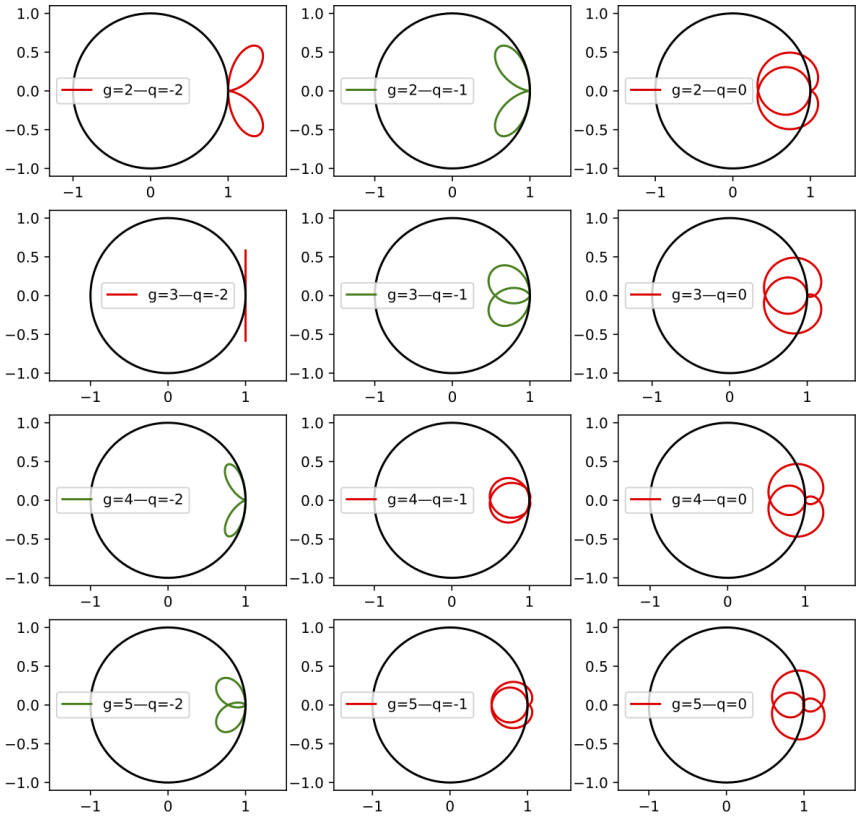

Figure 2.3: Amplification factor for glancing schemes $\Delta t/\Delta x = 0.9$.

The Figure 2.4 represents the parasitic glancing mode close to the boundary for various schemes (2.16). The numerical solutions corresponds to the initial data 1 for the outgoing transport problem and homogeneous Dirichlet boundary scheme. Now this is not surprinsing to numerically observe the expansion $u_j^n = 1 + (-1)^j v_\kappa(t, j\Delta x^{1/g})$ where the modulated profile v_κ solves (2.14) with appropriate boundary conditions and initial data.

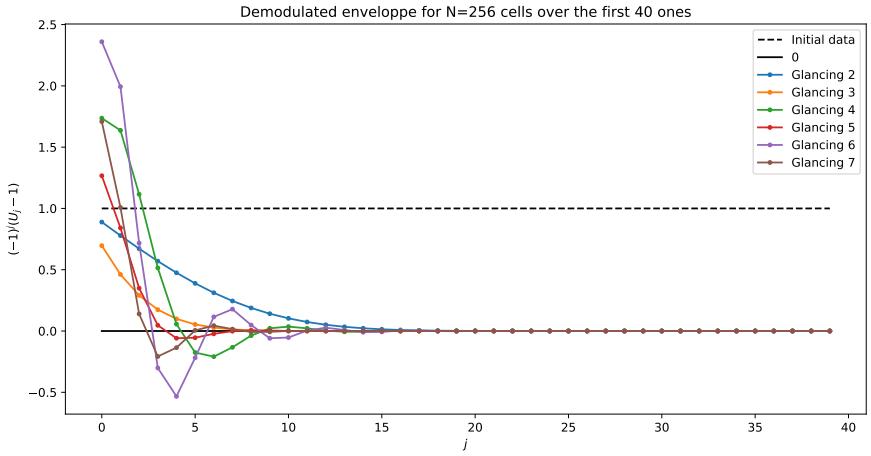

Figure 2.4: Demodulated effective envelope $(-1)^j(u_j - 1) \simeq v_\kappa(t, j\Delta x^{1/g})$ for homogeneous Dirichlet boundary condition with the numerical schemes (2.16).

2.4 Perspectives

(2A) Mixing different strategies. For any stable discrete IBVPs, an adapted improved convergence analysis is available from appropriate asymptotic expansions. They involve corrective terms for measuring the possible localized or propatating loss of consistency and thus of rate of convergence. They have to be determined from the various possible interactions: boundary scheme with interior scheme, initial condition with interior scheme, space corners, etc. Two complementary strategies are possible to improve the consistency.

1. Change g by removing appropriate terms without changing the boundary scheme. The benefits are that the strong stability result is not affected. This is the idea used to obtain ILW methods from Dirichlet in the Chapter 1.

Or 2. Change B so as to purge the expansion from any spurious modes (layers, propagative, glancing, etc.) The method of discrete transparent boundary condition as Ehrhardt and Arnold [EA01], Arnold, Ehrhardt, and Sofronov [AES03], Besse, Ehrhardt, and Lacroix-Violet [BEL16], Besse, Coulombel, and Noble [BCN20], and Kazakova and Noble [KN20] is precisely constructed in that way and guarantees at the same time the strong stability (see Coulombel [Cou19]). See also Ehrhardt [Ehr10] for absorbing boundary conditions with dissipative features.

This is not clear how a general mixing between the two strategies could be realized. In particular in the Chapter 1, we slightly modify the ILW method to the SILW, what changes g and B at the same time.

(2B) Explicit parasitic modes in general geometries. Another complementary project is concerned with the geometrical aspects of the boundary. The whole discussion in the manuscript is restricted to the poor onedimensional half-problem, even if archetypal of many principles. Some attention has been put on the stability theories for corners by Osher [Osh74a; Osh74b] and more recently by Benoit [Ben16; Ben17] and Besse,

Coulombel, and Noble [BCN20], and similarly for interval/strip domains by Trefethen [Tre85] and more recently Benoit [Ben20; Ben21] and also Inglard, Lagoutière, and Rugh [ILR20] where "ghost" solutions are analyzed for implicit schemes. The use of WKB-like expansions to understand the possible appearance of parasitic (even unstable) modes in such situations seems to be a promising perspective.

(2C) Space inhomogeneities. At the interface of the pseudospectral analysis (Perspective (1B)) and boundary layer expansions is found the analysis of pseudomodes. This topic has been studied for example by Trefethen and Chapman [TC04] and Trefethen [Tre05] that construct asymptotic localized pseudomodes for twisted-Toeplitz (corresponding to finite differences schemes on non-homogeneous domains). On the side of numerical computations, spurious effects are also observed, called discrete q-waves in the literature, see Le Roux [Le 12], Sengupta [Sen12], Marica and Zuazua [MZ15] and Biccari, Marica, and Zuazua [BMZ20]. By means of discrete pseudodifferential analysis, see Chodosh [Cho11], Schochet [Sch14], Faou and Grébert [FG21], a new analysis could be promising.

CHAPTER 3

Hyperbolic relaxation models

The present chapter is devoted to the numerical approximation of hyperbolic relaxation models, in the sense of Chen, Levermore, and Liu [CLL94], in the linear case with boundaries. The PDE part is subject to standard stability properties, as strict hyperbolicity of the first order terms and subcharacteristic dissipativity of the stiff zeroth order relaxation term. In addition the considered boundary condition satisfies the Kreiss-Lopatinskii condition for the non-stiff hyperbolic part. It is known from Xin and Xu [XX00] that a reinforced *Stiff Kreiss Condition* is actually necessary to preclude boundary instabilities during the limiting relaxation process. The aim is to develop numerical schemes and analyze carefully their uniform stability with respect to the relaxation parameter (and the discretization parameters as well of course). To do that, we make use of the Summation-by-Parts techniques, related to the energy method, and of the discrete Z-transform. Another strategy is based on the construction of discrete transparent boundary conditions but only for the extra discrete-boundary conditions. The presentation hereafter is related to the publications[1] [A/BNS20] and [P/BNS], obtained from the Master internship and then the PhD work of Thi Hoai Thuong Nguyen during the years 2017–2020, student co-advised with Nicolas Seguin.

[A/BNS20] B. Boutin, T. H. T. Nguyen, and N. Seguin. A stiffly stable semi-discrete scheme for the characteristic linear hyperbolic relaxation with boundary. *ESAIM Math. Model. Numer. Anal.* 54 5:1569–1596, 2020.

[P/BNS] B. Boutin, T. H. T. Nguyen, and N. Seguin. A stiffly stable fully discrete scheme for the damped wave equation using discrete transparent boundary condition.

[1] We warn the reader that some notations have been changed between this manuscript and the cited publications, so as to unify the current presentation.

3.1 Hyperbolic relaxation models with boundaries

Hyperbolic relaxation models Relaxation effects arise in many physical models. In particular several dynamical effects may coexist but at different time scales so that this is conceivable to study the reduced dynamic obtained for a partial equilibrium process. For example in the kinetic theory of the Boltzmann equation, if gases are supposed to be very close to local thermodynamic equilibrium (Maxwellian) then they are submitted to simpler fluid models. This assumption precisely consists of a limiting relaxation process in a partial direction of the dynamic, i.e. neglecting a time scale compared to another one. The usual method to investigate the subsequent limiting problem is based on Chapman-Enskog expansions. Liu [Liu87] and Chen, Levermore, and Liu [CLL94] study the stability of the limiting process in the case of hyperbolic conservation laws with relaxation and show the importance therein of the Whitham [Whi74] subcharacteristic condition. Indeed, it guarantees the dissipativity of the source term and activates the strong well-posedness of the system uniformly in the relaxation parameter and the hyperbolicity of the limiting equation.

Considering a vectorial unknown $U \in \mathbb{R}^N$, we are interested in relaxation systems of the form:

$$\begin{aligned} &\partial_t U + A\partial_x U = \frac{1}{\epsilon} S(U), && x \in \mathbb{R}_+,\ t \geq 0, \\ &\mathsf{B}U(t,0) = g(x), && t \geq 0, \\ &U(0,x) = f(x), && x \in \mathbb{R}_+. \end{aligned} \tag{3.1a}$$

The matrices A of size $N \times N$ and B are assumed to validate the one-dimensional Kreiss-Lopatinskii condition described in Chapter 1. On the other side, the relaxation operator $S : \mathbb{R}^N \to \mathbb{R}^N$ is subject to the usual dissipative requirements for relaxation problems. More precisely the results presented in this Chapter is restricted to the two-dimensional linear problem with the following matrices (where $a > 0$):

$$U = \begin{pmatrix} u \\ v \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 \\ a^2 & 0 \end{pmatrix}, \quad S = \begin{pmatrix} 0 & 0 \\ b & -1 \end{pmatrix}. \tag{3.1b}$$

$$\mathsf{B} = (\mathsf{B}_u, \mathsf{B}_v), \text{ with } \mathsf{B}_u \geq 0. \tag{3.1c}$$

Let us notice that mimicking the physical relaxation dynamics, some numerical method emerged called relaxation method, and based on a dimensional extension of the unknown and the modification of the dynamical system with a simpler linear hyperbolic structure, and additional stiff low order source terms to recover in the limit the original problem. For example, Jin and Xin [JX95] proposed to approximate the solution of the hyperbolic conservation law $\partial_t u + \partial_x f(u) = 0$ by designing schemes from the relaxation system of the form (3.1a) with additional relaxation variable v and:

$$A = \begin{pmatrix} 0 & 1 \\ a^2 & 0 \end{pmatrix}, \quad S(U) = \begin{pmatrix} 0 \\ f(u) - v \end{pmatrix} \tag{3.2}$$

In that case the subcharacteristic condition for the stiff well-posedness of the system (without

boundaries) and the convergence of its solution to equilibrium reads $a > |f'(u)|$ for all values of u involved. Many results are available from the works by Yong [Yon99b; Yon01; Yon02], Liu and Yong [LY01], and Yong and Jäger [YJ05].

In the framework of the linear relaxation system (3.1), the subcharacteristic condition is $a \geq b$ and the equilibrium solution is $U_{\mathsf{eq}} = (u_{\mathsf{eq}}, v_{\mathsf{eq}})$ solution to the algebraico-differential IBVP problem, that is well-posed:

$$\partial_t u_{\mathsf{eq}} + b\,\partial_x u_{\mathsf{eq}} = 0, \qquad v_{\mathsf{eq}} = b\,u_{\mathsf{eq}}, \tag{3.3}$$

with boundary condition

$$\begin{cases} (\mathsf{B}_u + b\mathsf{B}_v)u_{\mathsf{eq}}(t,0) = g(t) & \text{if } b > 0, \\ \text{no b.c.} & \text{if } b \leq 0. \end{cases} \tag{3.4}$$

Relaxation and boundaries Let us first introduce the following definition.

Definition 22. *Assume that the IBVP problem (3.1a) is such that for any $\epsilon > 0$ and any time $T > 0$, there exists a constant $C > 0$ such that for any $f \in L^2(\mathbb{R}^+)$ and $g \in L^2(\mathbb{R}^+)$, the corresponding solution satisfies*

$$\int_0^T \|U(t,\cdot)\|^2_{L^2(\mathbb{R}^+)}\,\mathsf{d}t + \int_0^T |U(t,0)|^2\,\mathsf{d}t \quad \leq \quad C\left(\|f\|^2_{L^2(\mathbb{R}^+)} + \int_0^T |g(t)|^2\,\mathsf{d}t\right).$$

Then the IBVP is said to be well-posed. *If in addition the constant C can be chosen independent of $\epsilon > 0$ (small) then the IBVP is said to be* stiffly well-posed.

As explained many times in the previous chapters, the boundary condition has to satisfy the Uniform Kreiss-Lopatinskii Condition for the hyperbolic part. For (3.1b), thanks to the Riemann invariants $au \pm v$, for the corresponding characteristic velocities $\pm a$, the UKLC, under which the problem (3.1) is well-posed, reads

$$\mathsf{B}_u + a\mathsf{B}_v \neq 0. \tag{3.5}$$

Now, it appears that the above UKLC is not completely sufficient to guarantee the limiting relaxation process and to preclude the existence of unstable relaxation boundary layers. Yong [Yon99a] proposed a Generalized Kreiss Condition (GKC) for general multi-dimensional linear constant coefficient relaxation systems, or one-dimensional nonlinear systems, with non-characteristic boundaries. This condition enables uniform (in ϵ) stability estimates and then justifies the reduced boundary condition for the corresponding equilibrium system. In the Jin-Xin system (3.2) with boundary condition (3.1c), Xin and Xu [XX00] (see also [XX02; XX04]) identify and rigorously justify the necessary and sufficient condition, called Stiff Kreiss Condition (SKC) for the stiff well-posedness. For higher dimensional linear relaxation problems, the SKC admits a determinantal form, and for the two-dimensional linear system (3.1) under interest, it reads

$$\left[\mathsf{B}_v = 0 \text{ or } \mathsf{B}_u\mathsf{B}_v^{-1} \notin [-a, -\tfrac{b+|b|}{2}]\right]. \tag{3.6}$$

Some insight into the SKC condition Following the classical methodology for first order IBVPs, the SKC condition is analyzed through the normal mode analysis and then the Laplace transform. Actually their form is modified so as to catch the scale involved in relaxation layer: $U^\epsilon = \mathsf{e}^{\tau t/\epsilon}\varphi(x/\epsilon)$ with $\mathsf{Re}\,\tau > 0$ and $\varphi \in L^2(\mathbb{R}_+, \mathbb{R}^2)$. Again a separation results is available. The modified UKLC condition is then concerned with the stable linear subspace $\mathbb{E}_\mathsf{s}(\tau)$ associated to the matrix $A^{-1}(S - \tau I)$. Its not-uniform version reads:

$$\mathsf{Ker}\,\mathsf{B} \cap \mathbb{E}_\mathsf{s}(\tau) = \{0\}.$$

Asymptotic expansions In addition to [Yon99b], the study by Xin and Xu [XX00] also covers the characteristic case with $b = 0$ and provides optimal asymptotic expansions for the limit process, including boundary and/or initial layers. The system (3.1) with $b = 0$ is also known as the *damped wave equation* and will be the subject of the forthcoming numerical study. In that case, the SKC in [XX00] then simply reduces to the following inequality:

$$\mathsf{B}_u + a\mathsf{B}_v > 0. \tag{3.7}$$

Under the subcharacteristic condition and the SKC (3.6), being given compatible data and boundary smooth data (in $H^2(\mathbb{R}_+)$ spaces), the asymptotic expansion towards the equilibrium solution U_eq is of the form

$$U^\epsilon(t,x) = U_\mathsf{eq}(t,x) + \begin{cases} 0 + h.o.t & \text{if } b > 0, \\ U^\mathsf{bl}(t, \frac{x}{\epsilon}) + h.o.t & \text{if } b < 0, \\ U^\mathsf{bl}(t, \frac{x}{\sqrt{\epsilon}}) + h.o.t & \text{if } b = 0. \end{cases} \tag{3.8}$$

The high order terms involve higher power of the relaxation parameter ϵ and are estimated in the time-Laplace norm topology (see Chapter 1).

For the interested reader, a review of these results is done by Zhang and Wang [ZW04] with additional source terms and the corresponding stiff full estimates. More recent works by Zhou and Yong [ZY21; ZY20] are concerned with the case of general hyperbolic linear relaxation systems with characteristic boundaries, either for the original hyperbolic part, or for the limiting hyperbolic part. Other works on the topic are the following ones: [CJL+14; Xu04; Ye04].

3.2 Stiffly stable schemes for the damped wave equation

From now on, the continuous IBVP under consideration is the following

$$\begin{aligned} &\partial_t u^\varepsilon + \partial_x v^\varepsilon = 0, \\ &\partial_t v^\varepsilon + a^2 \partial_x u^\varepsilon = -\varepsilon^{-1} v^\varepsilon, \end{aligned} \tag{3.9}$$

with the appropriate initial and boundary conditions from (3.1c). A first scheme is introduced at the semi-discrete level and based on the central scheme for the flux term, thus has the form

$$\begin{aligned} &\frac{\partial}{\partial t}U_j(t) + (\mathcal{Q}U)_j(t) = \frac{1}{\epsilon}SU_j(t), \quad && j \geq 1,\ t \geq 0, \\ &U_j(0) = f_j, && j \geq 0, \\ &BU_0(t) = g(t), && t \geq 0, \end{aligned} \tag{3.10a}$$

with

$$(\mathcal{Q}U)_j = \frac{1}{2\Delta x}A(U_{j+1} - U_{j-1}), \quad j \geq 0.$$

Following the summation by parts method from Strand [Str94], the energy method can be adapted to the central operator by choosing the extrapolated ghost value $U_{-1} = 2U_0 - U_1$, so that finally the boundary scheme is $(\mathcal{Q}U)_0 = \frac{1}{\Delta x}(U_1 - U_0)$. The convenient discrete integration by part (see [GKO13]) then applies for the scalar product

$$\langle U, V\rangle_{\Delta x} := \frac{\Delta x}{2}\langle U_0, V_0\rangle + \Delta x \sum_{j\geq 1}\langle U_j, V_j\rangle.$$

The physical boundary condition in (3.10a) determines one scalar boundary unknown in U_0 but the numerical scheme requires an additional second scalar unknown so as to fully define $U_0(t)$. This additional scalar boundary discrete boundary condition is proposed under the following ODE form

$$\Gamma\left(\frac{\partial}{\partial t}U_0(t) + (\mathcal{Q}U)_0(t)\right) = \frac{1}{\epsilon}\Gamma SU_0(t), \quad t \geq 0, \tag{3.10b}$$

where the rank one matrix is $\Gamma = \begin{pmatrix} -a^2\mathsf{B}_v & \mathsf{B}_u \end{pmatrix}$. The scheme is then fully defined.

The proof of the stiff stability of the semi-discrete scheme (3.10) is based on two ingredients and a superposition argument. Firstly, the summation by parts enables the treatment of homogeneous boundary conditions. Secondly, the case with non-homogeneous boundary conditions and homogeneous initial data is handled by means of the Laplace transform and the "stiff UKLC" estimate adapted to the problem (uniform in ϵ and Δx resolvent estimate).

Our result is the following:

Theorem 23 (Theorem 1.2 in [A/BNS20])**.** *Under the strict dissipativity condition* $\mathsf{B}_u\mathsf{B}_v > 0$*, for all* $T > 0$*, there exists* $C > 0$ *such that for all* $f \in \ell^2(\mathbb{N}, \mathbb{R}^2)$ *and all* $g \in \mathcal{C}^1(\mathbb{R}^+, \mathbb{R}) \cap L^2(\mathbb{R}^+, \mathbb{R})$ *the solution* $(U_j)_{j\geq 0}$ *to* (3.10) *satisfies:*

$$\int_0^T |U_0(t)|^2\,dt + \int_0^T \sum_{j\geq 0}\Delta x|U_j(t)|^2\,dt \leq C_T\left(\sum_{j\geq 0}\Delta x|f_j|^2 + \int_0^T |g(t)|^2\,dt\right),$$

where the constant C_T *is independent of* f*,* g *and of the relaxation parameter* $\epsilon \in (0, +\infty)$ *and the discretization parameter* $\Delta x \in (0, 1]$*.*

We now briefly explain some steps of the proof of that theorem.

Summation by parts method We first consider homogeneous boundary condition $g = 0$ and use the energy method. It enjoys the effect of the SBP method,with the symmetrizer $H = \mathsf{diag}(a^2, 1)$, so as to obtain the identity

$$\partial_t \langle U, HU \rangle_{\Delta x} + 2a^2 \frac{\mathsf{B}_u}{\mathsf{B}_v} u_0^2 + \frac{\Delta x}{\epsilon} v_0^2 = -\frac{2\Delta x}{\epsilon} \sum_{j \geq 1} v_j^2,$$

The dissipativity of the boundary condition B then follows under the condition:

$$2a^2 \frac{\mathsf{B}_u}{\mathsf{B}_v} + \frac{\Delta x}{\epsilon} \left(\frac{\mathsf{B}_u}{\mathsf{B}_v} \right)^2 > 0.$$

Then, using in addition the dissipative character of relaxation, the following stiff bound follows:

$$\langle U, HU \rangle_{\Delta x} + C \int_0^T |U_0|^2 \mathsf{d}t \leq \langle f, Hf \rangle_{\Delta x}.$$

Actually, if $\mathsf{B}_u \mathsf{B}_v > 0$ then the above condition is true, else if $\mathsf{B}_u \mathsf{B}_v < 0$, having a bound of the form $\Delta x \geq \delta \epsilon$ also suffices to conclude.

Normal mode analysis and stiff UKLC estimate As for the continuous case, the homogeneous initial data case $f = 0$ can be handled by the UKLC algebra and Laplace transform. However, the presence of the discretization in space now increase the dimension of the stable linear subspace associated to some time frequency τ with $\mathsf{Re}\, \tau > 0$. This is the point where the supplemented artificial boundary condition (3.10b) is involved. In [A/BNS20], we prove that the condition $\mathsf{B}_u \mathsf{B}_v > 0$ suffices to guarantee the uniform determinantal version of the UKLC here. In addition, some numerical experiments show that this UKLC is not always satisfied under the continuous SKC (3.6) only.

3.3 Extensions

- A first extension is concerned with the implicit time discretization of the previous semi-discrete scheme. The stiff stability is then obtained again under the strict dissipativity condition $\mathsf{B}_u \mathsf{B}_v > 0$ with an inverse "CFL" condition of the form $\Delta x < 3a\Delta t/8$.

- A second extension is again for the implicit time discretization, but now with the upwind flux in space, in the sense of the Riemann invariants of the hyperbolic part of the problem. The analysis is then successfully conducted only for the energy method part, i.e. for homogeneous boundary data g.

- A third extension is proposed and intended to obtain a fully discrete scheme that would be stiffly stable exactly in the same regime than the underlying PDE, that is under SKC (3.6) and not only the subclass $\mathsf{B}_u \mathsf{B}_v > 0$. Again the interior scheme is the implicit centered one. The idea now is to replace the ghost value U_{-1} by a convenient coming from the discrete transparent boundary technique, but only for the artificial part (3.10b). Now the eigenstructure involves four waves but only two matricial eigenprojectors Φ_-,

Φ_+ related to the continuous problem. The corresponding eigenvalues $\pm\kappa_\pm(z)$ satisfy $|\kappa_+| > 1$ and $|\kappa_-| < 1$. In the Z-transformed version of the scheme, the transparent boundary extension then simply reads:

$$\hat{U}_{-1} = \kappa_+(\Phi_- - \Phi_+)\hat{U}_0.$$

Coupled with the proposed strategy the final boundary scheme has finally the non-local in time-boundary form:

$$\Gamma\left[\frac{1}{\Delta t}(U_0^{n+1} - U_0^n) + \frac{1}{2\Delta x}\left(U_1^{n+1} - \sum_{k=0}^{n+1}\mathfrak{C}_{n+1-k}U_0^k\right) - \frac{1}{\epsilon}SU_0^{n+1}\right] = 0. \tag{3.11}$$

In [P/BNS], we can prove that under the SKC (and not only under the subclass of strictly dissipative conditions), the scheme is stiffly strongly stable (i.e. in the sense of zero initial data and nonzero boundary data), provided the previous inverse "CFL" condition is satisfied. A part of the analysis is still work in progress.

3.4 Perspectives

(3A) The main extension of this work is to propose asymptotic preserving schemes in the sense of Jin [Jin99] that encompass the boundary treatment for general situations with change of sign or vanishing in the characteristic velocities. The stiff stability under the full stiff Kreiss condition is then a crucial features in order to guarantee that the numerical scheme is not affected by unstable discrete boundary layers and is able to encompass in a correct way the (stable) relaxation boundary layer. A submitted work [P/ABC] with N. Crouseilles and M. Anandan deals with the development of accurate numerical schemes in the slightly different framework of kinetic relaxation equations in the diffusive scaling. The use of adapted half-moments strategies through micro-macro decompositions is often limited to first order in space. Increasing the order in space then requires appropriate treatments for the relaxation layer.

CHAPTER 4

Infinite dimensional QR eigenvalue method

> [The QR algorithm is] one of the most remarkable algorithm in scientific computing.
>
> Strang [Str80] *Linear algebra and its applications.*

The names of Rutishauser, Wilkinson, Lanczos and Francis are at the heart of the development of numerical algorithms for spectra computations mainly within the 1950s and the 1960s. The foundations of their results are found in some older analysis due to Jacques Hadamard in his thesis. It is concerned with characterizations of poles of meromorphic functions that follow themselves older results by Daniel Bernoulli. More recently, after the development of the first numerical algorithms for spectra approximations, two main directions have emerged. A first one is concerned with the improvement of the cost and the quality of the effective computations: reduction of the cost, improvement of the rate convergence, dimension reduction, etc. For a more complete overview of the developments of eigenvalue computational algorithm during the 20ts century, we refer the reader to the work by Golub and van der Vorst [GvdV00] and Golub and Uhlig [GU09]. A second one is motivated by several theoretical applications towards operator theory, based on the deep relationship between these algorithms and some infinite-dimensional dynamical systems (see Chu [Chu08]).

We present hereafter first a quick overview of the previously mentioned historical aspects, partially inspired from Gutknecht and Parlett [GP11]. After that, a discussion is proposed on some recent results [A/BR17] obtained a few years ago in a collaboration with N. Raymond. The topic is about formalizing in an unified way a family of isospectral differential systems acting on Hilbert-Schmidt operators. The abstract method is able to determine asymptotically the spectrum of such operators, with identified convergence rates, even in infinite dimension.

[A/BR17] B. Boutin and N. Raymond. Some remarks about flows of Hilbert-Schmidt operators. *J. Evol. Equ.* 17 2:805–826, 2017.

4.1 A little history about the QR method

The beginnings To make this presentation more concrete, let us consider a given matrix $A \in \mathcal{M}_n(\mathbb{C})$ whose spectral properties are under interest. Let us also introduce two vectors $x_0 \in \mathbb{C}^n$ and $y_0 \in \mathbb{C}^n$ and study the associated meromorphic function f defined from the resolvent of A by

$$f(z) := \langle y_0, (zI - A)^{-1} x_0 \rangle.$$

The poles of f are located among the eigenvalues of A. Namely, using its adjugate, the function f alternatively takes the form of a rational function

$$f(z) = \frac{\langle y_0, \mathsf{adj}(zI - A)x_0 \rangle}{\det(zI - A)}.$$

The denominator is the characteristic polynomial of A and the numerator is a polynomial whose degree is less than $n-1$. Actually the two polynomials may share some zeros, depending on particular choices for the vectors x_0, y_0. For z large enough i.e. $|z| > \rho(A)$, the function f admits an expansion as the series

$$f(z) = \sum_{m=0}^{\infty} \frac{a_m}{z^{m+1}},$$

where the complex coefficients $a_m = \langle y_0, A^m x_0 \rangle$ (also known as Schwarz constants) define the sequence of moments associated to the problem of determining the eigenvalues of A / poles of f. From the Cayley-Hamilton formula, the whole sequence $(a_m)_{m \geq 0}$ then solves a (constant coefficients) linear recurrence relation whose characteristic polynomial P necessarily divides the characteristic polynomial χ_A of A. The polynomial P again depends on possible particular properties of x_0, y_0 with respect to the eigenstructure of A. The d'Alembert's ratio test (d'Alembert [dAle68]) for convergent series, also known from the *method by Daniel Bernoulli [Ber32]* for roots, then applies. Assuming for a while, as usual, that A has a dominant root λ_1 and, in addition, that λ_1 is found among the roots of P, then the ratio a_{m+1}/a_m is convergent with limit λ_1.

The above method is clearly reminiscent of and related to the *power method* for computing the dominant eigenvalue. In his thesis, Hadamard [Had92] proposed an extension of Bernoulli's method so as to characterize successively any of the other poles of a meromorphic function from its moments. He first writes there *"Si la seule singularité située sur le cercle est un pôle, simple ou multiple, l'affixe de ce point est donnée par la limite du rapport a_m/a_{m+1}. [...] Cette condition nécessaire est aussi suffisante."*. To handle with the other roots, Hadamard makes use of the sequence of the associated *Hankel determinants*. We recall here that the Hankel determinant of order k at step m, associated to a sequence $(a_m)_{m \geq 0}$ is the quantity

$$H_m^{(k)} = \begin{vmatrix} a_m & a_{m+1} & \dots & a_{m+k-1} \\ a_{m+1} & a_{m+2} & \dots & a_{m+k} \\ \vdots & \vdots & & \\ a_{m+k-1} & a_{m+k} & \dots & a_{m+2k-2} \end{vmatrix}.$$

In particular, the sequence $(a_m)_{m\geq 0}$ solves a linear recurrence relation if and only if the sequence $(H_0^{(k)})_{k\geq 0}$ is null starting from some rank k_0. This is clearly the case here. The point of interest here lies in the behaviour of the quantity $H_m^{(k)}$ for large values m, but now fixed size k of the determinants. In this spirit, Hadamard proved the following result. Assume the modulus-separation property $|\lambda_1| > \ldots > |\lambda_k| > \Lambda > |\lambda_{k+1}|$ for some $\Lambda > 0$, then as $m \to \infty$:

$$H_m^{(k)} = (\lambda_1 \ldots \lambda_k)^m \left[1 + O\Big(\frac{\Lambda}{|\lambda_k|}\Big)^m\right].$$

From this result, this is then possible to identity any of the roots as being the limit of (heavily) computable quantities. Indeed, it has first to be noticed that the sequence of quotients $H_{m+1}^{(k)}/H_m^{(k)}$ converges to the product $\lambda_1 \ldots \lambda_k$, and thus

$$\lim_{m\to\infty} \frac{H_{m+1}^{(k)}}{H_m^{(k)}} \frac{H_{m+1}^{(k-1)}}{H_m^{(k-1)}} = \lambda_k. \tag{4.1}$$

Together with these results, the following Jacobi identity (compound determinant formula)

$$\Big(H_m^{(k)}\Big)^2 = H_{m-1}^{(k)} H_{m+1}^{(k)} + H_{m-1}^{(k+1)} H_{m+1}^{(k-1)}, \tag{4.2}$$

helps for designing, thanks to simple row and column determinant expansions for $H_m^{(k)}$, an efficient, though costly, algorithmic version of the method for approximating any λ_k. This was the idea proposed initially by Aitken [Ait26; Ait31] and that paved the way to further developments.

The algorithmic period The genuine beginning of the algorithmic methodology to approximate eigenvalues comes from the suggestion by Stiefel to Rutishauser to look again at the Schwarz constants a_m. By this way, Rutishauser [Rut54a; Rut54b; Rut54c] introduced the so-called *qd algorithm* (for *quotient-difference*) that avoids the explicit use of Hankel determinants, and thus reduces severely the complexity and the ill-posedness of the method. He reformulates, first for tridiagonal matrices, his algorithm into the LR algorithm, based on gaussian LU elimination steps. The idea is surprisingly simple in its formulation, and consists finally in the following matrix decomposition/recomposition process:

$$L_k R_k := A_k, \quad A_{k+1} := R_k L_k.$$

As a consequence, all matrices $(A_k)_{k\geq 0}$ share the same spectrum. The convergence to the diagonal matrix with values sorted by decreasing modulus was obtained by Rutishauser [Rut55] for real symmetric matrix with positive simple (thus separated) eigenvalues. A simple extension to general symmetric matrices with a quadratic convergence is obtained with Bauer in the same C. R. Acad. Sci. [RB55].

Motivated by handling with more general (i.e. non-symmetric) matrices, and by both the computational cost and the possible numerical instability, Francis [Fra61a; Fra61b] introduced the now famous variant, that makes use of unitary matrices rather than gaussian elimination

(see also Kublanovskaja [Kub61]): *Francis' QR algorithm.* A preprocessing step based on Householder matrices (see [Wil60]) is intended to reduce the matrix A first under the upper Hessenberg form so as to reduce the computational cost of further algorithmic steps. Actually, this first step produces a symmetric tridiagonal matrix when dealing with symmetric matrix A. The overall diagonalization algorithm is again very simple:

$$Q_k R_k := A_k, \quad A_{k+1} = R_k Q_k. \tag{4.3}$$

A variant with shifts (and even the double complex conjugate pairs shift technique) has made Francis' method very famous due to the significant increase in the convergence properties of the induced algorithm (cubic in the best case)

$$Q_k R_k := A_k - \sigma_k I_n, \quad A_{k+1} = R_k Q_k + \sigma_k I_n,$$

Isospectral dynamical systems The QR iteration (4.3) can be understood as a discrete dynamical system that preserves the spectrum along the orbits and that admits the subset of diagonal matrices as a manifold of equilibrium points. Concretely, this last property can be understood after observing then the vanishing of the Lie commutator (or bracket) $[Q, R] = QR - RQ$. In finite dimension, among all the isospectral diagonal equilibrium points, only a few exhibits attractivity properties in a more or less large isospectral stable manifold.

From Watkins [Wat84], it is known that the QR algorithm (and several generalizations for other abstract matrix factorization, e.g. LU and Cholesky, see Chu and Norris [CN88]) can be reinterpreted as the sampling at integer times of the solution to a differential system acting then on a time-dependent matrix, or linear operator, $A(t)$. These continuous dynamical systems also have a bracket structures and therefore manifest isospectrality in the flow as well. A more recent review about the reinterpretation of several linear algebra algorithms through such dynamical systems can be found in Chu [Chu08].

Our aim is to extend the previous ideas to the case of infinite dimension operators. In this spirit, Bach and Bru [BB10] tackle the example of the Brockett flow (presented later on) in infinite dimension. We analyze the structural and convergence properties of some bracket flows in a quite general framework (Hilbert-Schmidt operators) for which the diagonal operators turns out again to be an attractive manifold. There are deep applications in mathematical physics, for instance Bach and Bru [BB16] made use of this for diagonalizing unbounded operators in boson quantum field theory.

4.2 Double bracket flows of Hilbert-Schmidt operators

Some motivations: Toda lattice and Lax pairs The Toda lattice proposed by Toda [Tod67; Tod75] models a system of equal masses connected on a one-dimensional line by identical springs with a nonlinear exponential restoring force. The associated Hamiltonian has the following form

$$\mathsf{H}(p, q) = \frac{1}{2} \sum_n p_n^2 + \sum_n \mathsf{e}^{q_n - q_{n+1}},$$

where p_n and q_n are respectively the momentum and the displacement of the n-th mass from equilibrium. This model may seem to be quite artificial, however it exhibits many important features such as complete integrability, nonlinear normal modes, and it is in fact directly related to some approximation of solutions to the Korteweg-de-Vries equation for long-wave limits. The bracket structure behind this model is observed by Flaschka [Fla74a; Fla74b]. Namely, the dynamical hamiltonian system associated to H, that is

$$\frac{dq_n}{dt} = \frac{\partial \mathsf{H}}{\partial p_n}, \qquad \frac{dp_n}{dt} = -\frac{\partial \mathsf{H}}{\partial q_n},$$

can be, after the change of variables $a_n = -\frac{1}{2}p_n$, $b_n = \frac{1}{2}\mathsf{e}^{(q_n - q_{n+1})/2}$, formulated under the more tractable form

$$\frac{da_n}{dt} = 2(b_n^2 - b_{n-1}^2), \qquad \frac{db_n}{dt} = b_n(a_{n+1} - a_n).$$

Actually, this reformulation reveals the associated Lax [Lax68] pair

$$L(t) = \begin{pmatrix} a_1 & b_1 & & \\ b_1 & a_2 & \ddots & \\ & \ddots & \ddots & b_{N-1} \\ & & b_{N-1} & a_N \end{pmatrix}, \quad P(t) = \begin{pmatrix} 0 & -b_1 & & \\ b_1 & 0 & \ddots & \\ & \ddots & \ddots & -b_{N-1} \\ & & b_{N-1} & 0 \end{pmatrix}$$

and the associated bracket flow:

$$\frac{d}{dt}L = [L, P].$$

Global isospectral flows of Hilbert-Schmidt operators Let us now introduce some notations. We consider a separable complex Hilbert space $\mathcal{H}$, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and a Hilbert basis $(e_n)_{n\geq 0}$. $\mathcal{L}(\mathcal{H})$ denotes the set of bounded endomorphisms on $\mathcal{H}$ and $\mathcal{L}_2(\mathcal{H})$ the subset of Hilbert-Schmidt operators endowed with the norm $\|H\|_{\mathsf{HS}}^2 = \mathsf{tr}(H^\star H) = \sum_{n\geq 0} \|He_n\|^2 = \sum_{n=0}^{+\infty} |\lambda_n(H)|^2$ and has the canonical Hilbert basis $e_{i,j} = \langle e_i, \cdot \rangle e_j$. The subset of bounded operators that are diagonal with respect to the basis $(e_n)_{n\geq 0}$ is denoted $\mathcal{D}(\mathcal{H})$. We also denote $\mathcal{S}(\mathcal{H})$, $\mathcal{A}(\mathcal{H})$ and $\mathcal{U}(\mathcal{H})$ respectively the symmetric (i.e. hermitian), skew-symmetric and unitary operators in $\mathcal{L}(\mathcal{H})$, and $\mathcal{S}_2(\mathcal{H})$, $\mathcal{A}_2(\mathcal{H})$ their intersection with $\mathcal{L}_2(\mathcal{H})$. The Hilbert space $\mathcal{S}_2(\mathcal{H})$ is equipped with the canonical Hilbert basis: $E_{i,j}$ for $i \leq j$, with coefficients $\frac{1}{\sqrt{2}}(e_{i,j} + e_{j,i})$. The Hilbert space $\mathcal{A}_2(\mathcal{H})$ is equipped with the canonical Hilbert basis $E_{i,j}^{\pm}$ for $i < j$, with coefficients $\frac{1}{\sqrt{2}}(e_{i,j} - e_{j,i})$.

In this context the following result is quite natural to obtain (see Lax [Lax68] and Flaschka [Fla74a])

Proposition 24. *Let $\mathsf{G} : \mathcal{S}(\mathcal{H}) \to \mathcal{A}(\mathcal{H})$ be a locally Lipschitz mapping. Any Cauchy problem*

$$H' = [H, \mathsf{G}(H)], \qquad H(0) = H_0 \in \mathcal{S}_2(\mathcal{H}) \tag{4.4}$$

admits a unique global solution $H \in \mathcal{C}^1(\mathbb{R}, \mathcal{S}_2(\mathcal{H}))$. It is smoothly unitarily equivalent to the initial data: there exists $U \in \mathcal{C}^1(\mathbb{R}, \mathcal{U}(\mathcal{H}))$ such that the solution reads $H(t) = U^\star(t) H_0 U(t)$.

Actually, the change of basis U solves the linear differential problem

$$U' = U\mathsf{G}(H), \qquad U(0) = \mathsf{Id}. \tag{4.5}$$

The asymptotic dynamic in the neighbourhood of an diagonal equilibrium point $H_\infty = \mathsf{diag}(\lambda_i,\ i \geq 0)$ is actually directly related to the quantities g_{ij} involved in the identities $\mathsf{G}(E_{i,j}) = g_{ij}E^{\pm}_{i,j}$ for $i < j$, if available. To be more precise, the differential of the flow $\mathsf{F} : H \mapsto [H, \mathsf{G}(H)]$ at the point H_∞, considered as an endomorphism in $\mathcal{S}_2(\mathcal{H})$, is diagonalized in the basis $E_{i,j}$ with $\mathsf{dF}_{H_\infty}(E_{i,j}) = g_{ij}(\lambda_i - \lambda_j)E_{i,j}$. The first issue is then to identify the equilibrium points with non-positive eigenvalues. 0 is an eigenvalue associated in particular to any $E_{i,i}$.

- **Brockett sorting algorithm**

  The simplest archetypal example proposed from Brockett [Bro91] is related to the linear mapping $\mathsf{G}(H) = [H, A]$ where $A = \mathsf{diag}(a_i,\ i \geq 0) \in \mathcal{D}(\mathcal{H}) \cap \mathcal{S}_2(\mathcal{H})$ is a diagonal operator with $a_1 > a_2 > \ldots > 0$. In that case for $i < j$, one has $g_{ij} = a_j - a_i$ and thus the eigenvalues of dF_{H_∞} are $-(a_i - a_j)(\lambda_i - \lambda_j)$ for $i < j$. All of them are negative if and only if the order of sorting is the same for A and H_∞.

- **Toda flows**

  The previously discussed Toda flow also reads more compactly as $\mathsf{G}(H) = H^- - (H^-)^\star$ where $H^- = \sum_{0 \leq i \leq j} h_{i,j}e_i^\star e_j$ is the low truncation of H. Again, from the identity $\mathsf{G}(E_{i,j}) = -E^{\pm}_{i,j}$, we deduce the eigenvalues of dF_{H_∞} that are $-(\lambda_i - \lambda_j)$ for $i < j$.

- **Wegner flow**

  The flow by Wegner [Weg94] is defined by $\mathsf{G}(H) = [H, \mathsf{diag}(H)]$ where $\mathsf{diag}(H) = \sum_{0 \leq i} h_{i,i}e_i^\star e_i$ is the diagonal part of H. In that case, the mapping G does not act linearly but the spectrum of dF_{H_∞} at a given diagonal point $H_\infty = \mathsf{diag}(\lambda_i,\ i \geq 0)$ is then $-(\lambda_i - \lambda_j)^2$ for $i < j$.

Convergence results Contrary to the finite dimensional case, the unitary matrix U does not evolve in a compact set, thus other strategies have to be considered (a priori bounds, integrability properties and monotonicity arguments) to conclude on convergence.

Theorem 25 (Corollary 1.8 and 1.9 in [A/BR17])**.** *Assume $\mathsf{G} \in \mathcal{L}(\mathcal{S}_2(\mathcal{H}), \mathcal{A}_2(\mathcal{H}))$ and appropriate sign and lower bound assumptions on the "eigenvalues" g_{ij} of G. Then, the global solution $H(t)$ weakly converges in $\mathcal{S}_2(\mathcal{H})$ to $H_\infty = \mathsf{diag}(\lambda_i,\ i \geq 0) \in \mathcal{D}_2(\mathcal{H})$, with spectrum included in $\mathsf{spec}(H_0)$ with multiplicities.*
If $\dim \mathcal{H} < +\infty$ then the convergence is available in the strong topology and, if H_0 has simple eigenvalues then the convergence rate is $O(e^{-\gamma t})$ where $\gamma = \inf\{-g_{i,j}(\lambda_i - \lambda_j) > 0,\ i < j\}$.

The corresponding proof permits to extend the usual convergence result for the QR algorithm to the infinite dimensional case. We present this result hereafter with a quantification of the convergence of the off-diagonal terms.

Theorem 26 (Proposition 1.10 in [A/BR17]). *Let $H_0 \in \mathcal{L}_2(\mathcal{H})$ be diagonalizable with eigenvalues $(\lambda_j)_{j\geq 0}$ indexed by decreasing real parts. Assume in addition the existence of an invertible $P \in \mathcal{L}(\mathcal{H})$ such that $PH_0P^{-1} = \mathsf{diag}(\lambda_j,\ j \geq 0)$ and such that the minors $\mathsf{det}\,(\langle Pe_i, e_j\rangle)_{0\leq i,j\leq J}$ are invertible for all integer J. Then the solution $H(t)$ to the Toda flow satisfies for all finite $\ell \in \mathbb{N}$ and with $\delta_\ell := \min_{0\leq j\leq \ell} \mathsf{Re}\,(\lambda_j - \lambda_{j+1})$:*

$$H(t)e_\ell - \sum_{j>\ell} \langle H(t)e_\ell, e_j\rangle e_j = \lambda_\ell e_\ell + O(\mathsf{e}^{-t\delta_\ell}),$$

4.3 Numerical examples in finite dimension

We present hereafter some finite dimensional numerical illustrations for matrices in $\mathcal{M}_5(\mathbb{R})$ with a symmetric (full matrix) initial data, unitarily equivalent to the matrix $\mathsf{diag}\,([1, 4, 9, 16, 25])$. The ODE systems (4.4) plus (4.5) are approximated by means of the adaptive 4th-order Runge-Kutta scheme. Actually, there is no loss of isospectrality during the numerical discretization (up to the machine error) due to the use of the equation (4.5). This would not be true when using (4.4) directly (see the works by Calvo, Iserles, and Zanna [CIZ97] for more details on the numerical integration of bracket flows and Hairer, Lubich, and Wanner [HLW10] for general geometric integration results).

- **Brockett flow**

  The matrix-parameter is the diagonal matrix $A = \mathsf{diag}\,([5, 4, 3, 2, 1])$. On Figure 4.1, the convergence to the limit matrix $H_\infty = \mathsf{diag}\,([25, 16, 9, 4, 1])$ is observed, with a sorting of the eigenvalues in the descending order, in accordance to the ordering in the coefficients in A. The rate of convergence $O(\mathsf{e}^{-3t})$ is asymptotically observed, also in agreement with the theoretical value given by the values $(a_j - a_i)(\lambda_i - \lambda_j)$ in the following empty-low tabular:

$$\begin{bmatrix} . & -9 & -32 & -63 & -96 \\ & . & -7 & -24 & -45 \\ & & . & -5 & -16 \\ & & & . & -3 \\ & & & & . \end{bmatrix}.$$

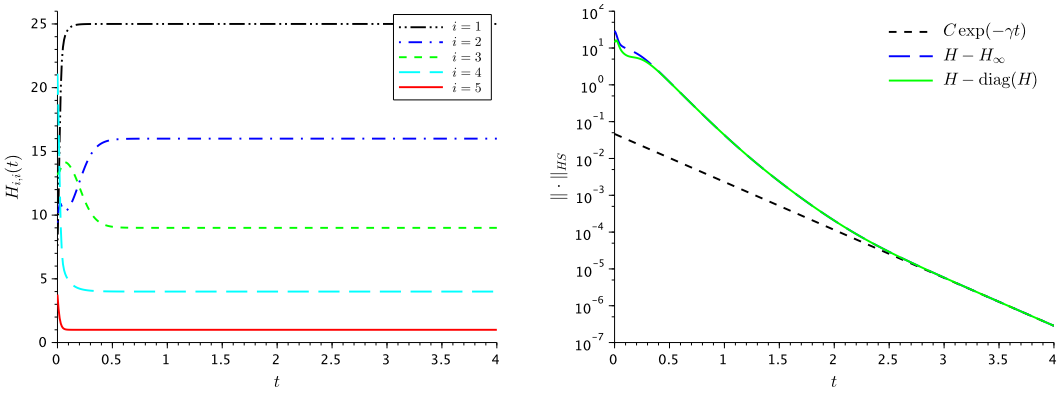

Figure 4.1: Brockett flow – Diagonal entries (left), rate of convergence(right).

- **Toda flow**

  The Toda flow gives similar results with the same rate of convergence. Now the empty-low tabular has coefficients $-(\lambda_i - \lambda_j)$:

$$\begin{bmatrix} . & -9 & -16 & -21 & -24 \\ & . & -7 & -12 & -15 \\ & & . & -5 & -8 \\ & & & . & -3 \\ & & & & . \end{bmatrix}.$$

  The Figure 4.2 illustrates the numerical counterpart of Theorem 26. For any $0 \leq \ell \leq 4$, we compute the norm of the residual column $\| \sum_{j>\ell} \langle H(t)e_\ell, e_j \rangle e_j \|$. The thick dashed curves represent the effective error and the thin solid ones correspond to reference rates, namely $\delta_\ell \in \{9, 7, 5, 3\}$.

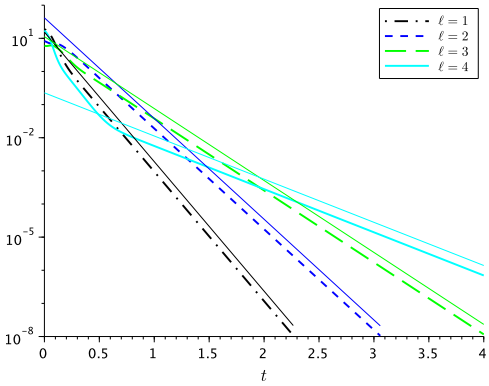

Figure 4.2: Toda flow – Convergence of extradiagonal columns.

- **Wegner flow**

  For the Wegner flow, the effective limiting solution appears to be the matrix $H_\infty = \mathsf{diag}\,([4, 9, 16, 25, 1])$ and the convergence rate $O(\mathsf{e}^{-9t})$ is again in concordance with the theory. Close to the limit H_∞ the empty-low "tabular-eigenvalue" measuring the linear-attractivity of H_∞ reads

$$\begin{bmatrix} . & -25 & -144 & -441 & -9 \\ & . & -49 & -256 & -64 \\ & & . & -81 & -225 \\ & & & . & -576 \\ & & & & . \end{bmatrix}.$$

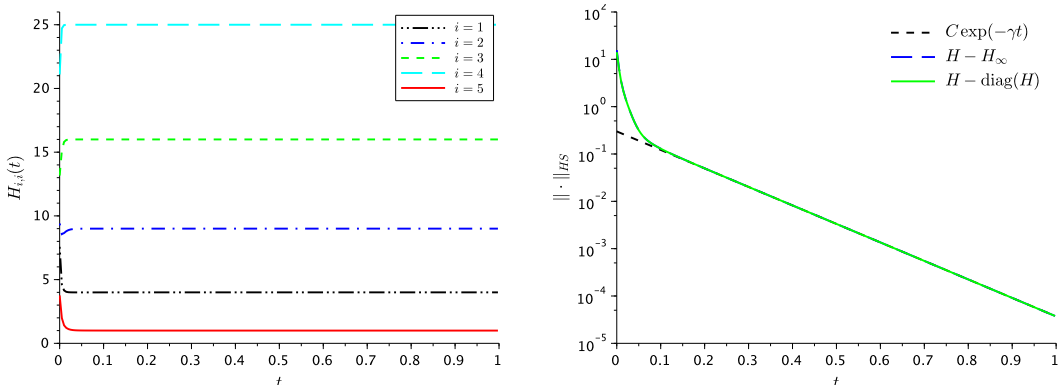

Figure 4.3: Wegner flow – Diagonal entries (left), rate of convergence(right).

CHAPTER 5

Dynamical systems in biology

This chapter is devoted to a brief presentation of the work [A/FJL+20]. This is an interdisciplinary research activity carried out with cellular biologists from the "Institute for Research in Health, Environment and Work" (IRSET – Université de Rennes), and more particularly with Denis Michel. The collaboration followed a first work [A/MBR16] which will not be discussed here.

[A/MBR16] D. Michel, B. Boutin, and P. Ruelle. The accuracy of biochemical interactions is ensured by endothermic stepwise kinetics. *Prog. Biophys. Mol. Biol.* 121 1:35–44, 2016.

[A/FJL+20] G. Flouriot, C. Jehanno, Y. Le Page, P. Le Goff, B. Boutin, and D. Michel. The basal level of gene expression associated with chromatin loosening shapes Waddington landscapes and controls cell differentiation. *J. Mol. Biol.* 432 7:2253–2270, 2020.

Problematic The mechanisms of cellular differentiation and dedifferentiation are under study. More precisely the balance between red and white blood cells (hematopoietic regulation) mostly results from a choice between the proteins called GATA1/2 and PU.1. The concentrations of these proteins, respectively denoted x and y, evolve according to positive and negative effects such as mutual sequestration and the basal gene expression b_1 and b_2 (Figure 5.1). The supported thesis is that the mutual inhibition between the two genes does not proceed through reduction of some basal level, but through preventing the self-stimulations of the proteins.

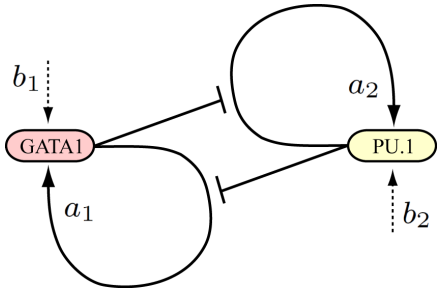

Figure 5.1: Reciprocal inhibition of self-stimulation with no specific regulation of the basal transcription frequencies.

To model these mechanisms, we study the following differential system:

$$\begin{aligned} \frac{dx}{dt} &= b_1 + a_1 \frac{x_f}{K_1 + x_f} - r_1\, x \\ \frac{dy}{dt} &= b_2 + a_2 \frac{y_f}{K_2 + y_f} - r_2\, y \end{aligned} \tag{5.1a}$$

where x_f and y_f are the concentrations of molecules not mutually interacting, because sequestrated in $x \bullet y$ complexes. Given the time scale separation between molecular interactions (very fast) and gene expression dynamics (much slower) the free concentrations are simply given by non-differential, algebraic equations, thus we have the free quantities

$$\begin{aligned} x_f &= x - x \bullet y \\ y_f &= y - x \bullet y \end{aligned} \tag{5.1b}$$

and the sequestrated complex is given by

$$x \bullet y = \frac{1}{2}\left[(D + x + y - \sqrt{(D + x + y)^2 - 4xy}\right]. \tag{5.1c}$$

The constant D is the equilibrium dimerisation constant between x and y. In the further presentation and the numerical simulations, the roles of GATA1/2 and PU.1 are supposed to be symmetric with identical parameters for both genes ($a_1 = a_2$, $b_1 = b_2$, $K_1 = K_2$ and $r_1 = r_2$).

Waddington epigenetic landscapes In parameterized dynamical systems, the analysis of bifurcations in the multistability landscape are important to correctly understand the possible behaviors of the concrete (non-deterministic) system. The dedicated tool for this study is the Waddington landscape, long envisioned as the ideal framework for conceptualizing cell differentiation and development. An epigenetic landscape, in the sense initiated by Waddington [Wad14], is a n-dimensional potential surface shaped by the mutual compatibility or incompatibility of the concentrations of the n cellular components, and modulated by the action of the genes. We reproduce hereafter on Figure 5.2 the nice pictures and heuristic proposed by Waddington. We refer also to the Wikipedia entry "Epigenetics" [Wik23] for more about the subject.

Strategy We propose hereafter a short informal discussion to present the methodology used to compute quasi-potential landscapes associated to the differential system (5.1). For mathematical results, the interested reader can refer to [FW12] or the other references hereafter.

Unlike unidimensional evolution system, not any differential system in higher dimension may be described from a simple scalar-valued potential function. When a Lyapunov function however exist, it may directly provide a scalar characterization of the attractive behavior of steady states and thus enables the possibility to draw a landscape. This is the case for example for any gradient-like systems. More generally, one may try to take into account

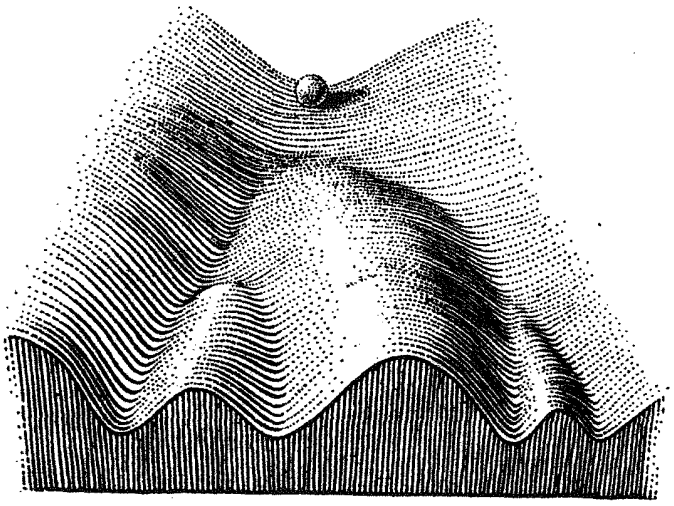

FIGURE 4

Part of an Epigenetic Landscape. The path followed by the ball, as it rolls down towards the spectator, corresponds to the developmental history of a particular part of the egg. There is first an alternative, towards the right or the left. Along the former path, a second alternative is offered; along the path to the left, the main channel continues leftwards, but there is an alternative path which, however, can only be reached over a threshold.

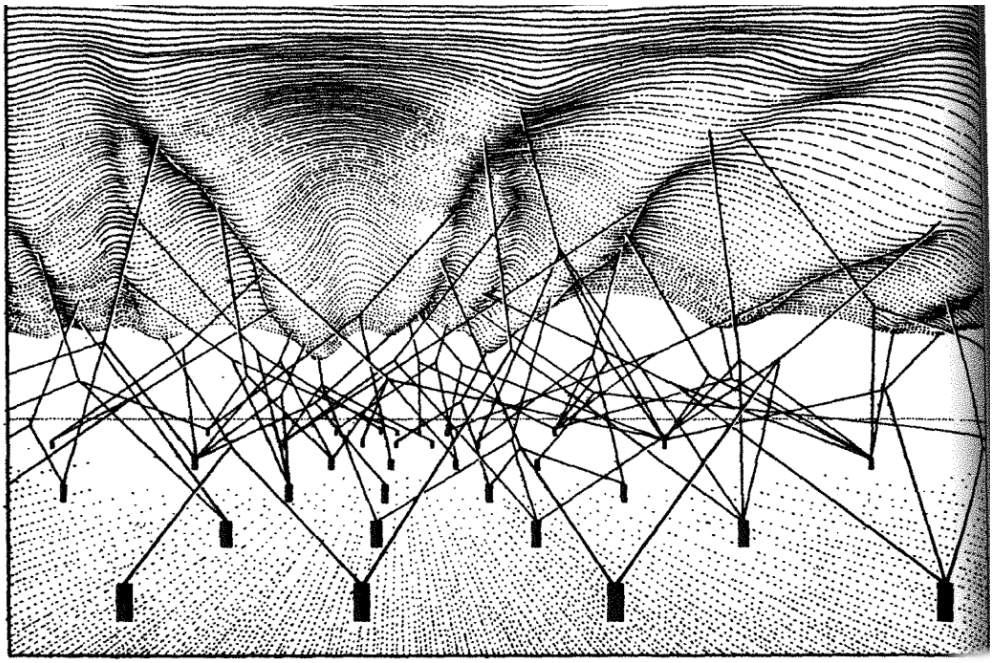

FIGURE 5

The complex system of interactions underlying the epigenetic landscape. The pegs in the ground represent genes; the strings leading from them the chemical tendencies which the genes produce. The modelling of the epigenetic landscape, which slopes down from above one's head towards the distance, is controlled by the pull of these numerous guy-ropes which are ultimately anchored to the genes.

Figure 5.2: Excerpts from the work of Waddington [Wad14].

the hamiltonian part of the dynamics, however this is not clear how to use the Hodge-Helmholtz-like decompositions to draw then a Waddington landscape [WZX+11; Wan15; Hua12; ZAA+12]. A very nice general overview of landscape theories with quasi-potentials is also available from Zhou and Li [ZL16].

The alternative selected approach is based on the probabilistic point of view and concerns the Freidlin and Wentzell [FW12] large deviations principle for invariant measures of stochastic convection-diffusion processes. In the phase space (x, y), we consider many trajectories of the deterministic system (5.1) perturbated with a small brownian motion. These trajectories, as random variables, evolve according to a stochastic differential equation (SDE) and their probability density $p(t, x, y)$ follows the corresponding Fokker-Planck equation, also known as Kolmogorov forward equation:

$$\frac{\partial p}{\partial t} + \frac{\partial (Xp)}{\partial x} + \frac{\partial (Yp)}{\partial y} = \epsilon \Delta_{x,y} p. \tag{5.2}$$

Here the dynamical drift field $X(x, y), Y(x, y)$ corresponds to the respective right-hand sides in (5.1) and the diffusive term in the right-hand-side takes into account isotropically the random brownian processes, thus rendering the mean effect of possible perturbations. In large time, depending on the structure of the vector field $(x, y) \mapsto (X, Y)$, most of the random trajectories of the SDE accumulates close to attractive steady states or singular trajectories of the dynamical system (5.1). They then evolves finally only through a fine balance between the brownian process and the deterministic dynamic. At the level of the PDE, the probability density p then becomes independent of time and converges to the so-called invariant measure of the stochastic process.

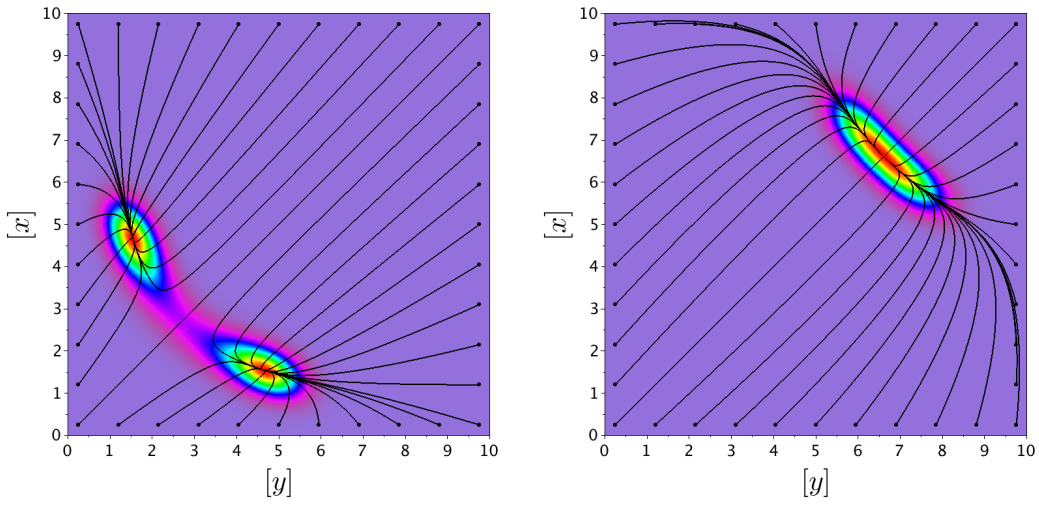

Figure 5.3: Invariant measure for the GATA1/2 (x): PU.1 (y) circuit. Bistable configuration with low basal expression $b = 1$ (left). Monostable configuration with high basal expression $b = 4$ (right).

The convection-diffusion PDE (5.2) is numerically processed, so as to identify the invariant measure for a given set of parameters (see Figure 5.3). The steady state probability p can then be used to figure out the landscape of multistable equilibria. It concentrates to high values

(red) in the neighborhood of such points but vanishes close to repulsing points (violet). Some deterministic trajectories of (5.1) are also represented on the same figure for a set of initial data (black curves). The situation depicted is the following. For a high basal expression $b = 4$ there is monostability ("indecise" cell with equivalent coexpression of the proteins GATA1/2 and PU.1), but for a low basal expression $b = 1$ two equilibria coexist and the probability of jump transition is also measured (pink bridge).

The transition from bistability to monostability is represented on Figure 5.4 with a convenient "projection" of the 4D mapping: $(x, y, b) \mapsto p$ onto a 3D mapping: $(x, b) \mapsto \tilde{p}$.

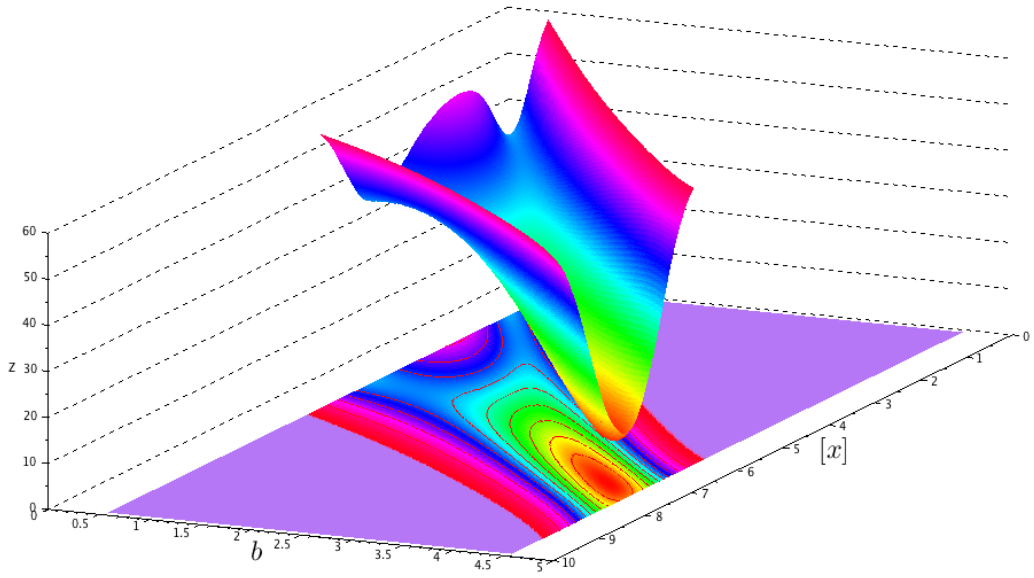

Figure 5.4: Projection of steady states on the axis of one of the variables as a function of b, showing the switch from bi- to monostability when increasing b, for other parameters fixed.

BIBLIOGRAPHY

[Ait26] A. C. Aitken. On Bernoulli's numerical solution of algebraic equations. *Proc. R. Soc. Edinburgh*, 46:289–305, 1926.

[Ait31] A. C. Aitken. Further numerical studies in algebraic equations and matrices. *Proc. R. Soc. Edinburgh*, 51:80–90, 1931.

[ABH+22] N. G. Al Hassanieh, J. W. Banks, W. D. Henshaw, and D. W. Schwendeman. Local compatibility boundary conditions for high-order accurate finite-difference approximations of PDEs. *SIAM J. Sci. Comput.* 44:A3645–A3672, 2022.

[And15] B. Andreianov. New approaches to describing admissibility of solutions of scalar conservation laws with discontinuous flux. In, *CANUM 2014 - 42e Congrès National d'Analyse Numérique*. Volume 50, ESAIM Proc. Surveys, pages 40–65. EDP Sci., Les Ulis, 2015.

[AKR11] B. Andreianov, K. H. Karlsen, and N. H. Risebro. A theory of L1-dissipative solvers for scalar conservation laws with discontinuous flux. *Arch. Ration. Mech. Anal.* 201:27–86, 2011.

[AES03] A. Arnold, M. Ehrhardt, and I. Sofronov. Discrete transparent boundary conditions for the Schrödinger equation: fast calculation, approximation, and stability. *Commun Math Sci*, 1:501–556, 2003.

[Aud11] C. Audiard. On mixed initial-boundary value problems for systems that are not strictly hyperbolic. *Appl. Math. Lett.* 24:757–761, 2011.

[BB10] V. Bach and J.-B. Bru. Rigorous foundations of the Brockett-Wegner flow for operators. *J. Evol. Equ.* 10:425–442, 2010.

[BB16] V. Bach and J.-B. Bru. *Diagonalizing quadratic bosonic operators by non-autonomous flow equation*, volume 1138 of *Mem. Am. Math. Soc.* American Mathematical Society (AMS), Providence, RI, 2016.

[BV06] F. Bachmann and J. Vovelle. Existence and uniqueness of entropy solution of scalar conservation laws with a flux function involving discontinuous coefficients. *Commun. Partial. Differ.* 31:371–395, 2006.

[BLN79] C. Bardos, A. Y. Leroux, and J.-C. Nédélec. First order quasilinear equations with boundary conditions. *Commun. Partial. Differ.* 4:1017–1034, 1979.

[BW93] R. M. Beam and R. F. Warming. The asymptotic spectra of banded Toeplitz and quasi-Toeplitz matrices. *SIAM J. Sci. Comput.* 14:971–1006, 1993.

[BL02] N. Bedjaoui and P. G. LeFloch. Diffusive-dispersive traveling waves and kinetic relations. I. Nonconvex hyperbolic conservation laws. *J. Differ. Equ.* 178:574–607, 2002.

[Ben16] A. Benoit. Geometric optics expansions for hyperbolic corner problems, I: self-interaction phenomenon. *Anal. PDE*, 9:1359–1418, 2016.

[Ben17] A. Benoit. Geometric optics expansions for hyperbolic corner problems II: From weak stability to violent instability. *SIAM J. Math. Anal.* 49:3335–3395, 2017.

[Ben20] A. Benoit. Lower exponential strong well-posedness of hyperbolic boundary value problems in a strip. *Indiana Univ. Math. J.* 69:2267–2323, 2020.

[Ben21] A. Benoit. Stability of finite difference schemes approximation for hyperbolic boundary value problems in an interval. *Math. Comp.* 2021.

[BS07] S. Benzoni-Gavage and D. Serre. *Multidimensional hyperbolic partial differential equations: first-order systems and applications*. Oxford mathematical monographs. Clarendon Press, Oxford ; New York, 2007.

[Ber32] D. Bernoulli. Observationes de seriebus quae formantur ex additione vel substractione quacunque terminorum se mutuo consequentium, ubi praesertim earundem insignis usus pro inveniendis radicum omnium aequationum algebraicarum ostenditur. *Commentarii Acad. Petropol*, 3:85–100, 1732.

[BEL16] C. Besse, M. Ehrhardt, and I. Lacroix-Violet. Discrete artificial boundary conditions for the linearized Korteweg–de Vries equation. *Numer. Methods Partial Differential Equations*, 32:1455–1484, 2016.

[BCN20] C. Besse, J.-F. Coulombel, and P. Noble. Discrete transparent boundary conditions for the two-dimensional leap-frog scheme. *ESAIM: Math. Model. Numer.* 2020.

[BGK+22] C. Besse, S. Gavrilyuk, M. Kazakova, and P. Noble. Perfectly matched layers methods for mixed hyperbolic–dispersive equations. *Water Waves*, 4:313–343, 2022.

[BMZ20] U. Biccari, A. Marica, and E. Zuazua. Propagation of one- and two-dimensional discrete waves under finite difference approximation. *Found. Comput. Math.* 20:1401–1438, 2020.

[BDS02] N. Borovykh, D. Drissi, and M. Spijker. A bound on powers of linear operators, with relevance to numerical stability. *Applied Mathematics Letters*, 15:47–53, 2002.

[BS00] N. Borovykh and M. Spijker. Resolvent conditions and bounds on the powers of matrices, with relevance to numerical stability of initial value problems. *J. Comput. Appl. Math.* 125:41–56, 2000.

[BS02] N. Borovykh and M. Spijker. Bounding partial sums of fourier series in weighted L2-norms, with applications to matrix analysis. *J. Comput. Appl. Math.* 147:349–368, 2002.

[Bri60] L. Brillouin. *Wave propagation and group velocity*. Pure and Applied Physics, Vol. 8. Academic Press, New York-London, 1960.

[Bro91] R. W. Brockett. Dynamical systems that sort lists, diagonalize matrices, and solve linear programming problems. *Linear Algebra Its Appl.* 146:79–91, 1991.

[CIZ97] M. P. Calvo, A. Iserles, and A. Zanna. Numerical solution of isospectral flows. *Math. Comp.* 66:1461–1486, 1997.

[CG01] C. Chainais-Hillairet and E. Grenier. Numerical boundary layers for hyperbolic systems in 1-d. *M2AN Math. Model. Numer. Anal.* 35:91–106, 2001.

[CFvN50] J. G. Charney, R. Fjörtoft, and J. von Neumann. Numerical integration of the barotropic vorticity equation. *Tellus*, 2:237–254, 1950.

[CLL94] G.-Q. Chen, C. D. Levermore, and T.-P. Liu. Hyperbolic conservation laws with stiff relaxation terms and entropy. *Comm. Pure Appl. Math.* 47:787–830, 1994.

[Cho11] O. Chodosh. Infinite matrix representations of isotropic pseudodifferential operators. *Methods Appl. Anal.* 18:351–372, 2011.

[Chu08] M. T. Chu. Linear algebra algorithms as dynamical systems. *Acta Numer.* 17:1–86, 2008.

[CN88] M. T. Chu and L. K. Norris. Isospectral flows and abstract matrix factorizations. *SIAM J. Numer. Anal.* 25:1383–1391, 1988.

[CJL+14] F. Coquel, S. Jin, J.-G. Liu, and L. Wang. Well-posedness and singular limit of a semilinear hyperbolic relaxation system with a two-scale discontinuous relaxation rate. *Archive for Rational Mechanics and Analysis*, 214:1051–1084, 2014.

[CG11] J.-F. Coulombel and A. Gloria. Semigroup stability of finite difference schemes for multidimensional hyperbolic initial boundary value problems. *Math. Comp.* 80:165–203, 2011.

[Cou13] J.-F. Coulombel. Stability of finite difference schemes for hyperbolic initial boundary value problems. In, *HCDTE lecture notes. Part I. Nonlinear hyperbolic PDEs, dispersive and transport equations*. Volume 6, AIMS Ser. Appl. Math. Page 146. Am. Inst. Math. Sci. (AIMS), Springfield, MO, 2013.

[Cou15] J.-F. Coulombel. The Leray-Gårding method for finite difference schemes. *Journal de l'École polytechnique - Mathématiques*, 2:297–331, 2015.

[Cou19] J.-F. Coulombel. Transparent numerical boundary conditions for evolution equations: Derivation and stability analysis. *Annales de la Faculté des Sciences de Toulouse. Mathématiques.* 28:259–327, 2019.

[Cou20] J.-F. Coulombel. The Leray-Gårding method for finite difference schemes. II. Smooth crossing modes. 2020. URL: http://arxiv.org/abs/2009.11657. preprint.

[CL20] J.-F. Coulombel and F. Lagoutière. The Neumann numerical boundary condition for transport equations. *Kinetic and Related Models*, 13:1–32, 2020.

[CFL28] R. Courant, K. Friedrichs, and H. Lewy. Über die partiellen Differenzengleichungen der mathematischen Physik. *Math. Ann.* 100:32–74, 1928.

[CN47] J. Crank and P. Nicolson. A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type. *Proceedings of the Cambridge Philosophical Society*, 43:50–67, 1947.

[dAle68] J. L. R. d'Alembert. *Opuscules mathématiques ou Mémoires sur différens sujets de géométrie, de méchanique, etc*, volume 5. 1768.

[Daf73] C. M. Dafermos. Solution of the Riemann problem for a class of hyperbolic systems of conservation laws by the viscosity method. *Archive for Rational Mechanics and Analysis*, 52:1–9, 1973.

[DDJ18] G. Dakin, B. Després, and S. Jaouen. Inverse Lax–Wendroff boundary treatment for compressible lagrange-remap hydrodynamics on cartesian grids. *J. Comput. Phys.* 353:228–257, 2018.

[DLM95] G. Dal Maso, P. G. LeFloch, and F. Murat. Definition and weak stability of nonconservative products. *Journal de Mathematiques Pures et Appliquées. Neuvieme Serie*, 74:483–548, 1995.

[dBoo05] C. de Boor. Divided differences. *Surv. Approx. Theory*, 1:46–69, 2005.

[Des09] B. Després. Uniform asymptotic stability of Strang's explicit compact schemes for linear advection. *SIAM J. Numer. Anal.* 47:3956–3976, 2009.

[DiP85] R. J. DiPerna. Measure-valued solutions to conservation laws. *Archive for Rational Mechanics and Analysis*, 88:223–270, 1985.

[DL88] F. Dubois and P. LeFloch. Boundary conditions for nonlinear hyperbolic systems of conservation laws. *J. Differ. Equ.* 71:93–122, 1988.

[EA01] M. Ehrhardt and A. Arnold. Discrete transparent boundary conditions for the Schrödinger equation. In, *Riv. Mat. Univ. Parma (6)*. Volume 4, pages 57–108. 2001.

[Ehr10] M. Ehrhardt. Absorbing boundary conditions for hyperbolic systems. *Numer. Math. Theory Methods Appl.* 3:295–337, 2010.

[EN88] T. Eirola and O. Nevanlinna. What do multistep methods approximate? *Numer. Math.* 53:559–569, 1988.

[FG21] E. Faou and B. Grébert. Discrete pseudo-differential operators and applications to numerical schemes. 2021. URL: https://arxiv.org/abs/2109.15186. preprint.

[Fla74a] H. Flaschka. The Toda lattice. I. Existence of integrals. *Phys. Rev. B (3)*, 9:1924–1925, 1974.

[Fla74b] H. Flaschka. On the Toda lattice. II. Inverse-scattering solution. *Progr. Theoret. Phys.* 51:703–716, 1974.

[Fra61a] J. G. F. Francis. The QR transformation: a unitary analogue to the LR transformation. I. *Comput. J.* 4:265–271, 1961.

[Fra61b] J. G. F. Francis. The QR transformation. II. *Comput. J.* 4:332–345, 1961.

[FW12] M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*, volume 260 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, Heidelberg, 2012.

[FL67] K. O. Friedrichs and P. D. Lax. On symmetrizable differential operators. In *Singular Integrals (Proc. Sympos. Pure Math., Chicago, Ill., 1966)*, pages 128–137. Amer. Math. Soc., Providence, R.I., 1967.

[GD12] J.-L. García Zapata and J. C. Díaz Martín. A geometric algorithm for winding number computation with complexity analysis. *Journal of Complexity*, 28:320–345, 2012.

[GT85] M. Giles and W. Thompkins. Propagation and stability of wavelike solutions of finite difference equations with variable coefficients. *J. Comput. Phys.* 58:349–360, 1985.

[GT83] M. Giles and W. T. Thompkins. Asymptotic analysis of numerical wave propagation in finite difference equations, Cambridge, Mass.: Gas Turbine & Plasma Dynamics Laboratory, Massachusetts, 1983.

[God03] P. Godillon. Green's function pointwise estimates for the modified Lax-Friedrichs scheme. *M2AN Math. Model. Numer. Anal.* 37:1–39, 2003.

[GR63] S. K. Godunov and V. S. Rjabenkii. Spectral criteria for the stability of boundary-value problems for non-selfadjoint difference equations. *Uspehi Mat. Nauk*, 18:3–14, 3 (111), 1963.

[Gol77] M. Goldberg. On a boundary extrapolation theorem by Kreiss. *Math. Comp.* 31:469–477, 1977.

[GT81] M. Goldberg and E. Tadmor. Scheme-independent stability criteria for difference approximations of hyperbolic initial-boundary value problems. II. *Math. Comp.* 36:603–626, 1981.

[GU09] G. H. Golub and F. Uhlig. The QR algorithm: 50 years later its genesis by John Francis and Vera Kublanovskaya and subsequent developments. *IMA J. Numer. Anal.* 29:467–485, 2009.

[GvdV00] G. H. Golub and H. A. van der Vorst. Eigenvalue computation in the 20th century. *J. Comput. Appl. Math.* 123:35–65, 2000.

[GL96] J. M. Greenberg and A. Y. Leroux. A well-balanced scheme for the numerical processing of source terms in hyperbolic equations. *SIAM J. Numer. Anal.* 33:1–16, 1996.

[GKS72] B. Gustafsson, H.-O. Kreiss, and A. Sundström. Stability theory of difference approximations for mixed initial boundary value problems. II. *Math. Comp.* 26:649–686, 1972.

[GKO13] B. Gustafsson, H.-O. Kreiss, and J. Oliger. *Time-dependent problems and difference methods*. Pure and Applied Mathematics (Hoboken). John Wiley & Sons, Inc., Hoboken, NJ, 2nd edition, 2013.

[GP11] M. H. Gutknecht and B. N. Parlett. From qd to LR, or, how were the qd and LR algorithms discovered ? *IMA J. Numer. Anal.* 31:741–754, 2011.

[Had92] J. Hadamard. Essay on the study of functions defined by their taylor expansion. *Journ. de Math. (4)*, 8:101–186, 1892.

[HLW10] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2010.

[HW96] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II*, volume 14 of *Springer Series in Computational Mathematics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1996.

[Her63] R. Hersh. Mixed problems in several variables. *J. Math. Mech.* 12:317–334, 1963.

[Hig86a] R. L. Higdon. Initial-boundary value problems for linear hyperbolic systems. *SIAM Rev.* 28:177–217, 1986.

[Hig86b] R. L. Higdon. Absorbing boundary conditions for difference approximations to the multidimensional wave equation. *Math. Comp.* 47:437–459, 1986.

[Hua12] S. Huang. The molecular and mathematical basis of Waddington's epigenetic landscape: a framework for post-Darwinian biology? *Bioessays*, 34:149–157, 2012.

[ILR20] M. Inglard, F. Lagoutière, and H. H. Rugh. Ghost solutions with centered schemes for one-dimensional transport equations with Neumann boundary conditions. *Ann. Fac. Sci. Toulouse Math. (6)*, 29:927–950, 2020.

[Jin99] S. Jin. Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations. *SIAM Journal on Scientific Computing*, 21:441–454, 1999.

[JX95] S. Jin and Z. Xin. The relaxation schemes for systems of conservation laws in arbitrary space dimensions. *Comm. Pure Appl. Math.* 48:235–276, 1995.

[KN20] M. Kazakova and P. Noble. Discrete transparent boundary conditions for the linearized Green-Naghdi system of equations. *SIAM J. Numer. Anal.* 58:657–683, 2020.

[Kre68] H.-O. Kreiss. Stability theory for difference approximations of mixed initial boundary value problems. I. *Math. Comp.* 22:703–714, 1968.

[Kre70] H.-O. Kreiss. Initial boundary value problems for hyperbolic systems. *Comm. Pure Appl. Math.* 23:277–298, 1970.

[Kub61] V. N. Kublanovskaja. Certain algorithms for the solution of the complete problem of eigenvalues. *Soviet Math. Dokl.* 2:17–19, 1961.

[Lax68] P. D. Lax. Integrals of nonlinear equations of evolution and solitary waves. *Comm. Pure Appl. Math.* 21:467–490, 1968.

[Le 23] P. Le Barbenchon. *Étude théorique et numérique de la stabilité GKS pour des schémas d'ordre élevé en présence de bords*, Université de Rennes, 2023.

[LN23] P. Le Barbenchon and P. Navaro. Boundaryscheme: python package for numerical scheme with boundary. 2023. URL: https://doi.org/10.5281/zenodo.7773741.

[Le 12] D. Y. Le Roux. Spurious inertial oscillations in shallow-water models. *J. Comput. Phys.* 231:7959–7987, 2012.

[LT99] P. G. LeFloch and A. E. Tzavaras. Representation of weak limits and definition of nonconservative products. *SIAM Journal on Mathematical Analysis*, 30:1309–1342 (electronic), 1999.

[LSZ17] T. Li, C.-W. Shu, and M. Zhang. Stability analysis of the inverse Lax–Wendroff boundary treatment for high order central difference schemes for diffusion equations. *J. Sci. Comput.* 70:576–607, 2017.

[LY01] H. Liu and W.-A. Yong. Time-asymptotic stability of boundary-layers for a hyperbolic relaxation system. *Communications in Partial Differential Equations*, 26:1323–1343, 2001.

[Liu87] T.-P. Liu. Hyperbolic conservation laws with relaxation. *Comm. Math. Phys.* 108:153–175, 1987.

[LM06] F. Lorcher and C.-D. Munz. Lax-Wendroff-type schemes of arbitrary order in several space dimensions. *IMA J. Numer. Anal.* 27:593–615, 2006.

[LN91] C. Lubich and O. Nevanlinna. On resolvent conditions and stability estimates. *BIT*, 31:293–313, 1991.

[MZ15] A. Marica and E. Zuazua. Propagation of 1D waves in regular discrete heterogeneous media: a Wigner measure approach. *Found. Comput. Math.* 15:1571–1636, 2015.

[Mét04] G. Métivier. *Small Viscosity and Boundary Layer Methods*. Modeling and Simulation in Science, Engineering and Technology. Birkhäuser Boston, Boston, MA, 2004.

[Mét17] G. Métivier. On the L2 well posedness of hyperbolic initial boundary value problems. *Annales de l'institut Fourier*, 67:1809–1863, 2017.

[Mic87] D. Michelson. Convergence theorem for difference approximations of hyperbolic quasilinear initial-boundary value problems. *Math. Comp.* 49:445–459, 1987.

[Nev01] O. Nevanlinna. Resolvent conditions and powers of operators. *Studia Mathematica*, 145:113–134, 2001.

[Ngu20] T. H. T. Nguyen. *Numerical approximation of boundary conditions and stiff source terms in hyperbolic equations*, Université de Rennes 1, 2020.

[Osh74a] S. Osher. An ill-posed problem for a strictly hyperbolic equation in two unknowns near a corner. *Bull. Amer. Math. Soc.* 80:705–708, 1974.

[Osh74b] S. Osher. Initial-boundary value problems for hyperbolic systems in regions with corners. II. *Trans. Amer. Math. Soc.* 198:155–175, 1974.

[Pop40] T. Popoviciu. Introduction à la théorie des différences divisées. *Bulletin mathématique de la Société Roumaine des Sciences*, 42:65–78, 1940.

[Rau72] J. Rauch. L2 is a continuable initial condition for Kreiss' mixed problems. *Comm. Pure Appl. Math.* 25:265–285, 1972.

[RT92] L. Reichel and L. N. Trefethen. Eigenvalues and pseudo-eigenvalues of Toeplitz matrices. *Linear Algebra and its Applications*, 162–164:153–185, 1992.

[Rut54a] H. Rutishauser. Anwendungen des Quotienten-Differenzen-Algorithmus. *Z. Angew. Math. Phys.* 5:496–508, 1954.

[Rut54b] H. Rutishauser. Der Quotienten-Differenzen-Algorithmus. *Z. Angew. Math. Phys.* 5:233–251, 1954.

[Rut54c] H. Rutishauser. Ein infinitesimales Analogon zum Quotienten-Differenzen-Algorithmus. *Arch. Math. (Basel)*, 5:132–137, 1954.

[Rut55] H. Rutishauser. Une méthode pour la détermination des valeurs propres d'une matrice. *C. R. Acad. Sci. Paris*, 240:34–36, 1955.

[RB55] H. Rutishauser and F. L. Bauer. Détermination des vecteurs propres d'une matrice par une méthode itérative avec convergence quadratique. *C. R. Acad. Sci. Paris*, 240:1680–1681, 1955.

[SS60] P. Schmidt and F. Spitzer. The Toeplitz matrices of an arbitrary Laurent polynomial. *Math. Scand.* 8:15–38, 1960.

[Sch14] S. Schochet. Singular limits of symmetric hyperbolic systems with large variable-coefficient terms. *Comm. Partial Differential Equations*, 39:842–875, 2014.

[Sen12] T. K. Sengupta. Spurious waves in discrete computation of wave phenomena and flow problems. *Applied Mathematics and Computation*:31, 2012.

[Spi17] M. N. Spijker. Stability and boundedness in the numerical solution of initial value problems. *Math. Comp.* 86:2777–2798, 2017.

[Str94] B. Strand. Summation by parts for finite difference approximations for d/dx. *Journal of Computational Physics*, 110:47–67, 1994.

[Str67] G. Strang. On strong hyperbolicity. *J. Math. Kyoto Univ.* 6:397–417, 1967.

[Str80] G. Strang. *Linear algebra and its applications.* Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, second edition, 1980.

[Str04] J. C. Strikwerda. *Finite difference schemes and partial differential equations. 2nd ed.* Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2nd edition, 2004.

[Sze14] O. Szehr. Eigenvalue estimates for the resolvent of a non-normal matrix. *J. Spectr. Theory*, 4:783–813, 2014.

[TS10] S. Tan and C.-W. Shu. Inverse Lax-Wendroff procedure for numerical boundary conditions of conservation laws. *J. Comput. Phys.* 229:8144–8166, 2010.

[Thu86] M. Thuné. Automatic GKS stability analysis. *SIAM J. Sci. Statist. Comput.* 7:959–977, 1986.

[Tod67] M. Toda. Vibration of a chain with nonlinear interaction. *J. Phys. Soc. Jpn.* 22:431–436, 1967.

[Tod75] M. Toda. Studies of a non-linear lattice. *Phys. Rep.* 18C:1–123, 1975.

[Tre85] L. N. Trefethen. Stability of finite-difference models containing two boundaries or interfaces:25, 1985.

[Tre05] L. N. Trefethen. Wave packet pseudomodes of variable coefficient differential operators. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 461:3099–3122, 2005.

[Tre82] L. N. Trefethen. Group velocity in finite difference schemes. *SIAM Review*, 24:113–136, 1982.

[Tre84] L. N. Trefethen. Instability of difference models for hyperbolic initial-boundary value problems. *Comm. Pure Appl. Math.* 37:329–367, 1984.

[TC04] L. N. Trefethen and S. J. Chapman. Wave packet pseudomodes of twisted Toeplitz matrices. *Comm. Pure Appl. Math.* 57:1233–1264, 2004.

[TE05] L. N. Trefethen and M. Embree. *Spectra and pseudospectra: the behavior of nonnormal matrices and operators*. Princeton University Press, Princeton, N.J, 2005, pages xviii + 606.

[Vic81a] R. Vichnevetsky. Energy and group velocity in semi discretizations of hyperbolic equations. *Mathematics and Computers in Simulation*, 23:333–343, 1981.

[Vic81b] R. Vichnevetsky. Propagation through numerical mesh refinement for hyperbolic equations. *Mathematics and Computers in Simulation*, 23:344–353, 1981.

[VS15] F. Vilar and C.-W. Shu. Development and stability analysis of the inverse Lax-Wendroff boundary treatment for central compact schemes. *ESAIM: M2AN*, 49:39–67, 2015.

[Vol67] A. I. Volpert. Spaces BV and quasilinear equations. *Mat. Sb. (N.S.)* 73 (115):255–302, 1967.

[Wad14] C. H. Waddington. *The Strategy of the Genes*. Routledge, London, 2014.

[Wan15] J. Wang. Landscape and flux theory of non-equilibrium dynamical systems with application to biology. *Advances in Physics*, 64:1–137, 2015.

[WZX+11] J. Wang, K. Zhang, L. Xu, and E. Wang. Quantifying the Waddington landscape and biological paths for development and differentiation. *Proceedings of the National Academy of Sciences*, 108:8257–8262, 2011.

[Wat84] D. S. Watkins. Isospectral flows. *SIAM Rev.* 26:379–391, 1984.

[Weg94] F. Wegner. Flow-equations for Hamiltonians. *Ann. Phys. (8)*, 3:77–91, 1994.

[Whi74] G. B. Whitham. *Linear and nonlinear waves.* John Wiley & Sons, Hoboken, NJ, 1974.

[Wik23] Wikipedia. Epigenetics. 2023. URL: https://en.wikipedia.org/w/index.php?title=Epigenetics&oldid=1141129834.

[Wil60] J. H. Wilkinson. Householder's method for the solution of the algebraic eigenproblem. *Comput. J.* 3:23–27, 1960.

[Wu95] L. Wu. The semigroup stability of the difference approximations for initial-boundary value problems. *Math. Comp.* 64:71–88, 1995.

[XX00] Z. Xin and W.-Q. Xu. Stiff well-posedness and asymptotic convergence for a class of linear relaxation systems in a quarter plane. *J. Differ. Equ.* 167:388–437, 2000.

[XX02] Z. Xin and W.-Q. Xu. Initial-boundary value problem to systems of conservation laws with relaxation. *Quart. Appl. Math.* 60:251–281, 2002.

[XX04] Z. Xin and W.-Q. Xu. Boundary conditions and boundary layers for a class of linear relaxation systems in a quarter plane. In N. B. Abdallah, I. M. Gamba, C. Ringhofer, A. Arnold, R. T. Glassey, P. Degond, and C. D. Levermore, editors. Redacted by D. N. Arnold and F. Santosa, *Transport in Transition Regimes*. Volume 135, pages 279–292. Springer New York, New York, NY, 2004.

[Xu04] W.-Q. Xu. Boundary conditions and boundary layers for a multi-dimensional relaxation model. *J. Differ. Equ.* 197:85–117, 2004.

[Ye04] M. Ye. Numerical boundary layers of conservation laws with relaxation extension. *Appl. Numer. Math.* 51:385–405, 2004.

[Yon99a] W.-A. Yong. Boundary conditions for hyperbolic systems with stiff source terms. *Indiana University Mathematics Journal*, 48, 1999.

[Yon99b] W.-A. Yong. Singular perturbations of first-order hyperbolic systems with stiff source terms. *J. Differ. Equ.* 155:89–132, 1999.

[Yon01] W.-A. Yong. Remarks on hyperbolic relaxation systems. In, *Hyperbolic problems: theory, numerics, applications, Vol. I, II (Magdeburg, 2000)*. Volume 141, Internat. Ser. Numer. Math., 140, pages 921–929. Birkhäuser, Basel, 2001.

[Yon02] W.-A. Yong. Basic structures of hyperbolic relaxation systems. *Proc. Roy. Soc. Edinburgh Sect. A*, 132:1259–1274, 2002.

[YJ05] W.-A. Yong and W. Jäger. On hyperbolic relaxation problems. In, *Analysis and numerics for conservation laws*, pages 495–520. Springer, Berlin, 2005.

[ZM13] J. L. G. Zapata and J. C. D. Martín. A geometrical root finding method for polynomials, with complexity analysis. 2013. URL: http://arxiv.org/abs/1308.4217. preprint.

[ZW04] S.-Y. Zhang and Y.-G. Wang. Well-posedness and asymptotics for initial boundary value problems of linear relaxation systems in one space variable. *Z. Anal. Anwend.*607–630, 2004.

[ZAA+12] J. X. Zhou, M. D. S. Aliyu, E. Aurell, and S. Huang. Quasi-potential landscape in complex multi-stable systems. *Journal of The Royal Society Interface*, 9:3539–3553, 2012.

[ZL16] P. Zhou and T. Li. Construction of the landscape for multi-stable systems: Potential landscape, quasi-potential, A-type integral and beyond. *J Chem Phys*, 144:094109, 2016.

[ZY20] Y. Zhou and W.-A. Yong. Boundary conditions for hyperbolic relaxation systems with characteristic boundaries of type II. 2020.

[ZY21] Y. Zhou and W.-A. Yong. Boundary conditions for hyperbolic relaxation systems with characteristic boundaries of type I. *J. Differential Equations*, 281:289–332, 2021.

COLLEGE DOCTORAL BRETAGNE
MATHS, TELECOMS
INFORMATIQUE, SIGNAL
SYSTEMES, ELECTRONIQUE

Université de Rennes

Titre : Méthodes de différences finies pour des problèmes hyperboliques avec bords : stabilité et analyse multi-échelle

Résumé : Ce manuscrit d'habilitation à diriger des recherches présente les travaux que j'ai effectués ces dernières années. Ils se concentrent sur l'étude de la stabilité et sur l'analyse multi-échelle de méthodes numériques de différences finies, mises en œuvre dans le cadre de l'approximation de problèmes hyperboliques linéaires avec bords. Dans un tel contexte, différentes échelles peuvent intervenir, liées par exemple aux phénomènes de viscosité, de relaxation ou de discrétisation. À ces échelles, les interactions entre le problème intérieur et le bord du domaine de calcul sont alors susceptibles d'engendrer des effets parasites inattendus, tels que des couches limites. Elles nuisent souvent sévèrement aux propriétés de stabilité dans l'asymptotique souhaitée et réduisent parfois la qualité et la précision de l'approximation. Il apparaît alors crucial de discriminer et d'écarter les situations pathologiques.

Les trois premiers chapitres portent successivement sur 1) la théorie générale de stabilité pour le problème discret en domaine borné et la vérification numérique de la condition de Kreiss-Lopatinskii uniforme discrète, 2) la construction et l'utilisation de développements asymptotiques multi-échelles dans l'analyse de consistance au bord, et 3) le caractère uniforme des propriétés de stabilité au bord en présence d'une limite de relaxation.

Les deux chapitres finaux portent sur des aspects géométriques de l'asymptotique en temps grand de systèmes dynamiques pour 4) des flots de crochet isospectraux en dimension infinie, directement inspirés de la méthode QR d'approximation spectrale de matrices, et 5) le calcul de paysages de quasi-potentiels en biologie cellulaire, concernant les propriétés de multistabilité dans les mécanismes de l'hématopoïèse.

Title: Finite difference methods for hyperbolic problems with boundaries: stability and multiscale analysis

Abstract: This habilitation manuscript gathers the work I have done in recent years. They mainly focus on the study of the stability and the multiscale analysis of finite difference methods for the approximation of linear hyperbolic problems with boundaries. In such a context, various scales are likely to be present, related for example to the phenomena of viscosity, relaxation or discretization. Then, the interactions at these scales between the interior problem and the boundary of the computational domain are liable for unexpected parasitic effects, such as boundary layers. They often severely impair the stability properties in the asymptotic process and sometimes reduce the quality and the accuracy of the approximation. Therefore, it appears crucial to discriminate and rule out pathological situations.

The first three chapters relate successively to 1) the general theory of stability for the discrete problem in a bounded domain and the numerical verification of the discrete uniform Kreiss-Lopatinskii condition, 2) the construction and the use of asymptotic multi-scale expansions for the consistency analysis at the boundary, and 3) the uniform character of boundary stability properties in the presence of a relaxation limit.

The next two chapters deal with geometric aspects in the large-time asymptotic of dynamical systems for 4) isospectral bracket flows in infinite dimension, directly inspired by the QR method for spectral approximation of matrices, and 5) the computation of quasi-potential landscapes in cellular biology, for the multistability properties in the hematopoiesis mechanisms.