



**HAL**  
open science

# Étude mathématique et numérique d'équations cinétiques et fluides multi-échelles pour la description d'un plasma de fusion

Baptiste Fedele

► **To cite this version:**

Baptiste Fedele. Étude mathématique et numérique d'équations cinétiques et fluides multi-échelles pour la description d'un plasma de fusion. Analyse fonctionnelle [math.FA]. Université Paul Sabatier - Toulouse III, 2019. Français. NNT : 2019TOU30097 . tel-02507965

**HAL Id: tel-02507965**

**<https://theses.hal.science/tel-02507965>**

Submitted on 13 Mar 2020

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

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



# THÈSE

En vue de l'obtention du

## DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

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

---

---

Présentée et soutenue le 30 Septembre 2019 par :  
**Baptiste Fedele**

**Étude mathématique et numérique d'équations cinétiques et fluides multi-échelles pour la description d'un plasma de fusion**

---

---

### JURY

FRÉDÉRIQUE CHARLES	Sorbonne Université	Examinatrice
NICOLAS CROUSEILLES	INRIA Rennes	Rapporteur
BRUNO DESPRÉS	Sorbonne Université	Rapporteur
FRANCIS FILBET	Université Toulouse 3	Examineur
CLAUDIA NEGULESCU	Université Toulouse 3	Directrice de thèse
MAURIZIO OTTAVIANI	CEA Cadarache	Directeur de thèse

---

### École doctorale et spécialité :

*MITT : Domaine Mathématiques : Mathématiques appliquées*

### Unité de Recherche :

*Institut de Mathématiques de Toulouse (UMR 5219)*

### Directeurs de Thèse :

*Claudia Negulescu et Maurizio Ottaviani*

### Rapporteurs :

*Nicolas Crouseilles et Bruno Després*



*Là où la volonté est grande, les difficultés diminuent*

Nicolas Machiavel

*À mon père,*



# Remerciements

Mes premiers remerciements s'adressent tout naturellement à Claudia Negulescu et Maurizio Ottaviani pour avoir accepté d'encadrer cette thèse. Merci Claudia pour m'avoir proposé un sujet très actuel, motivant et passionnant, pour tes (très nombreuses) relectures et corrections, pour tes encouragements et tes conseils dans les moments difficiles et enfin pour avoir adapté tes attentes à mes capacités. Merci Maurizio pour nos échanges et notre travail sur la vorticité, tu prouves un peu plus que la collaboration entre les mathématiques et la physique peut être fructueuse, et merci pour ces quelques repas partagés à Aix-en-Provence.

Je suis très reconnaissant et honoré que Bruno Després et Nicolas Crouseilles aient accepté de lire et d'évaluer mon travail de recherche. Par ailleurs, je remercie chaleureusement Francis Filbet et Frédérique Charles d'avoir accepté de compléter mon jury.

Je garde un excellent souvenir de ma collaboration avec Stefan Possanner durant la première moitié de la thèse. Merci Stefan pour tes nombreux conseils, tes idées et cette soirée à Munich lors de l'*Ochtoberfest*.

Mes années à l'IMT resteront un souvenir inoubliable, en partie grâce à la vie que ce laboratoire porte en lui. J'aimerais à ce titre remercier de très nombreuses personnes que j'ai pu croiser en ces murs.

Je commence par remercier quelques membres permanents du laboratoire. Tout d'abord Christophe Besse, Francis Filbet, Pascal Noble, et Jean-Michel Roquejoffre, pour leurs cours de Master Recherche qui m'ont initié (parfois douloureusement) au monde des équations aux dérivées partielles. Je tiens à remercier également Francis et Franck Boyer sur l'intérêt qu'ils ont porté à mes recherches à quelques occasions. Merci à Etienne Fieux et François Constantino pour ces années d'enseignement partagées avec vous. Mes pensées vont également à Fanny Delebecque qui au-delà de son investissement exemplaire pour les enseignements que nous avons partagés m'aura apporté un important soutien moral à de nombreuses reprises. Je n'oublie pas d'exprimer ma gratitude envers Sophie Jan pour son investissement en faveur de l'interaction des mathématiques avec le monde industriel. Enfin, je remercie Sepideh Mirrahimi pour avoir organisé la pause café de l'équipe MIP, qui participe chaque vendredi matin à une ambiance conviviale au sein de cette dernière.

L'administration universitaire, parfois qualifiée de kafkaïenne<sup>1</sup> survit grâce à l'engagement de son personnel. Je tiens ici à remercier ces personnes dont le travail permet au laboratoire de fonctionner correctement. Tout d'abord Monique et Marie-Line pour toutes ces conversations durant ces trois années et demi, qui vont beaucoup me manquer. Je salue également leur travail quotidien au service du laboratoire. Merci à Delphine Dalla-Riva pour son dévouement à l'amélioration du laboratoire et pour sa disponibilité. Par ailleurs, je remercie grandement Janani Chandran, Sabrina Belmelih, Tamara Azaiez, Agnès Requis et Martine Labruyère pour leur soutien dans nos (trop nombreuses) tâches administratives.

Les doctorants forment un sous-ensemble fini (dont le cardinal tend à dangereusement décroître avec le temps) de l'IMT dont chaque élément est unique. Je leur témoigne à toutes et tous ma profonde gratitude pour toutes ces années partagées. Je souhaite rendre hommage à beaucoup d'entre vous, en souvenir de tous ces instants de franche camaraderie :

---

1. Mot dérivé de l'auteur Franz Kafka, dénonçant une bureaucratie angoissante et oppressive.

Tout d'abord les (ex)-occupants du bureau 302 : Guillaume, numérotien parfois malheureux, véritable doublure de Jean-Paul Rouve, avec l'intelligence d'un Kasparov, je tiens à te féliciter pour ton poste de MCF ; Marc, enseignant adulé, qui a été jusque dans l'Empire du Milieu afin d'exercer sa pédagogie ; Joachim, aussi grand de taille que consciencieux dans son travail, un homme avec qui il faut s'être échauffé la nuque afin de lui adresser la parole ; Alexis, lorgnant pendant des années sur mon bureau, qu'il va enfin pouvoir acquérir au prix d'une longue patience ; Mathias, petit frère de thèse éphémère, incompris du monde et des astres qui, un jour je l'espère, percevront enfin à quel point il est beau et intelligent (voir [1],[2]<sup>2 3</sup>).

Passons à présent aux autres doctorants MIP qui n'ont pas eu la chance de travailler dans le bureau 302, mais avec qui j'ai passé de très bons moments malgré tout. Tout d'abord Michelle, qui en plus des mathématiques possède de multiples talents (incontestables) : organisatrice de pique-niques, décoratrice de parking d'université, grossiste en pâtisserie (libanaise je crois), footballeuse (avec un taux de passes réussies frôlant les 3%), et surtout véritable atout pour la vie du laboratoire, tu as été une formidable rencontre ; Anthony, homme de gauche revendiqué, tout droit sortie d'un film d'auteur, à l'emploi du temps chargé comme celui d'un ministre ; Joe, parfois au chômage, mais toujours avec la barbe impeccable ; Susely, pâtissière de talent, spécialiste de la goyave, et avec qui je vais partager ma date de soutenance. Je salue également Sourav, Kamilia, Khuntal, Hoang Phuong, et François. Enfin je souhaite bon courage à la jeune génération : Louis, Corentin, et Perla.

Au-delà des contrées du 1R3 se trouvent deux territoires de prime abord hostiles à tout MIP<sup>4</sup> qui se respecte : le bâtiment 1R2, siège de l'équipe Émile Picard, et le bâtiment 1R1, fief de l'équipe Probabilités et Statistiques. Pourtant, je ne saurai aujourd'hui compter le nombre d'amitiés que j'ai pu nouer en ces lieux et je recommande à tout doctorant MIP de traverser la rue (ou le couloir, ça dépend) pour rejoindre ces contrées pas si lointaines. Je tiens donc à saluer un certain nombre de personnes parmi ces deux équipes.

Tout d'abord l'équipe Picard : Julie, avec qui une merveilleuse amitié s'est créée sur les marches du 1R2, qui m'a tout appris quant au mélange des équipes du laboratoire ; Florian, footballeur (canapé) passionné, collègue exemplaire pour le travail de représentant des doctorants ; Jade, aussi franche que l'or ; Damien, véritable guide vivant sur toutes les formalités à accomplir pour les doctorants ; Jean-Marc, qui a su très vite me reconnaître comme Président des doctorants ; Jules, qui m'a beaucoup épaulé à mes débuts en tant que représentant des doctorants ; Massimo, avec qui j'ai pu entretenir quelque peu mon italien. Je salue également Dominique, Dimitri, Elena, Anne, et Sara.

À présent, j'adresse un hommage à un nombre important de personnes dans mon équipe d'adoption, l'équipe ESP. En premier lieu ce cher Antoine, adopteur de MIP, artiste dessinateur de talent, dont l'œuvre inspirée située dans l'exemplaire de ton manuscrit se retrouvera très probablement dans un musée. Pierre, véritable Janus des émotions, aussi bavard en SMS qu'ordonné sur son bureau ; Fabien, comique réconfortant, recalé de l'école du rire chaque année, dont les doctorants de l'IMT servent inlassablement de public test ; Mickaël, premier MOF<sup>5</sup> du gras de France, et plus jeune président d'université de l'histoire ; Clément, vendeur de meubles sur *giphy*, qui n'a jamais été aussi drôle que lors de ce dîner

---

2. [1] M. Dus, Les plasmas de fusion : comment enfin les faire fonctionner, *Journal de Mickey*

3. [2] M. Dus, Théorie du contrôle, calculs très très compliqués, *Okapi*

4. Les termes MIP, ESP et Picard seront obsolètes à compter du 1er Janvier 2020.

5. Meilleur Ouvrier de France



"aux pieds sous la table"; Valentin, qui a joué son rôle pour m'éviter un score stalinien lors d'une élection, homme de contradiction : extrêmement talentueux en musique mais pourtant grand adepte de Patrick Sébastien; Laetitia, qui grâce à nous réalise l'exploit de connaître l'histoire de *Games of Thrones* sans en avoir vu un seul épisode; Hugo, aveyronnais de coeur, à l'accent étrange (à mes oreilles), qui a su accepter les différences culturelles avec le ch'ti que je suis; Éva, rencontrée pour la première fois à la Mécanique des Fluides, qui depuis cherche sans cesse à améliorer la vie et l'intégration des doctorants de l'institut, et je n'oublie pas cette incroyable victoire à *Résistance*; Trang, à qui je tire mon chapeau pour son apprentissage de la langue de Molière; Camille, talentueuse patineuse; William, dont l'amitié s'est créée lors d'une partie d'échecs mémorable, à Lille, et qui s'est poursuivie intarissablement au gré des (peu nombreuses) pause cafés; je reste si fier d'avoir arracher ton coeur aux griffes d'Antoine (qui s'est dégonflé pour venir me le reprendre); Maylis, avec qui j'ai formé un duo incroyable d'efficacité en TD L1 (sans parler de notre expérience de bénévolat en remise à niveau), un peu moins lors des parties en rouge de *Time Bomb* : mais qu'importe les parties perdues, j'ai gagné une grande amie. Je salue également Guillaume & Ioana, Clément B., et Laure.

Par ailleurs, merci à Florian, Jules, Antoine et Éva pour leur investissement pour la représentation des doctorants, cela aura été un honneur de partager ce ministère à vos côtés. Merci à Antoine et Jules de m'avoir accueilli au sein de ce conseil restreint. Je pars rassuré de laisser entre les mains expertes d'Éva et de Florian la responsabilité d'accompagner et d'aider tous les doctorants présents à l'IMT.

Je salue également quelques autres doctorants que j'ai pu croiser lors de ma thèse. Tout d'abord ma demi-petite soeur de thèse : Nissa (oui bien sûrrrr); mes amis italiens Giovanni et Elisa - je garde un très bon souvenir de leur passage à Toulouse - et Anouk pour ces quelques conférences partagées.

Je n'oublie pas de penser à ces amis qui n'ont pas eu la folie de se lancer dans une thèse mais avec qui je n'ai pas perdu contact! En premier lieu Valentin, complice depuis tant d'années, j'ai été très heureux de te revoir à quelques occasions ces derniers temps. Je salue également les membres du groupe de BE/TD de l'N6K : les trois Pierre, Maxime, Clément, Romain, Éva, Anthony et Colin, ainsi que mes amis italiens Tommaso et Jurgen. C'est avec beaucoup de nostalgie que je salue également mes amis de la *prépa* à Valenciennes, comme Joris, Thomas L., Thomas C., Mathilde et Manon.

Cette thèse a été l'aboutissement de mes études, qui ont commencé avec l'enseignement dispensé en classes préparatoires au Lycée Henri Wallon. Je n'oublie pas ici de remercier mes excellents professeurs durant ces deux années à Valenciennes. Merci à Christelle Duquesnoy, Anne & Freddy Szymczak pour leur enseignement exceptionnel en chimie, matière qui aujourd'hui encore me fascine malgré son absence dans mon cursus depuis mon intégration à l'N6K (je m'y suis remis en regardant la série *Breaking Bad*). Merci à PYM, "taulier" de la PCSI 1, pour sa transmission de la recherche du sens physique dans toute équation; merci à Benoît Hébert pour sa préparation de très grande qualité en physique, c'est avec plaisir que je regarde sa chaîne Youtube en certaines occasions. Merci à Pierre Joly pour sa rigueur extrême en rédaction mathématique que j'essaie encore d'appliquer aujourd'hui. Enfin, merci à Séverine Duminy qui à une époque où la discipline d'Euler, de Cauchy et de Riemann n'était pas encore dans mon coeur a su insuffler la flamme des mathématiques qui vit encore aujourd'hui.

J'aimerais pour achever mes remerciements partager mes dernières pensées avec ma famille qui fut d'une aide immense.

Mes sentiments très affectueux s'adressent tout d'abord à ma maman pour son intense soutien tout au long de la thèse, qui a su trouver les mots réconfortant dans les moments de difficulté et de doute.

Mes pensées vont également à mes frère et soeur et leur conjoint pour leur soutien et leur amour fraternel, merci beaucoup Aurélie & Jeff et Mathieu & Aurélie. C'est avec une grande affection que je salue mes trois neveux qui me comblent de joie à chaque retrouvaille : Lucas, Louis et Lorenzo.

Je remercie également toute ma famille de l'Hérault, qui m'accueille toujours avec chaleur, ça aura été un plaisir de vous rendre visite lors de ces dernières années passées à Toulouse. Merci donc à Myriam & Gilbert, Vincent, Alexandre & Laetitia, Mélissa, Hugo, et Guillaume. J'associe à ces remerciements Lisiane et sa famille qui, depuis Rome, m'ont encouragé.

Je partage aussi mes remerciements affectueux avec mon parrain Roland et avec Édith ainsi qu'à leurs enfants : Tancrede, Romain et Tristan. L'amitié entre nos deux familles demeure une chose très précieuse à mes yeux.

J'ai également une pensée pour Jean et Françoise pour leur accueil toujours chaleureux et généreux à Carcassonne.

Et c'est avec beaucoup de tendresse et d'amour que je remercie Charlotte pour son soutien indéfectible, sa grande patience et son amour inconditionnel. Dans quelques jours, la distance géographique nous séparant va enfin considérablement se réduire. Je salue également toute ma future belle-famille, Grégory & Vanessa, Camille, ainsi que Chloé & Paul.

Mes ultimes remerciements iront à mon père. Papa, tu m'as encouragé au début de la thèse à ne rien lâcher, et je suis sûr que de là où tu es aujourd'hui, tu sauras être fier de moi.



---

# TABLE DES MATIÈRES

<b>1</b>	<b>Introduction</b>	<b>15</b>
1.1	Présentation générale des plasmas et de leur modélisation . . . . .	16
1.1.1	Quelques généralités à propos des plasmas . . . . .	16
1.1.2	Descriptions mathématiques d'un système de particules . . . . .	18
1.1.3	Quelques phénomènes physiques liés aux plasmas . . . . .	22
1.2	Problèmes multi-échelles . . . . .	25
1.2.1	Motivations et cadre général . . . . .	26
1.2.2	Plusieurs exemples de problèmes multi-échelles . . . . .	27
1.2.3	Quelques méthodes analytiques et numériques pour la résolution de problèmes multi-échelles . . . . .	28
1.2.4	Schémas préservant l'asymptotique . . . . .	30
1.3	Présentation des résultats obtenus dans cette thèse . . . . .	33
1.3.1	Équation de transport anisotrope avec caractéristiques droites ( [51] Chapitre 2) . . . . .	33
1.3.2	Équation de transport anisotrope avec caractéristiques circulaires ( [51,53] Chapitre 2 & 3) . . . . .	35
1.3.3	Contributions numériques du ( <i>DAMM</i> )- <i>scheme</i> sur le système de Vlasov-Poisson ( [53] Chapitre 3). . . . .	40
1.3.4	Étude d'une équation de Vorticité-Poisson, écoulement de Kolmogorov ( [52] Chapitre 4). . . . .	42
1.3.5	Conclusion générale. . . . .	44
1.3.6	Quelques perspectives de nos travaux . . . . .	45
<b>2</b>	<b>Numerical study of an anisotropic Vlasov equation arising in plasma physics</b>	<b>49</b>
2.1	Introduction . . . . .	50
2.2	Physical motivation and toy models . . . . .	51
2.2.1	First toy model - Polar, field-aligned configuration . . . . .	53
2.2.2	Second toy model - Cartesian, not field-aligned configuration . . . . .	54
2.2.3	Aim of the present paper . . . . .	54
2.3	First anisotropic Vlasov toy model and its mathematical study . . . . .	55
2.3.1	Singularly perturbed problem . . . . .	55
2.3.2	Limit model . . . . .	56
2.3.3	Weak convergence . . . . .	56
2.3.4	Two-scale convergence . . . . .	57

2.4	Numerical schemes for the anisotropic Vlasov equation . . . . .	59
2.4.1	Semi-discretization in time . . . . .	60
2.4.2	Space discretization for the IMEX scheme . . . . .	63
2.5	Numerical simulations . . . . .	64
2.5.1	Some results obtained with our schemes . . . . .	65
2.5.2	Convergence of the schemes for fixed $\epsilon > 0$ . . . . .	67
2.5.3	Asymptotic behavior as $\epsilon \rightarrow 0$ . . . . .	68
2.6	Numerical analysis . . . . .	70
2.6.1	IMEX scheme . . . . .	70
2.6.2	Lagrange-multiplier scheme . . . . .	72
2.6.3	AP-properties . . . . .	74
2.7	Second Vlasov toy-model with variable coefficients . . . . .	74
2.7.1	Exact solution by the characteristic method . . . . .	75
2.7.2	Limit solution of the problem . . . . .	75
2.7.3	Numerical schemes for the second Vlasov toy model . . . . .	76
2.7.4	Numerical simulations . . . . .	77
2.7.5	Numerical analysis . . . . .	80
2.8	Concluding remarks . . . . .	82
<b>3</b>	<b>Asymptotic-Preserving scheme for the resolution of evolution equations with stiff transport terms</b> . . . . .	<b>84</b>
3.1	Introduction . . . . .	85
3.2	Asymptotic-Preserving reformulation . . . . .	88
3.2.1	Identification of the limit model . . . . .	88
3.2.2	Micro-Macro reformulation . . . . .	90
3.2.3	Zero mean value . . . . .	90
3.2.4	Regularization . . . . .	91
3.2.5	The stabilized AP-reformulation . . . . .	92
3.3	Some mathematical observations . . . . .	92
3.4	Numerical discretization . . . . .	95
3.4.1	Discretization parameters . . . . .	95
3.4.2	Space semi-discretization . . . . .	96
3.4.3	Time discretization . . . . .	96
3.5	Verification of the AP-scheme in a mathematical test case . . . . .	97
3.5.1	Analytical solution for both $\epsilon$ -regimes . . . . .	97
3.5.2	AP-property . . . . .	99
3.5.3	Convergence property . . . . .	100
3.5.4	Influence of the stabilization . . . . .	100
3.5.5	Choice of the stabilization parameter $\sigma$ . . . . .	103
3.6	Numerical simulations for the Vlasov-Poisson test case . . . . .	105
3.6.1	The Vlasov-Poisson system and its numerical discretization . . . . .	105
3.6.2	Numerical simulations for weak Landau damping and $\epsilon = 1$ . . . . .	106
3.6.3	Numerical simulations for strong Landau damping and $\epsilon = 1$ . . . . .	107
3.6.4	Numerical simulations of the two-stream instability, study of the limit case $\epsilon \rightarrow 0$ . . . . .	109
3.7	Concluding remarks and perspectives . . . . .	110

<b>4</b>	<b>Numerical and analytical study of a forced vorticity equation and its asymptotic-preserving resolution</b>	<b>122</b>
4.1	Motivation, background and objectives	123
4.1.1	Non-dimensional Vorticity-Poisson system	125
4.1.2	Quantities of interest	126
4.2	Numerical scheme	128
4.2.1	Asymptotic-Preserving reformulation	128
4.2.2	Numerical discretization	130
4.3	Numerical study of the Taylor-Green vortex case	133
4.3.1	The Taylor-Green vortex for large values of $\epsilon$	133
4.4	Study of a forced case - Kolmogorov flow	136
4.4.1	Linear instability of the Kolmogorov flow	136
4.4.2	Kolmogorov flow : simulations of the instability	145
4.4.3	Kolmogorov flow for vanishing values of $\epsilon$	146
4.5	Concluding remarks and perspectives	148



---

---

# CHAPITRE 1

---

## INTRODUCTION

L'objet de cette thèse concerne l'introduction et l'étude de certaines méthodes numériques appliquées à des problèmes d'évolution contenant un terme de transport raide et apparaissant dans la physique des plasmas et de la mécanique des fluides. Bien souvent, ces problèmes où interviennent de multiples échelles, ne peuvent être résolus par des méthodes numériques classiques, et nécessitent ainsi une approche différente, multi-échelles. Nous allons donc proposer dans cette thèse un nouveau schéma numérique de type AP (*i.e* préservant l'asymptotique) pour la résolution de tels problèmes et nous allons tester numériquement les performances de ce schéma dans le cadre de l'équation de Vlasov à fort champ magnétique, le système de Vlasov-Poisson à temps long ou encore de l'équation de vorticit  dans le r gime de temps visqueux.

Dans ce chapitre introductif, nous allons dans une premi re partie rappeler quelques notions utiles   la compr hension des r sultats pr sent s dans ce manuscrit. D'abord, quelques notions de base de la physique des plasmas seront introduites, et les diff rentes descriptions math matiques de ces derniers seront explicit es. Puis, nous nous attarderons sur la notion de probl me multi- chelles et de quelques m thodes (analytiques et num riques) pour les traiter. Nous pr senterons   cette occasion les m thodes num riques pr servant l'asymptotique, qui occupent une large place dans ce manuscrit. La deuxi me partie de l'introduction sera d di e plus particuli rement   la pr sentation des diff rents mod les math matiques  tudi s dans cette th se, des m thodes num riques d velopp es ainsi qu'  la synth se des r sultats obtenus.

La suite du manuscrit regroupe les trois articles publi s/soumis durant ce travail de th se :

- B. Fedele and C. Negulescu, Numerical study of an anisotropic Vlasov equation arising in plasma physics, *Kinetic and Related Models*, **11**, (2018), 1395–1426.
- B. Fedele, C. Negulescu and S. Possanner, Asymptotic-Preserving scheme for the resolution of evolution equations with stiff transport terms, *SIAM MMS (Multiscale Model. Simul.)*, **17**, (2019), 307–343.
- B. Fedele, C. Negulescu, M. Ottaviani, Numerical and analytical study of a forced vorticity equation and its asymptotic-preserving resolution, soumis en Juillet 2019.



# 1.1 Présentation générale des plasmas et de leur modélisation

## 1.1.1 Quelques généralités à propos des plasmas

Un plasma désigne un milieu composé de molécules ou d'atomes, partiellement ou totalement ionisé. Il peut être obtenu à partir d'un gaz neutre soumis à des conditions physiques bien spécifiques, notamment une très forte température, induisant l'ionisation du gaz. Le terme plasma fut employé pour la première fois en 1928 par le physicien tchèque Irving Langmuir qui établit alors un parallèle avec la notion biologique du plasma sanguin [90]. La communauté scientifique qualifie l'état plasma de quatrième état de la matière. Il s'agit de l'état dominant dans l'Univers : le coeur des étoiles, les vents solaires ou encore le gaz intergalactique se trouvent à l'état plasma. Bénéficiant de la protection du champ magnétique terrestre, nous sommes relativement épargnés par la présence de ce quatrième état de la matière dans notre proche environnement. Néanmoins, le progrès technologique de ces dernières décennies a fait multiplier la présence des plasmas dans notre quotidien, citons par exemple les arcs électriques de soudure ou encore les lampes à plasma.

Ces dernières années, l'étude des plasmas a joué un rôle majeur dans plusieurs grands projets industriels, notamment les réacteurs d'engins spatiaux [97], ou encore l'ambitieux projet ITER (*International Thermonuclear Experimental Reactor*). Ce projet international a pour objectif de concevoir un démonstrateur technologique<sup>1</sup> de la fusion nucléaire par confinement magnétique. Ce confinement doit se produire dans un *Tokamak* (du russe *toroïdalnaïa kamera magnitnymi katushkami*), une immense chambre toroïdale ceinturée par de puissantes bobines (voir Figure 1.1 pour un schéma de principe d'un Tokamak tel qu'ITER). Lors de son fonctionnement, un gaz d'éléments légers (à savoir le deutérium et le tritium, les deux isotopes de l'hydrogène) est soumis à des températures et des pressions extrêmes afin qu'il se mue en plasma, un milieu où les espèces peuvent fusionner et ainsi créer de l'énergie. En effet, le réarrangement produit par la réaction de fusion conduit à une perte (partielle) de la masse totale du système, qui est convertie en énergie via la relation  $E = mc^2$ . Les puissantes bobines placées autour de l'enceinte toroïdale permettent de maintenir le plasma chaud et loin des parois, zones où il risque de perdre ses propriétés essentielles en cas de collisions avec ces dernières. De plus, les parois pourraient être endommagées suite à de telles collisions. Les objectifs du projet ITER sont très ambitieux, le principal étant d'obtenir une quantité d'énergie cinq à dix fois plus importante à celle nécessaire au fonctionnement de la machine. ITER prévoit également de tester la production d'un des réactifs de fusion, le tritium, au sein même de l'enceinte du Tokamak. En effet, contrairement au deutérium qui est présent sur Terre en quantité très abondante, le tritium est beaucoup plus rare. Le projet ITER envisage donc de réutiliser le flux neutronique issu des réactions de fusion ainsi que des parois en lithium (qui seront installées vers la fin du projet dans la couverture du réacteur) afin de recréer le tritium consommé. Dans cette thèse, nous étudierons des modèles simplifiés de certains phénomènes présents dans un plasma de Tokamak, en particulier ceux qui sont liés aux phénomènes de turbulence, qui peuvent mettre en péril le bon fonctionnement du processus de fusion. En particulier entraver le chauffage du plasma, et donc diminuer les performances de la machine. L'intérêt d'étudier de tels modèles est qu'ils possèdent des difficultés similaires à celles des modèles

---

1. Un autre réacteur à fusion du nom de DEMO devrait ensuite être construit, il est sensé être plus proche d'un réacteur de production industriel que ne l'est ITER.

plus réalistes et donc plus complexes.

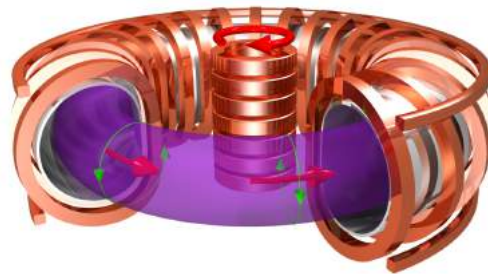


FIGURE 1.1 – Schéma d'un Tokamak.

Le comportement d'un plasma diffère sensiblement de celui d'un gaz neutre par les composés ionisés qu'il contient. Ainsi, à cause de la ionisation partielle ou totale du plasma, les forces dominantes dans ce dernier sont les forces électromagnétiques s'exerçant entre les particules chargées à longue distance. Les forces dites collisionnelles sont elles aussi des forces de nature électromagnétiques mais à courte distance, elles sont parfois négligées en raison de la faible collisionnalité d'un plasma (faible densité et températures importantes). Néanmoins, de nombreux modèles prennent également en compte les effets collisionnels dans les plasmas. Ainsi, les particules chargées, en plus des interactions avec leurs proches voisines, subissent l'interaction avec le champ électro-magnétique moyen créé par leur déplacement dans le plasma (et également d'un éventuel champ magnétique extérieur créé indépendamment).

Créer et maintenir un plasma à partir d'un gaz neutre, comme l'exige le projet ITER, n'est pas sans difficulté. Tout d'abord, une énergie considérable est nécessaire pour d'une part ioniser le gaz et d'autre part surpasser les forces répulsives à très courte distance de Van der Waals pour fusionner les particules entre elles. Mais à de tels niveaux d'énergie, des instabilités de nature électro-magnétique peuvent entraîner l'apparition de phénomènes de turbulence venant mettre en péril le confinement du plasma. L'étude de ces phénomènes, et la manière de les limiter sont donc deux éléments essentiels pour la réussite du projet ITER. Enfin, l'une des difficultés majeures inhérente à l'étude des plasmas est que dans certaines situations (comme dans les plasmas de Tokamak), la dynamique de ces derniers est fortement anisotrope : le mouvement des particules le long des lignes de champ magnétique et orthogonalement à celles-ci diffère de façon considérable. De plus, le rapport de masse très important entre les deux espèces constitutives du plasma (ions et électrons) entraîne l'existence de comportements différents de ces dernières, selon l'échelle temporelle d'observation considérée. L'un des enjeux de la simulation numérique des plasmas de Tokamak est de prendre en compte les multiples complexités de la dynamique du plasma qui viennent d'être énoncées. Un des objectifs principaux de cette thèse est de présenter quelques schémas numériques pouvant apporter une contribution à cet objectif.

Pour des ouvrages généraux sur la physique des plasmas, on pourra consulter [12, 21, 38]. Pour des ouvrages plus centrés sur la physique des Tokamaks, nous renvoyons à [26, 58, 83, 119].

## 1.1.2 Descriptions mathématiques d'un système de particules

Il existe plusieurs niveaux de description possibles afin de décrire la dynamique d'un système de particules tel que celui présent dans un plasma de Tokamak. Les paragraphes suivants ont pour objectif de les présenter succinctement.

### Approche microscopique.

Considérons un système de  $N$  particules d'une unique espèce. On note  $\mathbf{X}_i(t) \in \mathbb{R}^3$  resp.  $\mathbf{V}_i(t) \in \mathbb{R}^3$  la position resp. la vitesse de la particule  $i$  ( $1 \leq i \leq N$ ) à l'instant  $t \in \mathbb{R}^+$ . Si on note  $m$  resp.  $q$  la masse resp. la charge de l'espèce considérée, alors les champs  $\mathbf{X}_i(t)$  et  $\mathbf{V}_i(t)$  vérifient les équations de Newton :

$$(\mathcal{N}) \left\{ \begin{array}{l} \frac{d\mathbf{X}_i}{dt} = \mathbf{V}_i, \quad 1 \leq i \leq N, \\ m \frac{d\mathbf{V}_i}{dt} = q(\mathbf{E}_e(t, \mathbf{X}_i) + \mathbf{V}_i \times \mathbf{B}_e(t, \mathbf{X}_i)) + q \sum_{j \neq i}^N (\mathbf{E}_j(t, \mathbf{X}_i) + \mathbf{V}_i \times \mathbf{B}_j(t, \mathbf{X}_i)), \quad 1 \leq i \leq N, \end{array} \right.$$

où l'on a négligé toute autre force (en particulier de nature gravitationnelle) excepté les forces électro-magnétiques. Les champs électrique  $\mathbf{E}_e \equiv \mathbf{E}_e(t, \mathbf{x})$  et magnétique et  $\mathbf{B}_e \equiv \mathbf{B}_e(t, \mathbf{x})$  sont des champs externes appliqués au point  $\mathbf{x} \in \mathbb{R}^3$  à l'instant  $t$  tandis que les champs  $\mathbf{B}_j \equiv \mathbf{B}_j(t, \mathbf{x})$  et  $\mathbf{E}_j \equiv \mathbf{E}_j(t, \mathbf{x})$  sont les champs induits par la particule  $j$ , appliqués au point  $\mathbf{x}$  à l'instant  $t$ . La détermination de ces champs électro-magnétiques nécessite la résolution des équations de Maxwell :

$$(\mathcal{M}^m) \left\{ \begin{array}{l} \nabla \cdot \mathbf{E}^m = \frac{\rho^m}{\varepsilon_0}, \\ -\frac{1}{c^2} \partial_t \mathbf{E}^m + \nabla \times \mathbf{B}^m = \mu_0 \mathbf{J}^m, \\ \partial_t \mathbf{B}^m + \nabla \times \mathbf{E}^m = 0, \\ \nabla \cdot \mathbf{B}^m = 0, \end{array} \right.$$

où  $\mathbf{E}^m$  et  $\mathbf{B}^m$  désignent la résultante des champs électro-magnétiques induits par les particules du système (l'indice  $m$  indiquant qu'il s'agit de champs de nature microscopique),  $\varepsilon_0$ ,  $\mu_0$  et  $c := (\varepsilon_0 \mu_0)^{-1/2}$  représentent respectivement la constante de permittivité électrique, la constante de perméabilité magnétique et la vitesse de la lumière. La densité de charge  $\rho$  et le vecteur courant  $\mathbf{J}$  sont définis via :

$$\rho^m(t, \mathbf{x}) = \sum_{i=1}^N q \delta(\mathbf{x} - \mathbf{X}_i), \quad \mathbf{J}^m(t, \mathbf{x}) = \sum_{i=1}^N q \mathbf{V}_i \delta(\mathbf{x} - \mathbf{X}_i).$$

Dans le cas particulier (mais néanmoins courant dans de nombreux modèles) où les perturbations du champ magnétique sont négligées, le champ électrique  $\mathbf{E}^m$  dérive alors d'un potentiel scalaire  $\phi^m \equiv \phi^m(t, \mathbf{x})$  :

$$\mathbf{E}^m = -\nabla_{\mathbf{x}} \phi^m,$$

et ce potentiel vérifie alors l'équation de Poisson :

$$-\varepsilon_0 \Delta_{\mathbf{x}} \phi^m = \rho^m.$$

Cette description, bien que très précise pour décrire la dynamique de la population du système devient très peu pratique lorsque le nombre de particules est trop important. En particulier, toute tentative de résolution numérique d'un tel système d'équations lorsque  $N \rightarrow \infty$  se révèle infaisable. En guise d'exemple, le Tokamak en construction à travers le projet ITER prévoit une densité de  $10^{14}$  particules par centimètre cube [119]. La description microscopique n'est donc pas adaptée pour décrire la dynamique des populations denses d'un Tokamak. Une autre description mathématique est donc nécessaire, cette dernière devra offrir un cadre plus propice à la simulation numérique tout en modélisant de manière satisfaisante les phénomènes physiques que l'on souhaite étudier.

### Approche mésoscopique.

En réponse au problème soulevé dans le paragraphe précédent, l'approche mésoscopique est une alternative utile afin d'étudier la dynamique des particules. Elle s'appuie sur une description statistique de ces dernières via une fonction de distribution. Dans cette approche, appelée également approche cinétique, il n'est plus question de connaître la position et la vitesse de chaque particule  $i$  du système à chaque instant  $t$ . L'inconnue majeure est la fonction de distribution notée  $f_\alpha(t, \mathbf{x}, \mathbf{v})$ . Physiquement, la quantité  $f_\alpha(t, \mathbf{x}, \mathbf{v}) d\mathbf{x}d\mathbf{v}$  correspond au nombre moyen d'espèces  $\alpha$  contenu dans le volume infinitésimal  $d\mathbf{x}d\mathbf{v}$  centré au point  $(\mathbf{x}, \mathbf{v})$  dans l'espace de phase. La fonction de distribution  $f_\alpha$  vérifie l'équation dite de Vlasov pour l'espèce  $\alpha$  :

$$\partial_t f_\alpha + \mathbf{v} \cdot \nabla_{\mathbf{x}} f_\alpha + \frac{q_\alpha}{m_\alpha} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \nabla_{\mathbf{v}} f_\alpha = 0, \quad (t, \mathbf{x}, \mathbf{v}) \in \mathbb{R}^+ \times \Omega \times \mathbb{R}^d, \quad (1.1)$$

où  $\Omega$  est un sous-ensemble de  $\mathbb{R}^d$  avec  $d \leq 3$  et  $\alpha$  désigne dans le cadre de la physique des plasmas soit des ions soit des électrons. En particulier,  $q_\alpha = \pm e$  avec  $e$  la charge élémentaire et  $m_\alpha$  désigne soit la masse de l'électron ou de l'ion. Les champs électromagnétiques  $\mathbf{E}$  et  $\mathbf{B}$  sont déterminés via les équations de Maxwell moyennées :

$$(\mathcal{M}) \begin{cases} \nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0}, \\ -\frac{1}{c^2} \partial_t \mathbf{E} + \nabla \times \mathbf{B} = \mu_0 \mathbf{J}, \\ \partial_t \mathbf{B} + \nabla \times \mathbf{E} = 0, \\ \nabla \cdot \mathbf{B} = 0. \end{cases}$$

La densité de charge  $\rho$  et le vecteur courant  $\mathbf{J}$  sont reliés avec la fonction de distribution via :

$$\rho(t, \mathbf{x}) := \sum_{\alpha} q_{\alpha} \int_{\mathbb{R}^3} f_{\alpha}(t, \mathbf{x}, \mathbf{v}) d\mathbf{v}, \quad \mathbf{J}(t, \mathbf{x}) := \sum_{\alpha} q_{\alpha} \int_{\mathbb{R}^3} \mathbf{v} f_{\alpha}(t, \mathbf{x}, \mathbf{v}) d\mathbf{v}.$$

Notons que l'équation de Vlasov (1.1) peut-être obtenue simplement à partir des équations de Newton lorsque l'on considère uniquement des champs électromagnétiques extérieurs. Lorsque l'on considère en plus les interactions entre les particules, le passage des équations microscopiques au modèle mésoscopique est bien plus délicat.

L'équation de Vlasov (1.1) ne considère pas les effets collisionnels entre les particules. Afin d'en tenir compte, on modifie l'équation de Vlasov (1.1) de la façon suivante :

$$\partial_t f_\alpha + \mathbf{v} \cdot \nabla_{\mathbf{x}} f_\alpha + \frac{q_\alpha}{m_\alpha} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \nabla_{\mathbf{v}} f_\alpha = \mathcal{Q}_\alpha(f_\alpha). \quad (1.2)$$

où  $\mathcal{Q}_\alpha$  un opérateur de collision. Les premiers travaux sur les noyaux de collisions ont eu lieu à la fin des années 50 [120]. Nous renvoyons également à l'article de Villani [127] pour les aspects mathématiques de plusieurs noyaux de collision. Dans cette thèse, les équations du type (1.2) ne seront pas étudiées. C'est un choix courant et justifié dans l'étude des plasmas de Tokamak, qui intervient lorsque la fréquence des collisions est très inférieure à la fréquence des phénomènes considérés.

Bien que les modèles mésoscopiques (ou modèles cinétiques) ne se basent plus sur la détermination des grandeurs dynamiques (position et vitesse notamment) de chaque particule, contrairement au modèle microscopique, il demeure qu'ils nécessitent un important coût numérique dû au fait que le système possède toujours 6 degrés de liberté (sans compter le temps). Afin de réduire encore la complexité du modèle, il est possible dans le cas de plasmas fortement magnétisés de réduire le nombre de dimensions du système. Il est bien connu que le mouvement d'une particule dans un champ magnétique se décompose en un mouvement le long des lignes de champ et d'un mouvement giratoire autour de ces lignes. La trajectoire de la particule épouse donc la forme d'une hélice (voir Figure 1.2 pour la vue 3D et Figure 1.3 pour la vue projetée dans le plan  $(x, y)$ ). On décompose ainsi le vecteur vitesse  $\mathbf{v}$  de la particule selon un vecteur parallèle aux lignes de champ  $\mathbf{v}_{//}$  et un vecteur orthogonal à ces lignes  $\mathbf{v}_\perp$ . Ce dernier est fonction de l'angle de giration obtenu par changement de coordonnées du système. Si on moyenne sur ce dernier, alors l'équation de Vlasov se réduit à une équation dite giro-cinétique, qui modélise l'évolution du giro-centre dans un espace de phase ayant une dimension de moins que celui dans l'équation de Vlasov. Ce modèle n'est valide que lorsque l'échelle de temps des phénomènes observés est supérieure à l'inverse de la fréquence cyclotron  $\omega_c$  donnée par :

$$\omega_c := \frac{|q_\alpha|B}{m_\alpha},$$

qui est l'échelle temporelle caractéristique du mouvement giratoire, et où  $B$  correspond à l'intensité du champ magnétique. Cette approximation a été introduite historiquement par Littlejohn [96]. L'analyse mathématique de cette approximation giro-cinétique a été étudiée principalement à l'aide de techniques d'homogénéisation et d'analyse asymptotique [59, 60, 63, 65]. Cette approximation est de plus utilisée massivement dans l'élaboration de codes de calculs pour la simulations numériques des Tokamaks et des Stellarators<sup>2</sup> [62, 66, 71]. Dans les chapitres 2 et 3 de ce manuscrit, nous étudierons analytiquement et surtout numériquement un modèle simplifié de l'équation de Vlasov, en ne passant pas par l'approximation giro-cinétique.

---

2. Il s'agit d'une technologie alternative au Tokamak pour effectuer la fusion nucléaire à but civil.

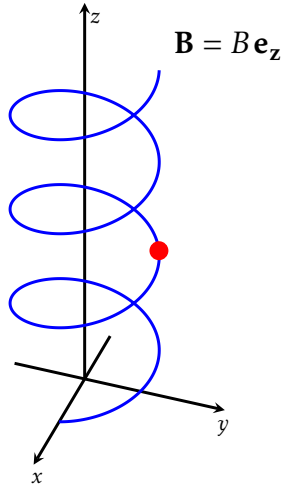


FIGURE 1.2 – Trajectoire (en bleu) d’une particule chargée (point en rouge) autour d’une ligne de champ magnétique verticale dirigée selon l’axe  $z$ .

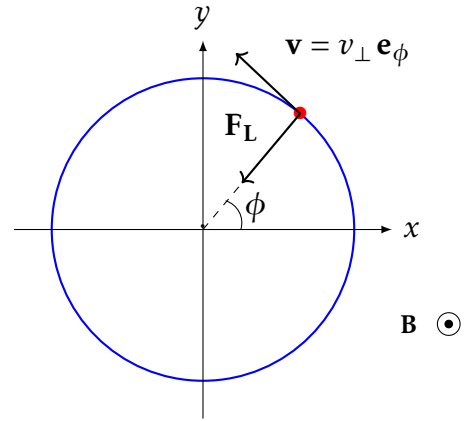


FIGURE 1.3 – Projection sur le plan  $(x, y)$  du mouvement de la particule chargée (point en rouge).  $F_L$  correspond à la force de Lorentz et  $v_{\perp}$  à la composante de la vitesse perpendiculaire au champ  $\mathbf{B}$ .

La description mathématique choisie dans cette thèse pour l’étude d’un plasma de Tokamak est l’approche cinétique. Nous nous intéresserons en particulier au Chapitre 3 au système de Vlasov-Poisson, constituant un modèle cinétique classique de la physique des plasmas, et qui permet de considérer les interactions électrostatiques auto-induites du système de particules. Pour une espèce  $\alpha$ , l’équation de Vlasov-Poisson avec champ magnétique extérieur (indépendant du temps) s’écrit :

$$(\mathcal{VP}) \begin{cases} \partial_t f_{\alpha} + \mathbf{v} \cdot \nabla_{\mathbf{x}} f_{\alpha} + \frac{e_{\alpha}}{m_{\alpha}} (\mathbf{E} + \mathbf{v} \times \mathbf{B}_{\mathbf{e}}) \cdot \nabla_{\mathbf{v}} f_{\alpha} = 0, \\ \mathbf{E} = -\nabla_{\mathbf{x}} \phi, \\ -\varepsilon_0 \Delta_{\mathbf{x}} \phi = \rho. \end{cases} \quad (1.3)$$

Une des simplifications mathématique courante de ce modèle est de considérer le champ magnétique extérieur  $\mathbf{B}_{\mathbf{e}}$  nul<sup>3</sup>. Nous retiendrons cette simplification dans le Chapitre 3. Ce système a été étudié de manière approfondie par la communauté scientifique tant au niveau mathématique que numérique. Sur les problématiques d’existence et d’unicité des solutions de ce système ainsi que de leur régularité, on peut citer les travaux d’Arsenev [9], de P. E. Jabin [79], de Diperna et P.L. Lions [44, 45]. À propos des travaux axés sur la résolution numérique de ce système, on peut se référer à la liste de publications suivantes [15, 85, 86, 123, 130]. Il est également possible d’étudier le système de Vlasov-Maxwell, où les effets des champs magnétiques externes et/ou auto-induits sont pris en compte. On pourra se référer aux travaux de F. Golse et L. Saint-Raymond [65] ou encore ceux de Diperna et P. L. Lions [43].

### Approche macroscopique.

Dans ce paragraphe, la question de l’approche macroscopique ou fluide est abordée. Les modèles macroscopiques peuvent être vus comme des limites des modèles cinétiques abor-

3. Pour ITER, cette simplification n’est pas pertinente, une première alternative, plus réaliste, est de considérer un champ  $\mathbf{B}_{\mathbf{e}}$  homogène.

dés précédemment dans le cadre fortement collisionnel. Ils sont de ce fait plus attractifs d'un point de vue numérique mais perdent des informations physiques contenues dans les modèles plus précis, cinétiques. Pour aboutir à des modèles fluides, on effectue des bilans de matières dans des volumes élémentaires afin d'aboutir à des équations aux dérivées partielles faisant apparaître des grandeurs moyennées, telles que la densité  $n_\alpha(t, \mathbf{x})$ , le champ de vitesse  $\mathbf{u}_\alpha(t, \mathbf{x})$  ou encore la température  $T_\alpha(t, \mathbf{x})$ . Ces grandeurs ne sont ni plus ni moins que les premiers moments en vitesse de la fonction de distribution  $f_\alpha$  évoquée dans l'approche cinétique :

$$\begin{aligned} n_\alpha(t, \mathbf{x}) &:= \int_{\mathbb{R}^d} f_\alpha(t, \mathbf{x}, \mathbf{v}) d\mathbf{v}, \\ \mathbf{u}_\alpha(t, \mathbf{x}) &:= \frac{1}{n_\alpha(t, \mathbf{x})} \int_{\mathbb{R}^d} \mathbf{v} f_\alpha(t, \mathbf{x}, \mathbf{v}) d\mathbf{v}, \\ T_\alpha(t, \mathbf{x}) &= \frac{m_\alpha}{d k_B n_\alpha(t, \mathbf{x})} \int_{\mathbb{R}^d} |\mathbf{v} - \mathbf{u}_\alpha(t, \mathbf{x})|^2 f_\alpha(t, \mathbf{x}, \mathbf{v}) d\mathbf{v}, \end{aligned}$$

où  $k_B$  désigne la constante de Boltzmann. Certains phénomènes inhérents aux plasmas, comme certains effets dissipatifs ne peuvent être pris en compte par une modélisation fluide. Ces phénomènes se basent en effet sur des mécanismes de résonance (interaction ondes-particules) pour lesquels toutes les particules d'une même distribution n'ont pas le même comportement. Seules les particules dont la vitesse est proche de la vitesse de phase des ondes vont pouvoir effectuer des échanges d'énergie. Cet effet est totalement invisible dans une modélisation fluide qui ne considère que l'évolution des moments de la fonction de distribution. Dans cette thèse, les modèles fluides pour les plasmas ne seront pas abordés. L'avantage des modèles fluides réside surtout dans le faible nombre de dimension du système à considérer, ce qui numériquement se traduit par des coûts de calcul raisonnables.

### 1.1.3 Quelques phénomènes physiques liés aux plasmas

Dans cette partie, nous étudions l'origine physique de certains phénomènes liés aux plasmas illustrés dans cette thèse. Rappelons dans un premier temps que l'enjeu de la fusion thermonucléaire dans un Tokamak est de maintenir le plasma à une température suffisamment élevée, afin de contrer les forces répulsives qui tendent à éloigner les espèces les unes des autres et ainsi de favoriser l'approche des noyaux pour que la force nucléaire soit suffisante afin de déclencher la réaction de fusion. Dans de telles conditions de température, de nombreuses ondes (de nature électro-magnétiques entre autres) sont présentes dans le plasma, ces dernières peuvent alors, lorsque certains modes sont excités, causer la dégradation du confinement que l'on souhaite établir. L'étude de ces phénomènes de nature ondulatoire est donc primordiale dans la perspective d'obtenir un confinement efficace. Il est tout particulièrement important de comprendre comment certains modes de ces ondes sont excités, et lorsque c'est le cas, s'ils sont amplifiés ou amortis.

Dans cette partie, on considère un plasma constitué d'ions et d'électrons, supposé globalement neutre, *i.e* les densités électronique et ionique sont quasiment les mêmes. Les ions seront supposés de masse infinie, et seront considérés comme immobiles par rapport aux électrons. De plus on supposera dans cette partie que le champ magnétique est infiniment grand, *i.e* que les électrons ont un mouvement purement longitudinal, le long des lignes de champ magnétique.

Lorsque l'on perturbe la densité d'un tel plasma en équilibre, les forces électrostatiques naissantes entraînent un mouvement oscillatoire des électrons. La fréquence de ces oscillations, appelée fréquence plasma est donnée par :

$$\omega_p := \sqrt{\frac{\bar{n} e^2}{\varepsilon_0 m_e}}, \quad (1.4)$$

avec  $\bar{n}$  la densité particulière, on trouvera une démonstration de ce résultat dans [26]. Ce mouvement oscillatoire très simple ne prend pas en compte un éventuel mouvement de nature thermique. Lorsque que ce dernier est présent, il est parfois nécessaire de considérer des modèles cinétiques afin d'expliquer de nombreux phénomènes d'interactions ondes-particules dans le plasma. Notant  $\theta$  la température caractéristique des électrons, la vitesse thermique caractéristique de ceux-ci est alors donnée par :

$$V_{th} := \sqrt{\frac{k_B \theta}{m_e}}.$$

On se place désormais dans le cadre où les électrons ont une température caractéristique  $\theta$ .

On considère un plasma en parfait équilibre thermodynamique avec une fonction de distribution  $f_0(\mathbf{v})$ . À l'instant  $t = 0$ , on perturbe cet équilibre via une fonction de distribution  $f_1(t, \mathbf{x}, \mathbf{v})$ . En considérant que la fluctuation est de faible amplitude, il est possible de linéariser le système de Vlasov-Poisson et d'obtenir une relation de dispersion de la forme

$$D(k, \omega/k) = 0, \quad (1.5)$$

où  $k$  désigne le nombre d'onde et  $\omega$  la fréquence de l'onde. Cette relation fera directement apparaître la fonction de distribution d'équilibre  $f_0(\mathbf{v})$ , on verra alors dans la suite que la forme de cet équilibre de départ peut jouer un rôle majeur dans l'évolution de la dynamique du plasma.

La totalité des phénomènes qui seront présentés dans la suite (amortissement Landau, instabilité bi-courant et équilibres BGK) concernent des interactions ondes-particules suite à la perturbation d'un état d'équilibre d'un plasma.

### Amortissement Landau.

Ce phénomène physique a été introduit par Landau [89] et possède la remarquable propriété d'être un amortissement de nature non collisionnelle. L'amortissement Landau est la décroissance exponentielle en temps de toute les petites perturbations dans un plasma stable, en particulier l'amplitude des ondes de Langmuir (qui correspondent aux oscillations de densité décrites dans la sous-section précédente).

En particulier, Landau a montré qu'un plasma possédant une distribution d'électrons ayant une pente négative au point correspondant à la vitesse de phase de l'onde est stable. C'est le cas en particulier d'une distribution Maxwellienne de vitesse (voir Figure 1.4). Néanmoins, cette condition de stabilité n'est que suffisante.

On pourra consulter les ouvrages [1, 26] pour un développement plus précis sur l'amortissement Landau, ainsi que l'article [50] qui adopte un point de vue microscopique. Pour



une approche mathématique de l’amortissement Landau, on renvoie aux travaux de C. Mouhot et C. Villani [128].

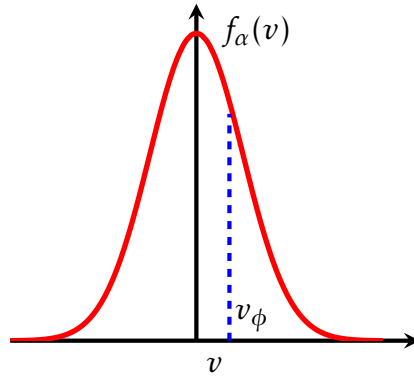


FIGURE 1.4 – Distribution Maxwellienne.

### Instabilité bi-courant.

On considère ici deux courants d’électrons d’égale densité mais se déplaçant en sens contraire dans un fond continu d’ions immobiles (voir Figure 1.5). Deux cas sont alors possibles, soit on considère un plasma chaud et les courants d’électrons possèdent alors un élargissement thermique dans leur fonction de distribution, soit on considère un plasma froid, et les fonctions de distributions sont associées à des distributions de Dirac. Ce dernier cas n’est pas illustré dans cette thèse. Contrairement à l’amortissement Landau, si la vitesse de phase  $v_\phi$  est bien choisie, une instabilité peut alors apparaître. Sur la Figure 1.5, on voit que pour la vitesse de phase choisie, il existe cette fois davantage de particules ayant une vitesse légèrement supérieure à  $v_\phi$  que de particules ayant une vitesse légèrement inférieure à  $v_\phi$ . Ainsi, les particules vont transmettre de l’énergie à l’onde qui va s’amplifier. Cette instabilité n’est néanmoins possible que si la vitesse relative entre les deux courants est supérieure à la vitesse thermique des distributions maxwelliennes [29]. Dans la limite du plasma froid (où  $V_{th} \approx 0$ ), la théorie linéaire, via la résolution de l’équation de dispersion (1.5), nous donne alors la fréquence et le taux de croissance des modes instables (néanmoins les résultats obtenus peuvent être comparés à des simulations effectuées pour des plasmas chauds, les effets thermiques ne jouant pas un rôle majeur dans cette instabilité [29]). Une fois que l’onde est suffisamment amplifiée, la théorie linéaire n’est plus valide. Dans ce cas, des structures cohérentes peuvent apparaître dans l’espace des phases, tels que les trous d’électrons. En effet, l’énergie électronique est donnée par  $\mathcal{E} = 1/2 m_e^2 v^2 - e\varphi(x)$ , où  $\varphi$  est le potentiel du champ électrique. Puisque les électrons résonnants (ceux qui interagissent avec l’onde) perdent de l’énergie au profit de l’onde, certaines zones de l’espace seront telles que  $\mathcal{E} < 0$ , et les électrons résonnants seront piégés dans ces zones (ils n’ont plus assez d’énergie cinétique pour sortir du puits de potentiel de l’onde) [26]. L’instabilité finit alors par saturer lorsque tous les électrons résonnants ont transmis leur énergie à l’onde. Ce phénomène fortement non-linéaire ne peut être capturé que par des modèles cinétiques. Pour plus de détails sur ce type d’instabilité, on pourra consulter les références indiquées dans ce paragraphe ainsi que [77], [5] et [38].

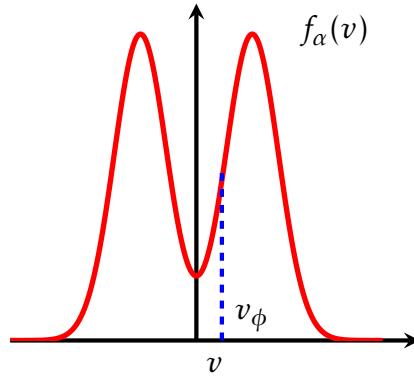


FIGURE 1.5 – Distributions maxwelliennes à deux courants opposés.

### Équilibres de Bernstein-Green-Kruskal

Les équilibres de Bernstein-Green-Kruskal (on utilisera désormais l'acronyme BGK pour les mentionner) ou ondes BGK sont des équilibres spatialement non-homogènes du système de Vlasov-Poisson introduits pour la première fois dans les travaux de Bernstein, Green et Kruskal [14]. Leur apparition découle de phénomènes de nature fortement non-linéaires, de fait, leur étude dépasse le cadre proposé par la théorie linéaire [29], consistant principalement à résoudre l'équation de dispersion (1.5), obtenue après linéarisation des équations de Vlasov. Les équilibres BGK peuvent être considérés comme des limites asymptotiques en temps de l'instabilité bi-courant évoquée précédemment [110, 116]. Ils sont ainsi reliés au phénomène de piégeage d'électrons évoqué dans le précédent paragraphe. Afin de mieux expliquer ici ce phénomène, on s'appuie sur les Figures 1.6 et 1.7. On a représenté sur la Figure 1.6 un plan de phase  $(x, v)$  possible d'une onde BGK, où les zones de piégeage (celles entourées par des lignes de niveaux fermées) ainsi que les zones libres (celles où les lignes de niveaux sont ouvertes) sont clairement identifiables. On constate bien que, pour un potentiel  $\varphi = \varphi_1$  fixé, les zones de piégeage correspondent aux zones de faibles vitesses (correspondant à une faible énergie cinétique) d'une part, et aux points de délimitation  $T_1$  et  $T_2$  (Figures 1.6 et 1.7) d'autre part. Les ondes BGK ont également fait l'office de nombreuses simulations numériques, on pourra se référer aux publications suivantes [40, 41, 74, 100, 103, 104]. Un résultat important de cette thèse correspond à l'obtention numérique de ces ondes BGK à partir de l'instabilité bi-courant en un temps de calcul très rapide grâce aux propriétés préservant l'asymptotique du schéma numérique que nous développerons.

## 1.2 Problèmes multi-échelles

À travers ces premiers paragraphes concernant les plasmas, nous avons montré le cadre fortement multi-échelles les concernant. En effet, de nombreuses échelles temporelles et spatiales coexistent au sein d'un plasma : nous avons déjà évoqué la fréquence des collisions, la fréquence plasma ainsi que la fréquence cyclotronique (différente selon l'espèce considérée) mais nous pouvons citer également la longueur caractéristique du milieu (comme le Tokamak par exemple), ou encore la longueur de Debye. Cette dernière correspond à l'ordre de grandeur spatial au delà duquel on obtient la neutralité du plasma.

Ainsi, dans ce manuscrit, notre intérêt se porte sur divers problèmes multi-échelles issus de la physique des plasmas et de la mécanique des fluides. Dans cette partie, nous allons définir la notion générale de problème multi-échelles et décrire les méthodes tant analy-

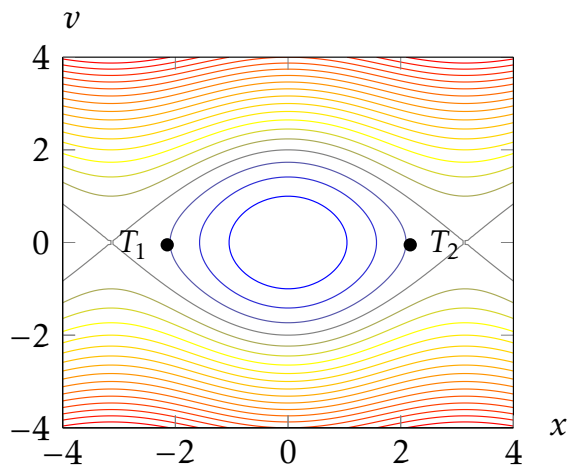


FIGURE 1.6 – Plan de phase  $(x, v)$  explicitant le phénomène de piégeage d'électrons.

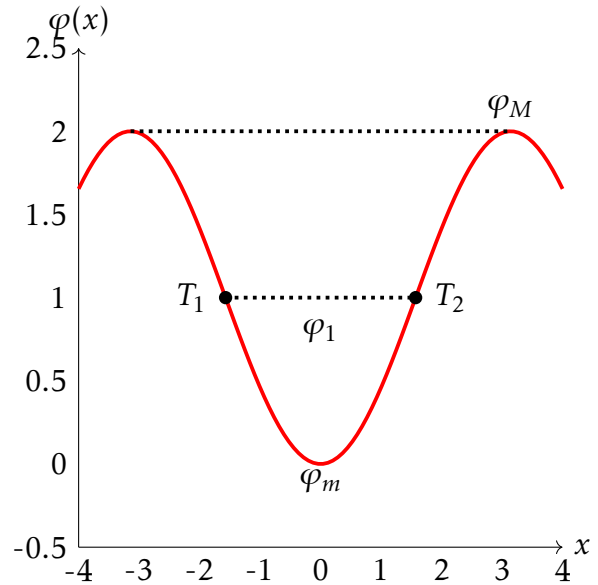


FIGURE 1.7 – Potentiel  $\varphi$ .

tiques que numériques mises en oeuvre afin de traiter ce type de problème (qui peuvent apparaître dans bien d'autres domaines que ceux précités, notamment en mécanique des structures). Nous nous attarderons en particulier sur les méthodes préservant l'asymptotique (dites méthodes AP), qui occupent une place majeure dans les travaux de cette thèse. Pour des ouvrages généraux sur les systèmes multi-échelles, on pourra consulter les ouvrages de Holmes, LeBris ou Weinan [23, 76, 129].

### 1.2.1 Motivations et cadre général

Un problème multi-échelle est un problème physique où des échelles temporelles et/ou spatiales très disparates coexistent. Dans l'approche classique d'une modélisation d'un problème physique, une échelle d'observation de ce dernier est privilégiée. Les effets des échelles plus fines sont pris en compte à travers des équations d'état, tandis que ceux des échelles plus larges sont négligés, invoquant l'homogénéité des grandeurs physiques à ces échelles. Par exemple, pour les fluides newtoniens, le tenseur des contraintes agissant sur ceux-ci est considéré comme un opérateur linéaire par rapport au gradient de vitesse du fluide, mais cette équation d'état échoue à rendre compte des effets microscopiques réels pour des fluides non-newtoniens. L'approche multi-échelles admet quant à elle l'existence d'une hiérarchie de modèles d'un même problème permettant son étude de manière approfondie.

Mentionnons pour clore cette introduction que les problèmes multi-échelles peuvent être classés selon deux grandes familles : celle considérant des problèmes possédant des singularités locales, comme des chocs, des couches limites, etc. (problèmes de type A), et celle considérant des problèmes où coexistent des échelles microscopiques et macroscopiques au sein d'un même domaine, comme par exemple les problèmes fortement oscillants (problèmes de type B).

## 1.2.2 Plusieurs exemples de problèmes multi-échelles

Nous allons présenter dans cette partie plusieurs exemples de problèmes, où dans chaque cas, un paramètre physique noté  $\epsilon$  témoigne du caractère multi-échelles du problème considéré.

### Limite hydrodynamique.

Dans un gaz avec un grand nombre de particules, la fonction de distribution  $f^\epsilon \equiv f^\epsilon(t, \mathbf{x}, \mathbf{v})$  vérifie l'équation de Boltzmann suivante :

$$\partial_t f^\epsilon + \mathbf{v} \cdot \nabla_{\mathbf{x}} f^\epsilon = \frac{1}{\epsilon} Q(f^\epsilon),$$

où  $Q$  désigne un opérateur de collision. Le paramètre  $\epsilon$  est le libre parcouru moyen  $l$  adimensionné avec l'échelle caractéristique de longueur  $L$ , ce paramètre sans dimension est appelé nombre de Knudsen :

$$\text{Kn} := \frac{l}{L} = \epsilon, \quad \epsilon \in (0, 1).$$

L'opérateur de collision agit comme une relaxation qui conduit le système à des équilibres thermodynamiques locaux : des solutions  $f^0$  telles que  $Q(f^0) = 0$ . À la limite  $\epsilon \rightarrow 0$ , cette équation tend vers le système d'Euler compressible qui donne l'évolution des quantités macroscopiques suivantes : densité, vitesse, température, etc. On pourra consulter les thèses de F. Filbet et L. Navoret [54, 111] pour plus de détails sur ce problème.

### Limite en champ magnétique intense.

On considère ici l'équation de Vlasov d'inconnue  $f^\epsilon$ , représentant la fonction de distribution des ions :

$$\partial_t f^\epsilon + \mathbf{v} \cdot \nabla_{\mathbf{x}} f^\epsilon + \left( \mathbf{E} + \frac{1}{\epsilon} \mathbf{v} \times \mathbf{B} \right) \cdot \nabla_{\mathbf{v}} f^\epsilon = 0,$$

où  $\epsilon$  est cette fois la fréquence cyclotronique  $\omega_c$  adimensionnée avec une échelle de temps caractéristique, notée  $T$  :

$$\epsilon := \frac{1}{T \omega_c}, \quad \epsilon \in (0, 1).$$

Cette fois, l'équation précédente n'est plus diffusive, mais fortement oscillante lorsque  $\epsilon \ll 1$ . Nous étudierons un modèle simplifié de cette équation dans les Chapitres 2 et 3 de ce manuscrit. On pourra également consulter les travaux de Bostan sur le sujet [20].

### Limite temps-visqueux de l'équation de vortacité.

En deux dimensions, la dynamique d'un fluide est régie par l'équation de vortacité :

$$\partial_t \omega^\epsilon + \frac{1}{\epsilon} \{ \omega^\epsilon, \Psi^\omega \} = \Delta \omega^\epsilon - \Delta \omega_{eq},$$

où  $\omega^\epsilon$  désigne le champ de vortacité,  $\omega_{eq}$  un terme de forçage supposé indépendant de  $\epsilon$ , et  $\Psi^\epsilon$  la fonction courant reliée par le champ de vortacité via l'équation de Poisson :

$$-\Delta \Psi^\epsilon = \omega^\epsilon.$$

Ici, le paramètre  $\epsilon$  est défini comme le rapport entre le temps d'observation  $T$  et le temps visqueux  $\tau_\nu$  :

$$\epsilon := \frac{T}{\tau_\nu} = \frac{T \nu}{L^2},$$

avec  $\nu$  la viscosité du fluide et  $L$  une échelle spatiale caractéristique du problème. La limite  $\epsilon \rightarrow 0$  correspond au cas où le temps visqueux est bien plus long que le temps d'observation. Dans le Chapitre 4 de ce manuscrit, ce problème avec forçage sinusoïdal sera abordé.

### 1.2.3 Quelques méthodes analytiques et numériques pour la résolution de problèmes multi-échelles

Dans cette partie, on présente brièvement quelques techniques analytiques et numériques développées dans le but de traiter les problèmes multi-échelles.

#### Méthodes de raccordement asymptotique.

Ces méthodes trouvent leur origine dans le cadre de la mécanique des fluides. Historiquement, Prandtl (1905) [118] fut le premier à utiliser ce type de techniques dans le cadre d'un écoulement d'air sur une voilure d'avion. Les équations régissant un fluide visqueux étant délicates à résoudre, Prandtl invoqua le fait que les effets visqueux n'étaient significatifs que dans une zone restreinte de l'espace, proche de la paroi solide, appelée couche limite afin de simplifier son problème qu'il pût ensuite résoudre. Les décennies suivantes permirent à d'autres pionniers de développer ces techniques à travers une multitude de problèmes physiques différents. À partir des années 1940, de nombreux chercheurs enrichirent la théorie de ces méthodes, par exemple Cole and Lagerstrom sur des problèmes d'écoulements visqueux et compressibles [88]. Kaplun et Lagerstrom se sont ensuite concentrés sur des écoulements à bas Reynolds en usant des techniques de raccordement asymptotiques [82]. Van Dick compléta la théorie avec un ouvrage de référence en 1964 [47]. Citons pour terminer Ashley et Landahl qui ont apporté de nombreuses contributions à l'aide de ces techniques sur des problèmes d'aérodynamique [10]. À présent, nous présentons brièvement le principe de ces méthodes.

Considérons une grandeur physique quelconque  $G^\epsilon$ , typiquement la solution d'une équation différentielle avec  $\epsilon$  un paramètre du problème, pouvant être petit. Le domaine physique peut alors être scindé en une partie dite externe où les variations de  $G^\epsilon$  sont d'ordre unité et une partie interne où les variations de  $G^\epsilon$  sont très rapides. Il est alors possible d'obtenir deux modèles différents dans chacune des zones du domaine, avec une solution correspondante. Les deux solutions internes et externes doivent ensuite être raccordées dans un domaine intermédiaire (appelé couramment *overlapping region*), *i.e.*, un domaine où les deux solutions sont valides asymptotiquement. On pourra se référer pour une introduction détaillée de ces méthodes aux ouvrages de références [47, 76, 87]. Ces méthodes sont propices dans le cadre des problèmes multi-échelles de type A. Nous illustrerons dans ce manuscrit au Chapitre 4 cette méthode sur un problème d'instabilité linéaire d'un écoulement de type *shear-flow*, le paramètre  $\epsilon$  sera dans ce cadre lié au rapport d'aspect du domaine spatial.

## Méthode d'homogénéisation.

Parmi les problèmes multi-échelles, ceux mettant en jeu à la fois une dynamique oscillatoire très rapide et une dynamique de dérive plus lente sont courants, c'est le cas du mouvement d'une particule confinée par un champ magnétique intense. Ce cadre est propice à l'utilisation des méthodes dites d'homogénéisation qui sont ainsi utiles à la résolution des problèmes multi-échelles de type B. Ces méthodes sont apparues dans le cadre de l'ingénierie des matériaux composites possédant naturellement plusieurs échelles caractéristiques : l'une microscopique et l'autre macroscopique, le ratio entre ces deux échelles pouvant être très faible.

Plus généralement, considérons un opérateur différentiel  $L_\epsilon$  dépendant d'un paramètre  $\epsilon \in (0, 1)$ , et une famille de solutions  $u^\epsilon$  vérifiant l'équation :

$$L_\epsilon u^\epsilon = f \quad \text{dans } \Omega, \quad (1.6)$$

où  $f$  est un terme source quelconque et  $\Omega$  un domaine spatial. Si nous supposons que la solution  $u^\epsilon$  converge (dans un certain sens) lorsque  $\epsilon \rightarrow 0$  vers une fonction  $u$ , nous sommes amenés à chercher un opérateur limite  $L$  tel que  $u$  soit solution de l'équation homogénéisée :

$$L u = f \quad \text{dans } \Omega. \quad (1.7)$$

Le passage entre ces deux équations (1.6) et (1.7), ainsi que la détermination de l'opérateur  $L$  constitue une méthode d'homogénéisation. L'équation (1.7) ne contient plus explicitement les deux échelles d'observation mais induit dans sa solution  $u$  les effets moyens de l'échelle microscopique. Une des premières méthodes d'homogénéisation consiste à considérer l'*Ansatz* à deux échelles suivant pour  $u^\epsilon$ <sup>4</sup> :

$$u^\epsilon(\mathbf{x}) = u^0\left(\mathbf{x}, \frac{\mathbf{x}}{\epsilon}\right) + \epsilon u^1\left(\mathbf{x}, \frac{\mathbf{x}}{\epsilon}\right) + \epsilon^2 u^2\left(\mathbf{x}, \frac{\mathbf{x}}{\epsilon}\right) + \dots, \quad (1.8)$$

où chaque fonction  $u^k(\mathbf{x}, \mathbf{y})$  est périodique selon la deuxième variable  $\mathbf{y}$ . En insérant (1.8) dans (1.6), cela permet d'obtenir une hiérarchie d'équations qui permet de déterminer *in fine* les fonctions  $u^k$ . Il est alors ensuite nécessaire de démontrer rigoureusement ces résultats formels. Par exemple, afin de justifier le premier terme de l'*Ansatz* (1.8), il est nécessaire de démontrer des convergences du type :

$$\left\| u^\epsilon(\mathbf{x}) - u^0\left(\mathbf{x}, \frac{\mathbf{x}}{\epsilon}\right) \right\| \rightarrow 0 \quad \text{ou} \quad u^\epsilon(\mathbf{x}) \xrightarrow{\epsilon \rightarrow 0} u^0\left(\mathbf{x}, \frac{\mathbf{x}}{\epsilon}\right),$$

où la norme  $\|\cdot\|$  est à préciser, et on procède de même pour les ordres suivants. Afin de démontrer de tels résultats, diverses méthodes ont vu le jour selon les problèmes considérés, comme la méthode d'énergie de Tartar basée sur des fonctions oscillantes [125] ou celle de Lions, Bensoussan et Papanicolaou basée sur le principe du maximum [13]. Notons que dans le cadre des techniques d'homogénéisation, a été également introduite la notion de convergence à deux échelles [3, 114] qui est un autre outil permettant de justifier rigoureusement l'*Ansatz* (1.8). Cette dernière est une convergence intermédiaire entre les convergences forte et faible dans le sens où la convergence forte implique la convergence à deux échelles, et cette dernière implique la convergence faible. Nous donnons au chapitre 2 un résultat basé sur la convergence à deux échelles.

4. Il s'agit d'une considération pour le moment formelle : rien ne garantit l'existence d'un tel *Ansatz* en l'état.

Nous terminons sur les méthodes d'homogénéisation en signalant que ces techniques ont été appliquées à la physique des plasmas et notamment sur l'équation de Vlasov ou sur le système de Vlasov-Poisson dans la limite du champ magnétique intense. Nous citons à ce sujet les travaux de Frénod et Sonnendrücker [59, 60].

### *Méthode de décomposition de domaine.*

Les méthodes de décomposition de domaine parfois appelées méthodes de Schwarz permettent de décomposer un domaine  $\Omega$  en un nombre fini de sous-domaines  $\Omega^i$  ( $i \in \mathbb{N}^*$ ) et de résoudre le problème sur ces sous-domaines plutôt que sur le domaine entier. Le processus est itéré jusqu'à obtenir une convergence de la solution. Différents types de méthode de décomposition de domaine existent, comme celles admettant une zone de recouvrement entre les sous-domaines et celles n'en admettant pas. Si ces méthodes au départ ont servi des buts théoriques, elles sont aujourd'hui également utilisées dans le cadre numérique.

En effet, afin de résoudre un problème multi-échelles sur un domaine  $\Omega$  via une méthode de simulation numérique (différences finies, éléments finis ou volumes finis), il est nécessaire de discrétiser le domaine  $\Omega$  afin d'obtenir un maillage de pas noté  $h$ . Si  $\epsilon$  représente une certaine échelle d'observation, il est alors impératif de choisir  $h$  tel que  $h < \epsilon$  si l'on veut conserver la description fine du milieu. Néanmoins, lorsque  $\epsilon \rightarrow 0$ , la condition  $h < \epsilon$  ne peut plus être respectée compte tenu de la limitation de la puissance des machines de calcul actuelles. En utilisant les méthodes de décomposition de domaine sur un problème possédant des zones de singularités, ces méthodes peuvent se révéler efficaces et les sous-domaines qui couvrent ces zones auront des maillages plus fins que les sous-domaines couvrant des zones régulières. Le domaine est vaste et nous renvoyons par exemple à [124] pour plus de détails. Notons enfin que ces méthodes sont propices à l'utilisation du calcul parallèle.

### *Méthodes multi-grilles.*

Ces méthodes sont aussi fortement adaptées lorsque le domaine considéré possède des zones de singularité (par exemple des couches limites), où un maillage fin est nécessaire pour la prise en compte des effets microscopiques. Elles consistent donc à avoir un maillage très fin sur les zones singulières et à laisser les zones régulières avec des maillages grossiers. Afin de déterminer le niveau de raffinement du domaine, deux stratégies sont proposées par les méthodes multi-grilles :

- Dans le cas où la connaissance du problème permet de savoir d'avance quelles seront les zones d'intérêt, il suffit de raffiner ces zones en question.
- Ou alors il faut opter pour une méthode plus dynamique en calculant à chaque itération une erreur locale en chaque point du maillage afin de raffiner les zones où cette erreur est importante.

On pourra se référer à l'ouvrage [70] pour une description détaillée de ces méthodes.

Dans ce manuscrit, nous n'aurons pas recours aux méthodes numériques précitées, nous allons dans la section suivante présenter en détail les méthodes que nous utiliserons.

## **1.2.4 Schémas préservant l'asymptotique**

### *Position du problème.*

Dans cette partie, on présente les méthodes préservant l'asymptotique (*Asymptotic Preserving* en anglais) qui occupent une place primordiale dans ce manuscrit. Dans la suite on désignera par les initiales AP le terme *Asymptotic-Preserving*. Présentons dans un premier temps le contexte où peuvent être utiles ces schémas.

Soit  $\epsilon$  un réel positif qu'on choisira inférieur ou égal à 1, sans perte de généralité. Un problème physique faisant intervenir un tel paramètre sera noté  $P^\epsilon$ . Nous supposons que la solution de ce problème notée  $u^\epsilon$  converge (dans un certain sens) vers une solution  $u^0$  d'un problème limite, noté  $P^0$ . De tels problèmes sont délicats à étudier aussi bien d'un point de vue analytique que numérique. En effet, les problèmes non-limite et limite peuvent être de nature mathématique complètement différente (le problème limite étant souvent plus simple que le problème initial). Et certaines méthodes analytiques et/ou numériques adaptées pour le problème non-limite ne le seront pas plus le problème limite (et réciproquement).

Les schémas préservant l'asymptotique trouvent leur motivation dans le contexte des problèmes multi-échelles évoqués précédemment. Les méthodes numériques classiques peuvent échouer à résoudre correctement de tels problèmes. En guise d'exemple, tout schéma numérique explicite souhaitant résoudre un problème  $P^\epsilon$  fera apparaître une condition de stabilité qui dépendra de  $\epsilon$ , ainsi dans de nombreux cas, à la limite  $\epsilon \rightarrow 0$ , il faudra prendre des pas de temps et/ou d'espace de l'ordre de  $\epsilon$  pour ne pas violer cette condition. Ce qui n'est pas tolérable pour des valeurs très faibles de  $\epsilon$ .

Une première idée afin de contrer ce problème serait de résoudre le problème limite  $P^0$  plutôt que le problème  $P^\epsilon$ . Mais pour cela, il faut que celui-ci soit bien connu, ce qui n'est pas toujours le cas. De plus, dans de nombreux cas physiques, se préoccuper uniquement du problème limite est réducteur. Les exemples de problèmes multi-échelles donnés précédemment ont montré que le paramètre  $\epsilon$  pouvait être relié à un paramètre physique. Cela permet de souligner le fait que le problème limite n'est pas toujours adapté pour décrire le phénomène physique considéré. Si  $\epsilon$  est très faible dans certaines zones du domaine, il peut être de l'ordre 1 dans d'autres. Le paramètre  $\epsilon$  peut donc varier en espace et en temps du fait de l'évolution des propriétés physiques du système.

Une deuxième idée serait de développer un schéma numérique pour le problème  $P^\epsilon$  et un autre pour le problème  $P^0$  (sous couvert de connaître ce problème limite). Mais cette approche pose le problème de la zone de transition entre les deux régimes et de la manière de coupler les deux modèles. De plus, il est nécessaire de connaître précisément la variation spatiale de  $\epsilon$  afin de construire une zone de transition appropriée. L'idée d'un schéma AP est de résoudre sans condition de stabilité qui dépendrait de  $\epsilon$  les problèmes  $P^\epsilon$  et  $P^0$  sans passer par le développement de deux méthodes numériques distinctes.

### ***Propriétés d'un schéma AP.***

Nous allons considérer  $P^{\epsilon, \mathbf{h}}$  une discrétisation du problème  $P^\epsilon$  et  $P^{0, \mathbf{h}}$  une discrétisation du problème  $P^0$ , avec  $\mathbf{h} = (\Delta \mathbf{x}, \Delta t)$ . Un schéma préservant l'asymptotique devra vérifier ces trois conditions :

- Le schéma développé devra être uniformément stable par rapport à  $\epsilon$ .
- Pour tout  $\epsilon > 0$  fixé,  $P^{\epsilon, \mathbf{h}}$  devra être une discrétisation consistante du problème  $P^\epsilon$ .
- Pour tout  $\mathbf{h}$  fixé,  $P^{0, \mathbf{h}}$  devra être une discrétisation consistante du problème  $P^0$ .



Ainsi, ces schémas sont en mesure d'obtenir les bonnes solutions numériques aux échelles microscopique et macroscopique en une même procédure numérique. Nous pouvons résumer notre propos dans le diagramme commutatif présent sur la Figure 1.8 :

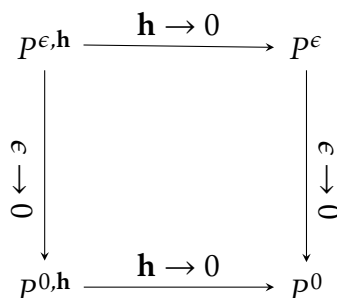


FIGURE 1.8 – Propriétés d'un schéma préservant l'asymptotique.

### *Construction d'un schéma AP.*

Si nous avons expliqué ce qu'est un schéma préservant l'asymptotique, il reste à donner la méthode pour construire un tel schéma à partir de notre problème de départ.

1. **Reformulation du problème à perturbation singulière.** Bien souvent, le problème  $P^\epsilon$  est singulier, ce qui conduit à un problème  $P^0$  de nature mathématique très différente du problème initial. Une des premières étapes du développement d'un schéma AP (bien qu'elle ne soit pas toujours nécessaire) est de reformuler mathématiquement le problème  $P^\epsilon$  afin d'obtenir un nouveau problème  $R^\epsilon$  équivalent au problème  $P^\epsilon$ . L'idée est que ce problème reformulé soit une perturbation régulière du problème limite  $P^0$ . Ils seront donc sous une même forme mathématique et le traitement numérique sera largement simplifié.
2. **Discrétisation en temps adaptée.** L'étape suivante est la discrétisation en temps du problème reformulé (ou du problème initial si ce dernier n'a pas été reformulé). Afin de s'assurer de la stabilité inconditionnelle du schéma, on implémente tous les termes raides du problème, s'il en reste après la régularisation du problème.
3. **Discrétisation en espace adaptée.** Enfin, il faut choisir une discrétisation en espace des opérateurs présents dans le problème reformulé. Cette étape est aussi importante que la discrétisation en temps. En particulier, même si tous les termes raides sont implicites, un mauvais choix de discrétisation spatiale peut conduire à un schéma numérique ne respectant pas les propriétés d'un schéma AP. Ce point sera illustré dans le Chapitre 2.

### *Quelques références pour finir.*

L'esprit des schémas AP a été introduit la première fois par Larsen & al. au sujet de l'obtention de la solution stable lors du transport de neutrons en régime diffusif [91]. Pour une revue détaillée des méthodes préservant l'asymptotique, on renvoie à [80, 112].

De nombreux schémas AP ont été développés pour résoudre une grande variété de problèmes, incluant des problèmes anisotropes elliptiques [35, 36] ou paraboliques [107], ainsi que des équations de Vlasov dans le régime hydrodynamique [55], le régime de dérive-diffusion [30, 84], ou encore le régime en champ magnétique fort [31, 81]. On peut également citer les problèmes liés au régime faible-Mach de l'équation d'Euler [37, 42].

## 1.3 Présentation des résultats obtenus dans cette thèse

Les paragraphes suivant décrivent en détail le contenu des différentes chapitres de ce manuscrit. Les trois chapitres ont donné lieu à trois publications : deux publiées [51, 53] et une soumise [52].

### 1.3.1 Équation de transport anisotrope avec caractéristiques droites ([51] Chapitre 2)

**Objectif général.**

Dans le chapitre 2, nous étudions tant numériquement qu'analytiquement le comportement du système suivant lorsque  $\epsilon \rightarrow 0$  :

$$(T)_\epsilon^1 \begin{cases} \partial_t f^\epsilon + a \partial_x f^\epsilon + \frac{b}{\epsilon} \partial_y f^\epsilon = 0, & \forall (t, x, y) \in [0, T] \times \Omega, \\ f^\epsilon(0, x, y) = f_{in}(x, y), & \forall (x, y) \in \Omega, \end{cases} \quad (1.9)$$

avec  $a$  et  $b$  deux constantes non nulles et  $\Omega$  un domaine de  $\mathbb{R}^2$  doublement périodique.

**Origine physique du système (1.9) et motivations de son étude.**

Le système (1.9) est issu de l'équation de Vlasov dans le régime giro-cinétique en posant  $\mathbf{E} \equiv 0$  et  $\mathbf{B} = \mathbf{e}_z$  et en étudiant sa dynamique dans le plan perpendiculaire au champ magnétique :

$$\partial_t f + \mathbf{v}_\perp \cdot \nabla_{\mathbf{x}} f + \frac{1}{\epsilon} (\mathbf{v} \times \mathbf{B}) \cdot \nabla_{\mathbf{v}} f = 0,$$

où  $\mathbf{v}_\perp = (v_x, v_y, 0)^T$ . Le paramètre  $\epsilon$  correspond dans ce cadre à la fréquence cyclotronique adimensionnée à une échelle de temps caractéristique. En réécrivant l'équation précédente en coordonnées polaires pour la vitesse et en choisissant une condition initiale indépendante de la coordonnée  $y$ , nous obtenons une équation pour l'inconnue  $F(t, x, y, r, \theta)$  :

$$\partial_t F + r \cos(\theta) \partial_x F - \frac{1}{\epsilon} \partial_\theta F = 0,$$

qui est une simple équation de transport anisotrope en dimension deux. Les principales motivations d'étudier un tel système, pourtant très simple, sont les suivantes :

- Le système (1.9) est certes un modèle jouet de l'équation de Vlasov, néanmoins ce dernier conserve son caractère anisotrope, qui est une difficulté bien présente dans les modèles plus complexes. Ainsi l'élaboration d'un schéma numérique efficace pour résoudre le système (1.9) peut être vu comme une première étape avant de s'attaquer à des modèles plus réalistes (et donc plus complexes).
- Le but étant sur ce modèle de développer plusieurs schémas numériques préservant l'asymptotique, un modèle simple permet d'obtenir des solutions analytiques, et donc d'effectuer une analyse numérique précise des schémas développés dans le but de dégager les avantages et les inconvénients de chacun.

## Convergences faible et à deux-échelles, deux résultats.

Nous donnons ici deux résultats analytiques sur le système (1.9) avant d'expliciter les résultats numériques afin de donner un cadre mathématique rigoureux à notre modèle. Le premier résultat explicite la convergence faible de la solution du modèle  $(T)_\epsilon^1$  vers la solution du modèle limite  $(T)_0^1$ , qui correspond au système suivant :

$$(T)_0^1 \begin{cases} \partial_t f^0 + a \partial_x f^0 = 0, & \forall (t, x) \in [0, T] \times \Omega_x, \\ f^0(0, x) = \bar{f}_{in}(x), & \forall x \in \Omega_x, \end{cases} \quad (1.10)$$

où  $\Omega_x = (0, L_x)$  et  $\bar{f}_{in}$  correspond à la moyenne en  $y$  de la condition initiale  $f_{in}$ .

**Théorème 1.3.1** *Considérons une condition initiale  $f_{in} \in H_{\#}^1(\Omega)$ . Alors l'unique solution de  $(T)_\epsilon^1$  resp.  $(T)_0^1$  satisfait  $f^\epsilon \in W^{1,\infty}(0, T; L_{\#}^2(\Omega)) \cap L^\infty(0, T; H_{\#}^1(\Omega))$  resp.  $f^0 \in W^{1,\infty}(0, T; L_{\#}^2(0, L_x)) \cap L^\infty(0, T; H_{\#}^1(0, L_x))$ . De plus, on a la convergence faible- $\star$  suivante*

$$f^\epsilon \xrightarrow[\epsilon \rightarrow 0]{*} f^0 \quad \text{in } L^\infty(0, T; L_{\#}^2(\Omega)). \quad (1.11)$$

La convergence faible ne permet pas de bien prendre en compte la dynamique fortement oscillante de la solution  $f^\epsilon$  lorsque  $\epsilon \rightarrow 0$ . En revanche, la convergence à deux-échelles permet de mieux rendre compte du caractère oscillatoire du problème (1.9). C'est l'objet de ce second résultat.

**Théorème 1.3.2** *Considérons une condition initiale  $f_{in} \in H_{\#}^1(\Omega)$  et  $f^\epsilon \in W^{1,\infty}(0, T; L_{\#}^2(\Omega)) \cap L^\infty(0, T; H_{\#}^1(\Omega))$  l'unique solution de  $(T)_\epsilon^1$ . Alors il existe  $F \in L^\infty(0, T; L_{\#}^\infty(0, L_\tau; L_{\#}^2(\Omega)))$ ,  $L_\tau := L_y/b$ -périodique pour la deuxième variable  $\tau$ , tel que*

$$f^\epsilon \xrightarrow[\epsilon \rightarrow 0]{2-s} F, \quad (1.12)$$

ce qui signifie que pour toute fonction  $\psi \in L^1(0, T; C_{\#}^1(0, L_\tau; L_{\#}^2(\Omega)))$ , nous avons

$$\int_0^T \int_{\Omega} f^\epsilon(t, x, y) \psi(t, t/\epsilon, x, y) dx dy dt \xrightarrow[\epsilon \rightarrow 0]{} \frac{1}{L_\tau} \int_0^T \int_{\Omega} \int_0^{L_\tau} F(t, \tau, x, y) \psi(t, \tau, x, y) d\tau dx dy dt.$$

De plus, nous avons la relation suivante, reliant  $F$  à la solution faible  $f^0$  de  $(T)_0^1$

$$f^0(t, x) = \frac{1}{L_\tau} \int_0^{L_\tau} F(t, \tau, x, y) d\tau.$$

Après avoir posé ce cadre mathématique, nous pouvons à présent donner les principaux résultats numériques obtenus pour ce modèle simplifié.

## Schémas différents mais résultats numériques identiques.

La simplicité du modèle (1.9) permet de développer une large classe de schémas pour sa résolution numérique. Néanmoins nous allons nous concentrer dans cette introduction sur deux d'entre eux<sup>5</sup> qui sont totalement différents dans leur approche de départ. L'idée initiale a été également de montrer les limites d'un schéma de type IMEX sur un maillage cartésien, en particulier lorsque les lignes de champ ne sont pas alignées avec le maillage.

5. Deux autres schémas ont été implémentés, nous renvoyons au Chapitre 2 pour plus de détails à leur sujet.

***Le schéma (IMEX), une discrétisation directe du modèle (1.9).***

Ce schéma est obtenu simplement en implicitant le terme raide dans (1.9) et laissant explicite l'autre terme de dérivée en espace. En particulier, ce schéma ne nécessite aucunement de reformuler le système (1.9). Les dérivées en espace sont discrétisées à l'aide d'un schéma upwind d'ordre 1.

***Le schéma Lagrange-Multiplicateur, une reformulation du modèle avant discrétisation.***

Pour élaborer ce schéma, nous effectuons la décomposition suivante pour  $f^\epsilon$  :

$$f^\epsilon = p^\epsilon + \epsilon q^\epsilon, \quad (1.13)$$

où  $p^\epsilon$  est choisi dans le noyau de l'opérateur dominant  $\frac{b}{\epsilon} \partial_y(\cdot)$ .

Pour assurer le caractère unique de la décomposition (1.13), nous fixons  $q^\epsilon$  sur le flot entrant  $\Gamma_{in}$  en le choisissant dans l'espace :

$$\mathcal{Q} := \{q^\epsilon \in \mathcal{V}, q^\epsilon|_{\Gamma_{in}} = 0\}, \quad \mathcal{V} := W^{1,\infty}(0, T; L^2_\#(\Omega)) \cap L^\infty(0, T; H^1_\#(\Omega)).$$

Et finalement, en injectant (1.13) dans (1.9), nous obtenons un système plus régulier :

$$(La)_\epsilon \begin{cases} \partial_t f^\epsilon + a \partial_x f^\epsilon + b \partial_y q^\epsilon = 0, & \forall (t, x, y) \in [0, T] \times \Omega \\ \partial_y f^\epsilon = \epsilon \partial_y q^\epsilon, & \forall (t, x, y) \in [0, T] \times \Omega \\ q^\epsilon|_{\Gamma_{in}} = 0. \end{cases} \quad (1.14)$$

Les discrétisations des dérivées spatiales de (1.14) sont également assurées ici à l'aide d'un schéma upwind d'ordre 1 et seul le terme  $a \partial_x f^\epsilon$  est laissé explicite dans (1.14).

***Constatations numériques.***

Mais, bien que les deux schémas présentés à l'instant soient très différents, les résultats numériques ne se distinguent que très peu dans ce premier modèle (1.9), il faudra attendre le modèle suivant (voir la sous-section suivante) pour laisser apparaître de réelles différences entre les deux approches.

En effet, considérant la condition initiale :

$$f_{in}(x, y) = \sin(x) (\cos(2y) + 1),$$

les deux schémas numériques conduisent à la limite  $\epsilon \rightarrow 0$  à la bonne solution (faible) du problème limite qui n'admet plus aucune variation selon  $y$ , comme attendu. Tout laisse à penser que le schéma *Lagrange-Multiplicateur* est inutilement lourd (puisque possédant un doublet d'inconnues contre une inconnue pour le schéma (IMEX)). La seule différence notable à signaler est que le conditionnement du système linéaire est indépendant de  $\epsilon$  pour le schéma *Lagrange-Multiplicateur* alors qu'il est de l'ordre  $1/\epsilon$  pour le schéma (IMEX).

### 1.3.2 Équation de transport anisotrope avec caractéristiques circulaires ([51, 53] Chapitre 2 & 3)

**Objectif principal.**

Ici, notre intérêt est porté sur l'étude de la limite  $\epsilon \rightarrow 0$  du système suivant :

$$(T)_\epsilon^2 \begin{cases} \partial_t f^\epsilon + \frac{y}{\epsilon} \partial_x f^\epsilon - \frac{x}{\epsilon} \partial_y f^\epsilon = 0, & \forall (t, x, y) \in [0, T] \times \Omega, \\ f^\epsilon(0, x, y) = f_{in}(x, y), & \forall (x, y) \in \Omega. \end{cases} \quad (1.15)$$

### Origine physique du système (1.15) et motivations de son étude.

Afin d'aboutir au système (1.15), nous partons toujours de l'équation de Vlasov dans le régime gyro-cinétique avec un champ magnétique intense. En choisissant de nouveau  $\mathbf{E} \equiv 0$  et  $\mathbf{B} = \mathbf{e}_z$  et en étudiant la dynamique perpendiculaire au champ  $\mathbf{B}$  avec cette fois une condition initiale indépendante des variables de position, nous aboutissons en coordonnées cartésiennes à l'équation :

$$\partial_t f + \frac{v_y}{\epsilon} \partial_{v_x} f - \frac{v_x}{\epsilon} \partial_{v_y} f = 0.$$

Les caractéristiques de cette équation de transport sont cette fois-ci circulaires dans l'espace des vitesses. Numériquement, ce fait, couplé avec la présence d'un maillage cartésien entrainera des difficultés nouvelles par rapport au premier modèle étudié (1.9).

### Retour sur les schémas numériques développés.

Les différences notables entre les modèles (1.9) et (1.15) nous imposent quelques modifications sur les schémas (IMEX) et *Lagrange-Multiplier*, même si leurs natures profondes resteront inchangées.

#### *Le schéma (IMEX) devient totalement implicite.*

Une nouvelle fois, la discrétisation s'applique directement sur le modèle (1.15), qui n'est donc pas reformulé. Néanmoins, les deux termes de dérivées spatiales étant raides, ils seront tous les deux implicites. En particulier, il n'y a plus de terme explicite, et dans la suite ce schéma sera renommé schéma (IMP), les dérivées spatiales sont discrétisées via un schéma upwind d'ordre 1.

#### *Le schéma Lagrange-Multiplier doit fixer différemment $q^\epsilon$ .*

Si la décomposition (1.13) demeure identique<sup>6</sup> il est néanmoins bien plus délicat de fixer  $q^\epsilon$  sur le flot entrant lorsque les caractéristiques sont circulaires. Nous allons donc fixer  $q^\epsilon$  en insérant un terme de régularisation de la forme  $\sigma q^\epsilon$  dans le système reformulé :

$$(La)_2^\epsilon \begin{cases} \partial_t f^\epsilon + y \partial_x q^\epsilon - x \partial_y q^\epsilon = 0, & \forall (t, x, y) \in [0, T] \times \Omega, \\ y \partial_x f^\epsilon - x \partial_y f^\epsilon - \epsilon (y \partial_x q^\epsilon - x \partial_y q^\epsilon) + \sigma q^\epsilon = 0, & \forall (t, x, y) \in [0, T] \times \Omega, \end{cases} \quad (1.16)$$

où  $\sigma > 0$  est un petit paramètre, le choix de ce dernier sera discuté dans la suite. Pour la discrétisation de (1.16), aucun terme n'est laissé explicite et les dérivées spatiales sont de nouveau discrétisées via un schéma upwind.

---

6. à ceci près que dans ce cas, l'opérateur dominant est  $y \partial_x(\cdot) - x \partial_y(\cdot)$  ;  $p^\epsilon$  étant choisi dans le noyau de ce dernier.

### Des résultats numériques bien différents.

Pour nos simulations avec le modèle (1.15), la condition initiale retenue sera une gaussienne centrée sur le domaine, de sorte que la solution attendue soit stationnaire.

Alors que les schémas (IMEX) et *Lagrange-Multiplier* n'avaient pas marqué leur différence sur le premier modèle (1.9), ils vont ici fortement se distinguer. En effet, alors que le schéma *Lagrange-Multiplier* entraînera une légère diffusion de la solution, due à la discrétisation (qui est donc contrôlable en diminuant les pas de temps et d'espace), le schéma (IMP) entraîne à la limite  $\epsilon \rightarrow 0$  une diffusion très importante, dans toutes les directions, conduisant au final à une solution constante non consistante avec celle attendue analytiquement. Ces deux constats sont visibles sur la Figure 1.9 ci-dessous.

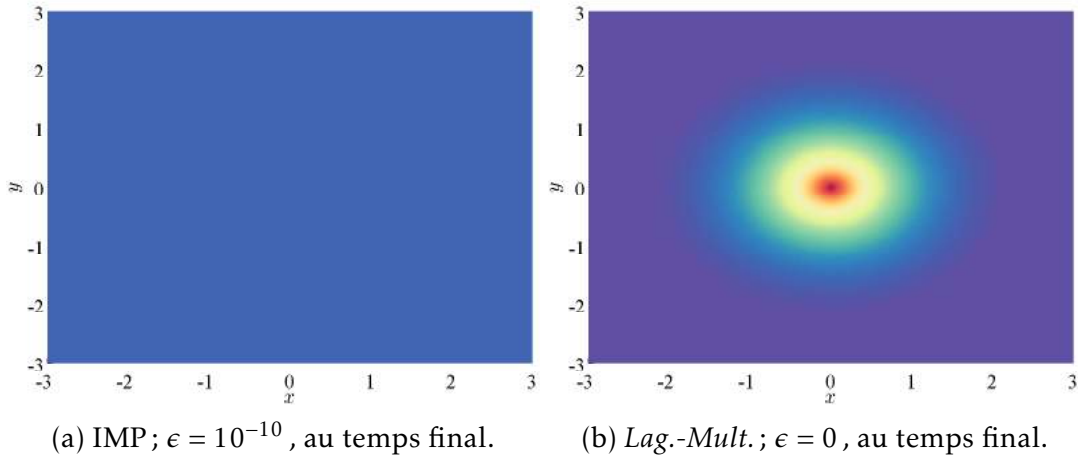


FIGURE 1.9 – Si le schéma *Lagrange-Multiplier* réussit à conserver la stationnarité de la solution à la limite  $\epsilon \rightarrow 0$ , le schéma (IMP) n'en est pas capable.

### Des résultats confirmés par l'analyse numérique.

Ces résultats numériques trouvent une confirmation dans l'analyse numérique des deux schémas (IMP) et *Lagrange-Multiplier*. En effet l'étude des erreurs de troncature vient montrer que dans le cas du schéma (IMP), le terme en premier ordre des pas de temps et d'espace est multiplié par  $1/\epsilon$ , ce qui n'est pas le cas pour le schéma *Lagrange-Multiplier*. L'erreur de troncature de ce dernier fait en effet apparaître une dépendance en  $1/\epsilon$  uniquement dans la direction des lignes de champs, tandis que dans la direction perpendiculaire à ces lignes, la dépendance en  $1/\epsilon$  disparaît.

### Un résultat mathématique du système reformulé.

Si des premiers résultats sur le schéma *Lagrange-Multiplier* ont été explicités, aucune analyse mathématique rigoureuse du système de transport reformulé après la décomposition Micro-Macro (1.13) n'a été effectuée. C'est l'objet de la suite de ce paragraphe.

Nous allons pour cela considérer une équation de transport anisotrope plus générale que dans les modèles précédents :

$$\partial_t f^\epsilon + \frac{1}{\epsilon} \mathbf{b} \cdot \nabla_{\mathbf{x}} f^\epsilon = 0, \quad \forall (t, \mathbf{x}) \in (0, T) \times \Omega, \quad \text{avec } \Omega \subset \mathbb{R}^2,$$

où le champ  $\mathbf{b}$  sera considéré de divergence nulle et indépendant du temps.

Considérons à présent le système semi-discrétisé en temps après avoir effectué la décomposition Micro-Macro (1.13) de l'équation précédente<sup>7</sup>, couplée avec la technique de régularisation illustrée dans le modèle précédent (1.15) :

$$(MM)^{\epsilon, \sigma, n} \begin{cases} f^{\epsilon, \sigma, n+1} + \Delta t \mathbf{b} \cdot \nabla q^{\epsilon, \sigma, n+1} = f^{\epsilon, \sigma, n}, \\ \mathbf{b} \cdot \nabla f^{\epsilon, \sigma, n+1} = \epsilon \mathbf{b} \cdot \nabla q^{\epsilon, \sigma, n+1} - \sigma q^{\epsilon, \sigma, n+1}, \end{cases} \quad (1.17)$$

avec  $n \in \mathbb{N}^*$ .

Afin de contextualiser notre futur résultat, nous nous plaçons dans le cadre suivant :

**Hypothèses 1** *Le domaine  $\Omega$  sera soit  $\mathbb{R}^2$  ou une bande infinie  $(L_1, L_2) \times \mathbb{R}$ . Dans le deuxième cas, on suppose en plus des conditions aux bords périodiques en  $x$  et le champ  $\mathbf{b}$  est supposé également périodique en  $x$ .*

Nous introduisons de plus les espaces fonctionnels suivant :

$$V := L^2(\Omega),$$

$$Q := \{v \in L^2(\Omega) \mid \mathbf{b} \cdot \nabla v \in L^2(\Omega), v \text{ satisfaisant les Hypothèses 1}\}.$$

Dans le cadre de ces hypothèses, on démontre alors le théorème suivant :

**Théorème 1.3.3** *Pour tout  $f^{\epsilon, \sigma, n} \in V^*$  le problème (1.17) admet pour tout  $\epsilon \geq 0$ ,  $\sigma > 0$ ,  $\Delta t > 0$  et  $n \in [0, N_t]$  une unique solution faible  $(f^{\epsilon, \sigma, n+1}, q^{\epsilon, \sigma, n+1})$  dans  $V \times Q$ .*

*Ce doublet de solutions vérifie les estimations suivantes, pour tout  $\epsilon \geq 0$ ,  $\sigma > 0$ ,  $\Delta t > 0$ , et  $n \in [0, N_t]$  :*

$$\|f^{\epsilon, \sigma, n+1}\|_V \leq \|f^{\epsilon, \sigma, n}\|_{V^*}, \quad (1.18)$$

$$\|q^{\epsilon, \sigma, n+1}\|_V^2 \leq \frac{1}{\sigma \Delta t} \|f^{\epsilon, \sigma, n}\|_{V^*}^2, \quad \|\mathbf{b} \cdot \nabla q^{\epsilon, \sigma, n+1}\|_V \leq \frac{2}{\Delta t} \|f^{\epsilon, \sigma, n}\|_{V^*}. \quad (1.19)$$

**Idée de la preuve :** La preuve de ce théorème repose sur la transformation du problème (1.17) en un problème variationnel. Néanmoins, un défaut de coercivité d'un des opérateurs de ce dernier nous contraindra à introduire l'opérateur régularisé associé, qui lui sera coercif. Le théorème de Lax-Milgram pourra ainsi être appliqué. Des manipulations judicieuses des équations permettront alors d'établir les estimations introduites dans l'énoncé du théorème.

Ce résultat permet d'enrichir le cadre mathématique de notre reformulation AP. L'idée à présent est d'améliorer significativement la discrétisation en temps et en espace de notre schéma, en vue d'obtenir des résultats numériques non plus sur des modèles jouets mais sur des modèles plus physiques et donc plus complexes.

7. ici,  $p^\epsilon$  est choisi dans le noyau de l'opérateur  $\mathbf{b} \cdot \nabla$ .

## La naissance du (DAMM)-scheme.

Les discrétisations opérées jusqu'à présent étaient toujours d'ordre 1 en temps et en espace à l'aide respectivement de schémas d'Euler (explicite ou implicite selon les cas) et du schéma upwind. Afin d'obtenir une procédure numérique plus performante, nous proposons à présent les discrétisations suivantes :

- L'opérateur  $\mathbf{b} \cdot \nabla$  sera discrétisé via le schéma d'Arakawa, qui historiquement a été conçu pour ce type d'opérateur, que l'on peut réécrire sous forme d'un crochet de Poisson. Il permet d'obtenir de très bonnes propriétés de conservations, notamment des grandeurs énergétiques. Nous renvoyons aux travaux originels d'Arakawa à ce sujet pour plus de détails sur ce schéma [6].
- La discrétisation temporelle sera elle assurée par une méthode DIRK (*Diagonal Implicit Runge-Kutta*) [2].

Lorsque ces discrétisations seront précédées de la décomposition Micro-Macro (1.13), alors cette procédure numérique sera référencée par le nom de (DAMM)-scheme (pour DIRK Arakawa Micro-Macro).

## Le choix du paramètre $\sigma$ .

Comme nous l'avons vu, il est nécessaire de fixer  $q^\epsilon$  afin d'obtenir une décomposition (1.13) unique et donc des systèmes reformulés bien posés. Pour cela, nous avons fait le choix d'une régularisation de type Tikhonov en ajoutant un terme de la forme  $\sigma q^\epsilon$  dans le système, ce qui entraîne bien sûr la perte de l'équivalence du nouveau système avec le système initial.

Se pose naturellement la question du choix du paramètre  $\sigma$ , la seule contrainte évidente étant de le choisir strictement positif, afin d'assurer la stabilité du schéma. Si le paramètre  $\sigma$  est choisi trop faible, le temps de calcul numérique en sera très défavorablement impacté, le conditionnement du système linéaire étant de l'ordre de  $\sigma^{-1}$ . Au contraire, le choisir trop élevé nous éloigne du modèle initial que l'on souhaite résoudre. Il est donc nécessaire de trouver un bon compromis.

Il est assez naturel de poser  $\sigma$  comme une puissance du pas d'espace  $h$  en visualisant le système (1.17). Considérant un champ  $\mathbf{b}$  de la forme  $\mathbf{b} := (y, -x)$ , il est alors possible de comparer la solution numérique avec la solution analytique. En faisant varier  $\sigma$ , nous pouvons alors évaluer son influence sur l'erreur numérique calculée. À l'aide de courbes d'erreur, nous montrons alors qu'un bon compromis est de choisir  $\sigma = h^2$  pour  $\epsilon$  grand et  $\sigma = h$  pour  $\epsilon \rightarrow 0$ .

## Conclusion de cette partie.

Nous avons ainsi explicité le (DAMM)-scheme et expliqué dans quelle mesure choisir le paramètre  $\sigma$  afin d'avoir un système bien posé (et numériquement un temps de calcul raisonnable) sans introduire une erreur trop conséquente dans le système reformulé. Nous avons aussi montré que dans le cas où les lignes de champ ne sont plus alignées avec les axes, un schéma de type IMEX n'est plus adapté. Il est temps à présent de mettre à profit ses propriétés AP sur des problématiques physiques, plus complexes. C'est l'objet des deux sous-sections suivantes.



### 1.3.3 Contributions numériques du (DAMM)-*scheme* sur le système de Vlasov-Poisson ( [53] Chapitre 3).

#### Contexte et objectifs.

Nous avons déjà introduit le système de Vlasov-Poisson dans la première partie de cette introduction, nous allons à présent considérer le système de Vlasov-Poisson en temps long obtenu à la limite  $\epsilon \rightarrow 0$  à partir du système suivant :

$$(VP)^\epsilon \begin{cases} \partial_t f^\epsilon + \frac{1}{\epsilon} \{f^\epsilon, \Psi^\epsilon\} = 0, & \forall (t, \mathbf{x}) \in [0, T] \times \Omega. \\ -\partial_{xx} \varphi^\epsilon(t, x) = 1 - \int_{\mathbb{R}} f^\epsilon(t, x, v) dv, & \forall (t, x) \in [0, T] \times \Omega_x. \end{cases} \quad (1.20)$$

où  $\Psi^\epsilon(t, x, v) = v^2/2 - \varphi^\epsilon(t, x, v)$  et  $\{\cdot, \cdot\}$  désigne le crochet de Poisson, défini par :  $\{\theta, \xi\} = \partial_x \theta \partial_v \xi - \partial_v \theta \partial_x \xi$ . Les motivations de l'étude du système (1.20) sont les suivantes :

- Après avoir validé le (DAMM)-*scheme* sur un modèle jouet linéaire, l'objectif est maintenant de le valider dans un cadre physique, sur des problèmes non-linéaires<sup>8</sup>. C'est pourquoi, dans un premier temps, nous fixons  $\epsilon = 1$  dans (1.20), et nous testons notre schéma sur quelques cas tests classiques du système de Vlasov-Poisson.
- Une fois cette nouvelle validation effectuée, nous allons utiliser les propriétés AP de notre schéma pour traiter le régime  $\epsilon \rightarrow 0$  de (1.20). Cela permettra d'étudier les ondes BGK avec un schéma uniformément stable en  $\epsilon$ .

#### Résultats numériques classiques retrouvés par le (DAMM)-*scheme*.

Afin d'amorcer les résultats plus significatifs concernant les ondes BGK, nous validons d'abord le (DAMM)-*scheme* à l'aide des résultats classiques associés au système de Vlasov-Poisson pour  $\epsilon = 1$ . Cependant, contrairement aux modèles jouets, la validation numérique se fait de manière indirecte : il n'est pas ici question de comparer notre solution numérique à une solution analytique (non connue) mais de retrouver numériquement quelques résultats analytiques connus comme les taux d'évolution de l'énergie de l'amortissement Landau et de l'instabilité bi-courant. Afin de simuler ces deux phénomènes, nous considérons les conditions initiales suivantes :

$$f_{in}^A(x, y) = \frac{1}{\sqrt{2\pi}} (1 + \gamma \cos(kx)) e^{-v^2/2}, \quad k = 0.5, \quad \gamma = 0.001, \quad \text{Amortissement Landau,}$$

$$f_{in}^B(x, y) = \frac{1}{2\sqrt{2\pi}} (1 + \gamma \cos(kx)) (e^{-(v-3)^2/2} + e^{-(v+3)^2/2}), \quad k = 0.2, \quad \gamma = 0.001, \quad \text{Instabilité bi-courant.}$$

En plus de simulations effectuées avec le (DAMM)-*scheme*, nous avons développé un schéma spectral afin de conforter nos résultats. Sur la Figure 1.10, nous montrons que nous retrouvons les taux d'évolution souhaités pour l'amortissement Landau (a) et l'instabilité bi-courant (b). Néanmoins, les propriétés AP de notre schéma n'ont pas été mises à contribution ici (car  $\epsilon = 1$ ), et nous remarquons sur la même Figure 1.10 que le schéma spectral, bien plus standard dans son implémentation, obtient des résultats quasi-identiques que le (DAMM)-*scheme*. Mais les résultats suivants vont venir mettre en évidence les avantages du (DAMM)-*scheme*, avantages liés aux propriétés AP du schéma.

8. Afin de traiter cette non-linéarité, nous implémentons dans le (DAMM)-*scheme* une méthode de point fixe pour résoudre le système linéaire à chaque itération en temps.

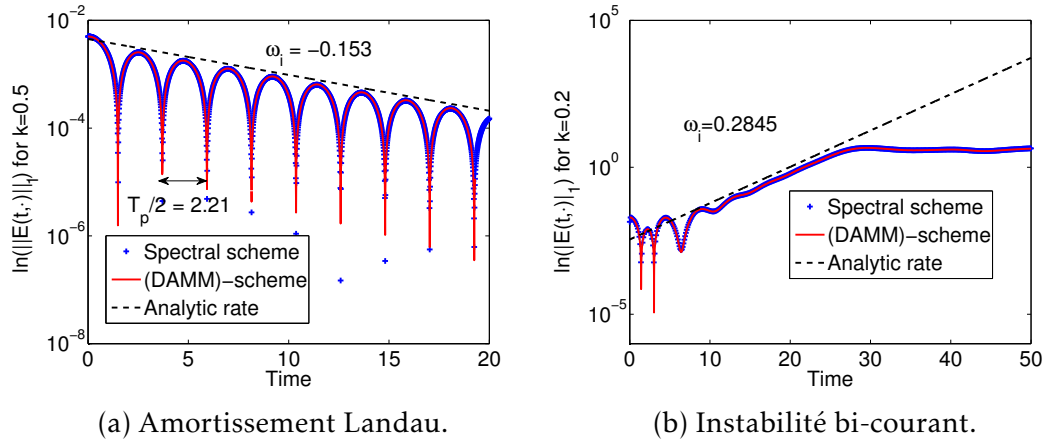


FIGURE 1.10 – Les résultats classiques du système de Vlasov-Poisson pour  $\epsilon = 1$  sont retrouvés par le *(DAMM)-scheme*.

### Propriétés AP du *(DAMM)-scheme* et ondes BGK.

Nous rappelons que les ondes BGK constituent une solution stationnaire mais non-homogène du système de Vlasov-Poisson. Elles peuvent être obtenues comme solutions limites en temps long de l'instabilité bi-courant. Or, considérer la limite  $\epsilon \rightarrow 0$  dans (1.20) revient à considérer  $t \rightarrow \infty$ , les propriétés AP du *(DAMM)-scheme* permettront donc d'obtenir la solution stationnaire directement lorsque  $\epsilon \rightarrow 0$  sans passer par les états intermédiaires du système, comme le ferait un schéma standard.

C'est là que se situe la grande force de notre schéma : si nous sommes intéressés uniquement par l'état final d'un système, et non par les états intermédiaires, alors le *(DAMM)-scheme* permet de l'obtenir avec un temps de calcul très raisonnable, qui aurait été beaucoup plus conséquent avec une procédure numérique classique, et cela dû aux conditions restrictives dépendantes de  $\epsilon$  de stabilité d'un schéma standard.

Afin d'obtenir les ondes BGK, la condition initiale suivante est utilisée :

$$f_{in}^C(x, y) = \frac{1}{\sqrt{2\pi}}(1 + \gamma \cos(kx))v^2 e^{-v^2/2}, \quad k = 0.5, \quad \gamma = 0.05.$$

Les résultats de la simulation numérique via le *(DAMM)-scheme* en prenant  $\epsilon = 0$ , et après quelques itérations en temps (phase de transition), sont visibles sur la Figure 1.11. On distingue la fonction de distribution sur le panel (a) et le diagramme  $f^0 - \Psi^0$  sur le panel (b), démontrant la relation fonctionnelle entre les deux quantités à la limite  $\epsilon \rightarrow 0$ . Afin de valider notre résultat nous avons rajouté sur ce panel (b) la courbe issue d'un fitting proposée dans [74].

### Conclusion de cette partie.

Nous avons mis en exergue à travers le système de Vlasov-Poisson les avantages du *(DAMM)-scheme*, à savoir ses propriétés AP. Ce premier modèle physique a permis de dépasser le cadre des modèles jouets du chapitre 2 comme cadre d'application de notre schéma. Dans la dernière sous-section, nous quitterons le champ de la physique des plasmas pour celui de la mécanique des fluides, pour donner un dernier exemple physique d'application de notre schéma.

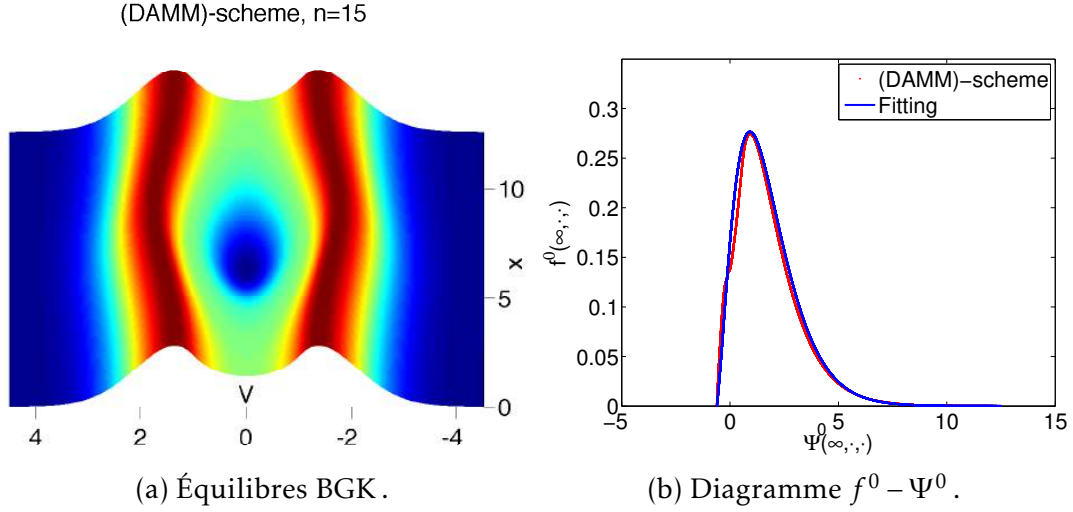


FIGURE 1.11 – Les équilibres BGK sont obtenus au bout de quelques itérations en temps par le (DAMM)-scheme, en prenant  $\epsilon = 0$ .

### 1.3.4 Étude d’une équation de Vorticité-Poisson, écoulement de Kolmogorov ( [52] Chapitre 4).

#### Objectif principaux et motivations.

L’objectif principal de cet ultime chapitre est l’étude du système de Vorticité-Poisson suivant :

$$(\mathcal{V})^\epsilon \begin{cases} \partial_t \omega^\epsilon + \frac{1}{\epsilon} \mathbf{u}^\epsilon \cdot \nabla \omega^\epsilon = (\Delta \omega^\epsilon - \Delta \omega_{eq}), & (t, \mathbf{x}) \in \mathbb{R}^+ \times \Omega, \\ -\Delta \Psi^\epsilon = \omega^\epsilon, & \mathbf{u}^\epsilon = {}^\perp \nabla \Psi^\epsilon, \quad (t, \mathbf{x}) \in \mathbb{R}^+ \times \Omega. \end{cases} \quad (1.21)$$

Une nouvelle fois, on travaille dans  $\mathbb{R}^2$ . Si dans un premier temps, nous oublions le second membre de la première équation de (1.21), alors on retrouve le système d’Euler incompressible, qui partage de fortes similitudes avec le système de Vlasov-Poisson, comme l’ont déjà signalé Majda et Bertozzi dans [102]. La présence d’un terme de dissipation  $\Delta \omega^\epsilon$  ainsi qu’un terme de forçage  $\Delta \omega_{eq}$  supposé indépendant du temps, viennent complexifier le modèle. Les buts de ce chapitre sont les suivants :

- Étudier numériquement la limite  $\epsilon \rightarrow 0$  de l’écoulement de Kolmogorov obtenu en considérant un forçage sinusoïdal et un domaine périodique possédant un rapport d’aspect supérieur à 1.
- Nous verrons que dans ces conditions, une instabilité apparait, et une analyse analytique de la phase linéaire de cette dernière sera effectuée. Nous apporterons ainsi une contribution à la compréhension cette instabilité en reliant son taux de croissance avec le rapport d’aspect du domaine.

Dans le contexte de la physique des plasmas, ce type de modèles peuvent également intervenir. En effet le système (1.21) est dérivé d’un système proposé par Hasegawa et Mima [17]. Ce dernier décrit l’évolution du potentiel électrique dans un plan perpendiculaire au champ magnétique et est un peu plus général que le modèle que nous étudions ici.

## L'écoulement de Kolmogorov.

En considérant un domaine  $\Omega$  de la forme  $\Omega = (0, 2\pi) \times (0, 2\pi/K)$  avec  $K > 0$  et un forçage  $\omega_{eq} = \cos(x)$  associé au système (1.21), nous étudions l'écoulement de Kolmogorov qui possède une large littérature [7, 8, 27, 78, 99, 108, 115]. En particulier, certaines de ces publications, notamment [78, 108] montrent que lorsque  $K < 1$ , un écoulement instable peut se produire pourvu que la viscosité soit suffisamment faible. Nous nous plaçons désormais dans ce cadre, afin d'apporter quelques contributions à l'étude de cette instabilité.

## Phase linéaire de l'instabilité de Kolmogorov.

Afin d'étudier cette phase linéaire, nous linéarisons le système (1.21) autour de l'équilibre  $\omega_{eq}$  afin d'obtenir l'équation suivante, qui est un cas particulier (considérant un champ de vitesse sinusoïdal) de l'équation d'Orr-Sommerfeld, de la seule variable  $\widetilde{\Psi}^\epsilon := \Psi^\epsilon - \Psi_{eq}$  :

$$-i\widehat{\gamma}(\widetilde{\Psi}''(x) - (1 - \eta)\widetilde{\Psi}(x)) + \sin(x)(\widetilde{\Psi}''(x) + \eta\widetilde{\Psi}(x)) = \epsilon(\widetilde{\Psi}^{(4)}(x) - 2K^2\widetilde{\Psi}''(x) + K^4\Psi(x)),$$

où  $\eta := 1 - K^2$  et  $\gamma$  correspond au taux d'instabilité. Si le terme visqueux est négligé et en posant  $\widehat{\gamma} = \gamma/K$ , l'équation précédente se réécrit :

$$-i\widehat{\gamma}(\widetilde{\Psi}''(x) - (1 - \eta)\widetilde{\Psi}(x)) + \sin(x)(\widetilde{\Psi}''(x) + \eta\widetilde{\Psi}(x)) = 0, \quad (1.22)$$

à laquelle on adjoint des conditions aux bords périodiques. Cette dernière équation est un cas particulier de l'équation de Rayleigh. Les équations d'Orr-Sommerfeld et de Rayleigh constituent des problèmes aux valeurs propres généralisés. On renvoie à [46] et aux références associées pour une étude approfondie de ceux-ci. De plus, des résultats numériques (appuyés par un résultat analytique partiel) [11] permettent de supposer que pour une valeur de  $K$  dans  $(0, 1)$ , il n'existe qu'une seule valeur propre  $\widehat{\gamma}$  réelle et strictement positive, conduisant à l'instabilité. On se place désormais dans ce cadre.

### *Lien entre le taux de croissance $\widehat{\gamma}$ et le rapport d'aspect $\eta$ .*

L'équation (1.22) est trop délicate à résoudre en l'état. Néanmoins, l'usage de certaines méthodes peut permettre une résolution approximative. En considérant  $\eta$  et  $\widehat{\gamma}$  comme de petits paramètres nous allons par deux méthodes différentes obtenir le même résultat, qui relie le rapport d'aspect du domaine avec le taux de croissance de l'instabilité linéaire, à savoir :

$$\eta = \widehat{\gamma}. \quad (1.23)$$

Cette égalité n'est vraie qu'au premier ordre en  $\widehat{\gamma}$ . Les deux méthodes qui permettent d'aboutir à ce résultat sont brièvement décrites ci-dessous :

1. On utilise dans un premier temps les méthodes de raccordement asymptotiques qui ont été décrites dans la première partie de cette introduction. En particulier, après avoir calculé les solutions externes et internes, on montre qu'il est nécessaire d'avoir la relation (1.23) afin de raccorder ces deux dernières.
2. L'autre approche consiste à réécrire l'équation (1.22) sous la forme  $\widetilde{\Psi}''(x) = Q(x)\widetilde{\Psi}(x)$  et à considérer la fonction  $\widetilde{\Psi}$  comme somme d'une constante (choisie arbitrairement à 1) et d'une partie fluctuante  $\widetilde{\Phi} \ll 1$  qui est induite par la présence des paramètres  $\eta$  et  $\widehat{\gamma}$ . On est ainsi ramené à résoudre  $\widetilde{\Phi}'' = Q(x)$  au premier ordre en  $\widehat{\gamma}$  et  $\eta$ . De

nouveau, on montrera qu'il sera nécessaire d'avoir la relation (1.23) afin de respecter cette fois la condition de périodicité de la fonction  $\widetilde{\Psi}$ .

Nous validons ensuite ce résultat numériquement en utilisant le *(DAMM)-scheme*. D'une part en comparant le taux d'instabilité linéaire numérique avec le taux d'instabilité théorique pour plusieurs valeurs de rapport d'aspect. D'autre part en comparant la phase de  $\widetilde{\Psi}(\pm\pi/2)$  numérique et théorique, qui peut aussi être relié au rapport d'aspect. En dépit d'un léger écart des valeurs théoriques et numériques pour des raisons soit purement numériques soit analytiques (l'égalité  $\widehat{\gamma} = \eta$  n'étant valide que pour des valeurs faibles de  $\widehat{\gamma}$ ), cette validation est satisfaisante et conforte ce résultat.

### Phase non-linéaire et saturation de l'instabilité de Kolmogorov.

La phase linéaire laisse place après un certain temps à une phase non-linéaire où des phénomènes de filamentation et de mélange des vortex sont visibles. Cette phase est bien plus délicate à étudier analytiquement. Nos premières simulations numériques (où la condition initiale est égale à  $\omega_{eq}$ ) montrent qu'après cette phase non-linéaire, lui succède une phase de saturation où un nouvel équilibre se forme.

#### *Le (DAMM)-scheme pour atteindre directement le nouvel équilibre.*

Naturellement, nous nous posons la question de savoir si le *(DAMM)-scheme* peut atteindre directement ce nouvel équilibre, comme il avait pu le faire pour les équilibres BGK avec le système de Vlasov-Poisson. En effectuant de nouvelles simulations numériques avec  $\epsilon \rightarrow 0$ , nous atteignons le nouvel équilibre dès les premières itérations en temps, sans passer par les états intermédiaires de mélange et de filamentation. Le nouvel équilibre est visible sur la Figure 1.12 (a).

#### *Un fitting pour renforcer nos résultats.*

Comme nous l'avions constaté pour les ondes BGK, un diagramme  $\omega^0 - \Psi^0$  permet de montrer l'existence d'une relation fonctionnelle entre les deux quantités, et nous conjecturons une relation de la forme  $\omega^0 = a \sinh(b\Psi^0)$  (Figure 1.12 (b)). À l'aide de [69], il nous a alors été possible de proposer un *fitting* de la fonction courant, à savoir

$$\Psi_{fit}(x, y) := \frac{4}{b} \tanh^{-1} \left( \frac{\sqrt{k} \operatorname{cn}(rx, k) - \sqrt{k_1} \operatorname{cn}(sy + \varphi, k_1)}{1 + \sqrt{k k_1} \operatorname{cn}(rx, k) \operatorname{cn}(sy + \varphi, k_1)} \right),$$

où  $\operatorname{cn}$  correspond à la fonction cosinus elliptique de Jacobi, et  $(r, s, k, k_1, \varphi)$  étant des paramètres déterminés numériquement. Nous comparons ce *fitting* avec la solution numérique obtenue. Les deux solutions semblent très similaires, confirmant la relation fonctionnelle évoquée plus haut (voir Figure 1.12 (b)).

### 1.3.5 Conclusion générale.

Le *(DAMM)-scheme* constitue le fil rouge de ce manuscrit. La décomposition Micro-Macro, première étape du *(DAMM)-scheme* a été effectuée sur les modèles jouets issus du régime giro-cinétique de l'équation de Vlasov. Cette première étape a permis de montrer les premiers avantages de notre procédure numérique par rapport aux méthodes classiques, de type IMEX ou spectral. Le *(DAMM)-scheme* s'est vu ensuite renforcer, en lui conférant des discrétisations temporelles et spatiales d'ordre 2 pour pouvoir l'appliquer sur des

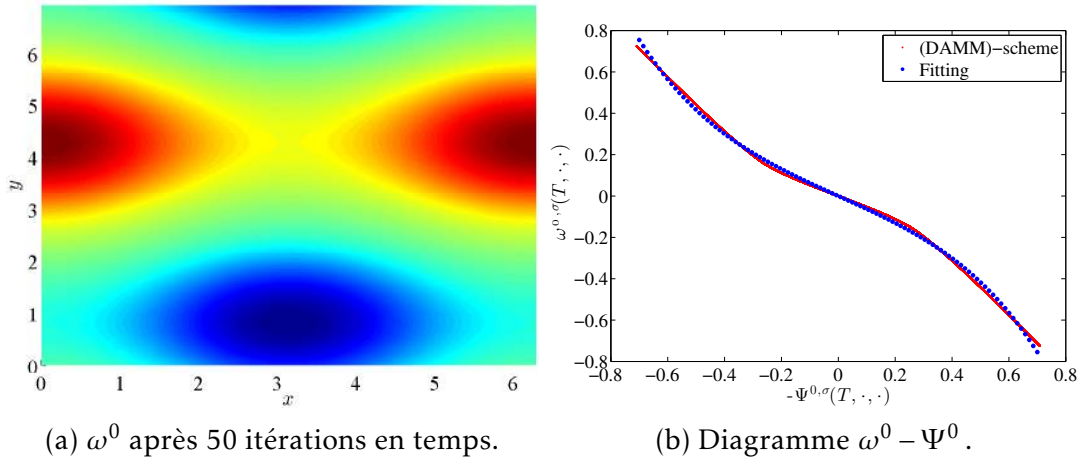


FIGURE 1.12 – Comme pour les ondes BGK, le *(DAMM)-scheme* permet d’atteindre l’état final du système en quelques itérations en temps.

problèmes plus complexes. Ainsi, les modèles de Vlasov-Poisson et de Vorticité-Poisson ont mis en exergue les propriétés *AP* du *(DAMM)-scheme*, sa principale force se révélant lorsqu’un problème physique (non-linéaire) possède un équilibre en temps long, que notre schéma peut atteindre directement, en édulcorant les états intermédiaires du système. Comme les avantages arrivent souvent avec des désavantages, notre schéma en admet aussi, une variable supplémentaire est nécessaire ainsi qu’une procédure de régularisation où le choix d’un paramètre est à discuter. Nous terminons cette introduction avec quelques perspectives de nos travaux.

### 1.3.6 Quelques perspectives de nos travaux

#### Équation de Vlasov à quatre dimensions.

Dans le Chapitre 2 de ce manuscrit nous avons considéré deux modèles simplifiés de l’équation de Vlasov qui permettaient de se ramener dans chaque cas à une équation d’évolution à deux variables (sans compter le temps). Une perspective possible est l’étude du modèle de Vlasov (pour les ions) à quatre dimensions (deux en espace, deux en vitesse, dans le plan perpendiculaire au champ  $\mathbf{B}$ ), dans le régime à fort champ magnétique.

$$\partial_t f^\epsilon + \mathbf{v} \cdot \nabla_{\mathbf{x}} f^\epsilon + \left( \mathbf{E} + \frac{1}{\epsilon} \mathbf{v} \times \mathbf{B} \right) \cdot \nabla_{\mathbf{v}} f^\epsilon = 0. \quad (1.24)$$

Nous pourrions considérer que le champ magnétique est dirigé selon  $\mathbf{e}_z$  :  $\mathbf{B} = (0, 0, b)^T$ , avec  $b$  donné. Le champ électrique  $\mathbf{E}$  sera lui aussi considéré donné. Physiquement,  $\epsilon$  représente la fréquence cyclotronique (adimensionnée avec une fréquence caractéristique d’observation) et sa limite  $\epsilon \rightarrow 0$  correspond à l’asymptotique en champ magnétique intense.

#### *Motivations.*

La motivation évidente de l’étude d’un tel système est qu’il se rapproche de la complexité d’un modèle plus réaliste quant à la dynamique d’un plasma de fusion. Même si ici, une seule espèce est considérée (le problème de couplage des ions et des électrons constituerait l’étape suivante), nous dépassons largement le cadre des modèles jouets explorés

jusqu'ici dans ce cadre. Ayant montré les avantages de la procédure numérique implémentée dans les différents cas tests de ce manuscrit, nous pouvons ainsi espérer de montrer son efficacité sur ce type de modèles, avec un nombre de dimensions plus important.

### **Reformulation de l'équation (1.24).**

En introduisant les deux opérateurs  $\mathcal{L} := \mathbf{v} \cdot \nabla_{\mathbf{x}}$  et  $\mathcal{T} := (\mathbf{v} \times \mathbf{B}) \cdot \nabla_{\mathbf{v}}$ , nous pouvons réécrire l'équation (1.24) sous la forme :

$$\partial_t f^\epsilon + \mathcal{L} f^\epsilon + \mathbf{E} \cdot \nabla_{\mathbf{v}} f^\epsilon + \frac{1}{\epsilon} \mathcal{T} f^\epsilon = 0. \quad (1.25)$$

Afin d'étudier le régime  $\epsilon \rightarrow 0$  numériquement, nous pouvons utiliser la décomposition Micro-Macro habituelle :  $f^\epsilon = p^\epsilon + \epsilon q^\epsilon$  où  $p^\epsilon \in \ker(\mathcal{T})$ . Puis en régularisant le système obtenu avoir injecté la précédente décomposition dans (1.25), nous obtenons :

$$\begin{cases} \partial_t f^\epsilon + \mathcal{L} f^\epsilon + \mathbf{E} \cdot \nabla_{\mathbf{v}} f^\epsilon + \mathcal{T} q^\epsilon = 0, \\ \mathcal{T} f^\epsilon - \epsilon \mathcal{T} q^\epsilon + \sigma f^\epsilon = 0. \end{cases} \quad (1.26)$$

Une discrétisation adaptée en temps et en espace de (1.26) permettrait alors l'obtention d'un schéma préservant l'asymptotique pouvant étudier la limite en champ intense  $\epsilon \rightarrow 0$ . Néanmoins, contrairement à toutes les simulations effectuées dans ce manuscrit (à l'aide de MATLAB), l'utilisation du langage C ou de Fortran se révélerait nécessaire compte tenu du nombre de degré de liberté dans le cas de ce modèle de Vlasov à quatre dimensions.

### **Système de Vlasov-Poisson avec noyau de collisions de type Fokker-Planck.**

Dans le Chapitre 3, nous nous sommes limités à étudier le système de Vlasov-Poisson 1D1V sans considérer de collisions. Naturellement, une des étapes suivantes possible serait de rajouter un noyau de collisions au modèle, par exemple de type Fokker-Planck :

$$\begin{cases} \partial_t f^\epsilon + \frac{1}{\epsilon} v \partial_x f^\epsilon - \frac{1}{\epsilon} E^\epsilon \partial_v f^\epsilon = \frac{1}{\epsilon^\beta} \partial_v [(v - \epsilon u^\epsilon) f^\epsilon + \partial_v f^\epsilon], \\ -\partial_{xx} \varphi^\epsilon(t, x) = 1 - n^\epsilon \quad E^\epsilon = -\partial_x \varphi^\epsilon, \end{cases} \quad (1.27)$$

avec  $n^\epsilon = \int_{\mathbb{R}} f^\epsilon(t, x, v) dv$  et  $\epsilon n^\epsilon u^\epsilon = \int_{\mathbb{R}} v f^\epsilon dv$  et  $\beta \in (0, 1)$ . L'objectif est d'étudier la limite  $\epsilon \rightarrow 0$  de ce système. Cette limite correspond à la fois à la limite en temps long de l'évolution des électrons sous l'influence d'un faible terme de collision. On retrouve à la limite  $\epsilon \rightarrow 0$  la relation de Boltzmann pour les électrons.

### **Motivations.**

Une nouvelle fois, l'objectif de l'étude du système (1.27) est de confronter la nouvelle complexité de ce modèle à notre procédure numérique. Comme on le verra au paragraphe suivant, le système (1.27) possède de nouvelles difficultés qui entraineront certaines modifications dans notre procédure numérique afin de le résoudre, tout en conservant son essence originelle. Le système (1.27) est un premier pas vers le couplage ions/électrons où en plus des collisions intra-espèces, la présence de collisions inter-espèces devra être prise en compte.

### **Reformulation AP et premiers résultats numériques.**

L'implémentation du (*DAMM*)-*scheme* afin de résoudre le système (1.27) est plus délicate que dans le cas du système de Vlasov-Poisson sans collisions. En effet, ici deux opérateurs, un de transport, l'autre de collisions sont présents dans le modèle. Ces deux derniers sont raides, mais pas de la même manière (puisque  $\beta \neq 1$ ). Une manière de s'inspirer du (*DAMM*)-*scheme* tout en s'adaptant à ces nouvelles difficultés est de considérer une double décomposition Micro-Macro telles que :  $f^\epsilon = p^\epsilon + \epsilon q^\epsilon$  et  $f^\epsilon = \xi^\epsilon + \epsilon^\beta \chi^\epsilon$  en choisissant  $p^\epsilon$  dans le noyau de l'opérateur de transport et  $\xi^\epsilon$  dans celui de l'opérateur de collisions. Cela permet d'aboutir à un système à trois inconnues ( $f^\epsilon, q^\epsilon, \chi^\epsilon$ ) sans aucun terme raide. Il est néanmoins nécessaire de régulariser ce système à l'aide de deux petits paramètres  $\sigma$  et  $\sigma'$  contre un habituellement, où seule une inconnue supplémentaire venait s'ajouter à l'inconnue de départ  $f^\epsilon$ .

Nous avons implémenté la procédure numérique ainsi décrite que nous avons testé sur le cas test issu de [25]. En particulier, ce dernier considère un champ électrique donné, sinusoïdal. Nous retrouvons en particulier la relation de Boltzmann à la limite  $\epsilon \rightarrow 0$  :

$$f^\epsilon(t, x, v) \xrightarrow{\epsilon \rightarrow 0} c(t) e^{-\varphi(t, x)} e^{-v^2/2} \quad c(t) = \frac{\int_0^1 n_0(x) dx}{\int_0^1 e^{-\varphi} dx}, \quad n_0(x) = \frac{1}{\sqrt{2\pi\vartheta}} \exp\left(-\frac{(x-0.5)^2}{2\vartheta^2}\right),$$

avec  $\vartheta = 0.05$ . Ces premiers résultats encourageants permettraient de considérer dans un travail futur non plus un champ électrique donné, mais auto-consistant en résolvant l'équation de Poisson.

### Quelques perspectives analytiques et numériques sur l'écoulement de Kolmogorov :

Dans le Chapitre 4, l'écoulement instable de Kolmogorov a été étudié. Si un résultat analytique a pu être apporté sur la phase linéaire de l'instabilité, aucun résultat en revanche n'est venu soutenir les constatations numériques du régime non-linéaire et de la saturation de l'équilibre final obtenu.

En particulier, certaines simulations numériques ont montré (peu après la phase linéaire de l'instabilité) la présence d'une phase où l'amplitude du champ de vortacité suivait une évolution d'oscillations amorties. Cette évolution a également été constaté dans [68]<sup>9</sup>. Une investigation plus poussé de ce phénomène, notamment de l'éventuelle influence du rapport d'aspect sur la décroissance des oscillations pourrait être l'objet d'un travail ultérieur.

D'un point de vue analytique, la caractérisation des équilibres stables de l'écoulement de Kolmogorov semble être un prolongement logique. Car si les simulations numériques semblent montrer que la phase de saturation de l'instabilité débouche sur l'apparition d'un équilibre, aucun argument analytique ne vient appuyer sa stabilité.

Ceci clos le chapitre d'introduction. Les trois chapitres suivant reprennent en détail les points énoncés dans les résumés exposés dans ce chapitre d'introduction.

---

9. Dans ces travaux, le modèle présenté diffère légèrement du nôtre, par l'ajout d'un terme visco-élastique, modélisant un plasma "à grains".





---

---

## CHAPITRE 2

---

# NUMERICAL STUDY OF AN ANISOTROPIC VLASOV EQUATION ARISING IN PLASMA PHYSICS

Le contenu de ce chapitre correspond à un article écrit avec C. Negulescu. Il a donné lieu à la publication suivante :

B. Fedele and C. Negulescu, **Numerical study of an anisotropic Vlasov equation arising in plasma physics**, *Kinetic and Related Models*, **11** (2018), 1395–1426.

### Abstract.

Le but de ce chapitre est d'étudier différents schémas numériques pour la résolution de deux équations de Vlasov anisotropes. Ces deux modèles étant obtenus à partir d'une description cinétique des équations régissant la dynamique des plasmas de tokamak, confinés par un champ magnétique intense. La simplicité des modèles étudiés permettront une compréhension profonde des schémas numériques utilisés, en particulier leurs propriétés préservant l'asymptotique. Cette approche nous permettra de choisir le schéma numérique le plus adapté pour des modèles plus réalistes et donc plus complexes.

*La suite de ce chapitre est écrite en anglais.*

## 2.1 Introduction

The present study addresses a new approach for an efficient numerical resolution of anisotropic transport models, which simplified are of the type

$$\begin{cases} \partial_t f^\epsilon + \frac{\mathbf{u}}{\epsilon} \cdot \nabla f^\epsilon = 0, & \forall (t, x, y) \in [0, T] \times \Omega, \\ f^\epsilon(0, x, y) = f_{in}(x, y), \end{cases} \quad (2.1)$$

subject to appropriate boundary conditions (here periodic ones). The unknown  $f^\epsilon$  stands for the quantity (distribution function) which is advected along the given (or self-consistently computed) field  $\mathbf{u}$  in the domain  $\Omega := [0, L_x] \times [0, L_y]$  and the small scaling parameter  $\epsilon \ll 1$  indicates that we have to deal with very strong advection fields  $\mathbf{u}$  or equivalently with the long-time asymptotics of  $f^\epsilon$ . Such anisotropic transport models arise very often in physics, as simplifications of more complex systems. In Section 2.2 we detail some examples coming from plasma physics, as the Vlasov equation for the ion dynamics in the gyrokinetic regime. There are however several other examples arising in physics and leading to a simplified transport equation as (2.1), for example when one studies the long-time asymptotics of the incompressible Euler 2D equations, (2.1) representing then the vorticity equation, which has to be coupled (via  $\mathbf{u}$ ) with a Poisson equation for the stream-function computation [102].

A numerical resolution of problems of the type (2.1) is rather challenging in the regime  $\epsilon \ll 1$ , due to the singularity of the mathematical problem as  $\epsilon \rightarrow 0$ . Certainly, the exact solution of the simple transport-case (2.1) is known for  $\epsilon > 0$ , however not in general situations, when  $\mathbf{u}$  is self-consistently computed via  $f^\epsilon$  and when other (not-stiff) terms are present. These general situations require then an efficient numerical treatment of (2.1). From a physical point of view we can say that we have to cope with a multiscale problem, the parameter  $\epsilon$  being the stiffness parameter. Standard schemes (explicit hyperbolic approaches) require very restrictive CFL-conditions (dependent on  $\epsilon$ ) in order to accurately account for the microscopic  $\epsilon$ -scales. Very often in such situations people are impliciting the stiff term [56], in order to avoid these too restrictive CFL-conditions. This can work in some situations, for example when the grid is aligned with the anisotropy, and only for a certain range of  $\epsilon$ -values. However in more general configurations, not-aligned grids and  $\epsilon$ -values covering all the interval  $[0, 1]$ , impliciting the stiff term is no more sufficient, as shall be seen in this paper. We propose thus in this work a new numerical procedure, based on *Asymptotic-Preserving* arguments, being able to solve (2.1) in an efficient manner, uniformly accurate and stable in  $\epsilon$ , and this on a simple, Cartesian grid. *Asymptotic-Preserving* methods are efficient, as they are designed in order to mimic on the discrete level the asymptotic behavior of the singularly perturbed problem solutions (see [80, 112] for a detailed introduction).

This paper was initiated by the repetitive remarks/questions one of the authors got during conferences, meaning that impliciting the stiff term in (2.1) is enough to get an efficient AP-scheme, which behaves well even in the limit  $\epsilon \rightarrow 0$ . The aim of this paper is to prove the contrary, AP-schemes are more than impliciting the stiff term. In order to understand in detail the main features of the here proposed AP-scheme, we preferred to keep the investigated model as simple as possible, so that a detailed numerical analysis is possible, permitting to perceive the differences of our scheme when compared to standard (implicit) schemes. We hope that doing so, we are able to resolve some of the confusion that surround AP-schemes. However, even if the here presented results are based on a simplified model as (2.1), the same *Asymptotic-Preserving* approach can be used for more involved anisotropic

transport problems, such as those presented in Section 2.2 and which shall be the objective of an upcoming work.

The AP-procedure we propose here was employed in other contexts by the authors (elliptic [35, 36], parabolic [107]). The present setting is more stimulating, as we have to cope with highly oscillating problems when  $\epsilon \ll 1$  and no more dissipative ones. In the present oscillating case, the limit (weak)  $\epsilon \rightarrow 0$  is more challenging, and has to be defined adequately. We refer the reader to [31, 32, 93] for other AP-scheme references.

This paper is laid as follows. Section 2.2 deals with the presentation of a physical situation leading, after scaling and simplification, to the anisotropic transport equation (2.1). Two simplified models which will be studied in the following, are presented. Section 3.2 reviews the mathematical framework necessary to study the first toy model, and investigates the asymptotic limit  $\epsilon \rightarrow 0$ . Section 3.2.3 introduces several numerical schemes that we shall apply for the resolution of the first toy model. Then, we present the numerical results obtained with these schemes in Section 2.5 and the numerical analysis in Section 2.6. The last section is dedicated to the mathematical and numerical study of the second toy model which considers variable coefficients. A conclusion gives some hints for our upcoming work, concerning the more realistic Vlasov equation (2.4).

## 2.2 Physical motivation and toy models

Let us shortly say here some words about the physical motivation of the present work and introduce the two simplified models we shall investigate numerically in the next sections. These simplified models are caricatures of typical asymptotic regimes encountered in plasma physics, as for example the gyro-kinetic regime, and contain all the numerical difficulties arising in the more complex real physical systems.

The core tokamak plasma can be considered as collisionless, such that the most appropriate model for the description of its dynamics is the Vlasov equation for each particle species ( $\alpha = e$  for electrons and  $\alpha = i$  for ions), *i.e.*

$$\partial_t f_\alpha + \mathbf{v} \cdot \nabla_{\mathbf{x}} f_\alpha + \frac{e_\alpha}{m_\alpha} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \nabla_{\mathbf{v}} f_\alpha = 0, \quad (2.2)$$

where  $e_\alpha = \pm e$  resp.  $m_\alpha$  are the particle elementary charge resp. mass and  $\mathbf{E}(t, \mathbf{x})$  resp.  $\mathbf{B}(t, \mathbf{x})$  are the electric respectively magnetic fields, determined self-consistently from Maxwell's equations. In the electrostatic case (given field  $\mathbf{B}$ ), Maxwell's equations have to be replaced by Poisson's equation

$$-\epsilon_0 \Delta \Phi = \rho, \quad \rho(t, \mathbf{x}) := \sum_{\alpha} e_{\alpha} \int_{\mathbb{R}^3} f_{\alpha}(t, \mathbf{x}, \mathbf{v}) d\mathbf{v}, \quad (2.3)$$

where  $\Phi$  is the electrostatic potential, related to the electric field  $\mathbf{E}$  by  $\mathbf{E}(t, \mathbf{x}) = -\nabla \Phi(t, \mathbf{x})$ . For more details about the modelling of magnetically confined fusion plasmas, we refer the interested reader to the textbooks [26, 64, 73].

From a numerical point of view, solving the system (2.2)-(2.3) is rather arduous, due among others to its high dimensionality (6 dimensional in the phase space  $(\mathbf{x}, \mathbf{v})$ ) and to the presence of several time and space scales in the dynamics, introduced for ex. by the strong magnetic field  $\mathbf{B}$  which confines the plasma in the tokamak. We shall be concerned

in the present work with the multi-scale aspects of the kinetic problem, difficulties which are described mathematically by the following rescaling of the Vlasov equation for the ions (see [19, 62, 63, 67, 113] for the gyrokinetic scaling)

$$\partial_t f + \mathbf{v} \cdot \nabla_x f + \left[ \mathbf{E} + \frac{1}{\epsilon} (\mathbf{v} \times \mathbf{B}) \right] \cdot \nabla_v f = 0, \quad (2.4)$$

where  $\epsilon$  stands for the ratio of the particle cyclotron period to the observation time. The electrons experience the appearance of a second small parameter, related to the small electron to ion mass ratio  $m_e/m_i$ , leading to additional numerical burden, we shall not consider here (see [25]). The effect of the intense magnetic field on the particle dynamics is that it introduces a strong anisotropy, the motion of the charged particles being splitted into a fast gyration around the magnetic field lines and a slow dynamics along these lines, separation which necessarily causes numerical complications.

From the equation (2.2), choosing now characteristic scales for our phenomenon shall permit to identify small parameters, describing specific plasma regimes. In particular, let us write each quantity under the form  $n(\cdot) = \bar{n} n'(\cdot)$ , where the constant  $\bar{n}$  stands for the characteristic scale of this quantity and  $n'(\cdot)$  is the dimensionless variable. Using this notation, let us denote by  $\bar{t}$  the observation time and by  $\bar{x}$  the characteristic space scale. The characteristic microscopic velocity is then fixed as  $\bar{v} := \bar{x}/\bar{t}$ . Furthermore, let the kinetic energy be of the same order as the potential energy of the problem, implying thus

$$m_i \bar{v}^2 = e_i \bar{\phi} \quad \Rightarrow \quad \bar{E} = \frac{\bar{\phi}}{\bar{x}} = \frac{m_i \bar{v}^2}{e_i \bar{x}}.$$

Finally, let us introduce the cyclotron frequency of the ions

$$\omega_c := \frac{e_i \bar{B}}{m_i}.$$

The plasma regime we shall be interested in, concerns a strongly confined plasma situation, in other words, we shall suppose that the observation time  $\bar{t}$  is much larger that the time scale  $\omega_c^{-1}$  of the cyclotron oscillations, *i.e.*

$$\epsilon := \frac{1}{\bar{t} \omega_c} \ll 1.$$

Writing thus the Vlasov equation (2.2) in dimensionless variables, namely

$$\partial_{t'} f' + \frac{\bar{t} \bar{v}}{\bar{x}} \mathbf{v}' \cdot \nabla_{x'} f' + \frac{e_i \bar{t} \bar{E}}{m_i \bar{v}} \mathbf{E}' \cdot \nabla_{v'} f' + \frac{e_i \bar{t} \bar{B}}{m_i} (\mathbf{v}' \times \mathbf{B}') \cdot \nabla_{v'} f' = 0,$$

and using the relations just introduced for the characteristic scales, leads finally to the rescaled Vlasov equation we are interested to study, *i.e.*

$$\partial_t f + \mathbf{v} \cdot \nabla_x f + \left[ \mathbf{E} + \frac{1}{\epsilon} (\mathbf{v} \times \mathbf{B}) \right] \cdot \nabla_v f = 0, \quad (2.5)$$

where we omitted for simplicity reasons the primes.

Let us introduce now two simplified toy models, which contain all the numerical difficulties of the initial model. In the rest of this paper we shall consider a homogeneous

magnetic field  $\mathbf{B} = b \mathbf{b}$  with fixed direction  $\mathbf{b} := e_z$  and constant magnitude  $|\mathbf{B}| = b \equiv 1$ . Furthermore, let us also introduce the following notation

$$\mathbf{v}_{\parallel} = (0, 0, v_z)^t, \quad \mathbf{v}_{\perp} = (v_x, v_y, 0)^t, \quad {}^{\perp}\mathbf{v} := (v_y, -v_x, 0)^t = \mathbf{v} \times \mathbf{B}.$$

Sometimes it is more convenient to shift in (2.4) from Cartesian coordinates to polar coordinates for the velocity, *i.e.*

$$v = (v_x, v_y, v_z) \Leftrightarrow (r, \theta, v_z), \quad \begin{cases} v_x := r \cos(\theta) \\ v_y := r \sin(\theta) \end{cases}, \quad \begin{cases} \theta \in [0, 2\pi) \\ r \geq 0 \end{cases}.$$

The Vlasov equation (2.4), written in polar coordinates, has then the form

$$\begin{aligned} \partial_t F + v_z \partial_z F + E_z \partial_{v_z} F + (E_x \cos \theta + E_y \sin \theta) \partial_r F - \frac{1}{r} (E_x \sin \theta - E_y \cos \theta) \partial_{\theta} F \\ + r (\cos \theta \partial_x F + \sin \theta \partial_y F) - \frac{1}{\epsilon} \partial_{\theta} F = 0, \end{aligned} \quad (2.6)$$

where the unknown is now  $F(t, x, y, z, r, \theta, v_z)$ .

The two formulations, (2.4) resp. (2.6), corresponding to a Cartesian (not field-aligned) resp. polar (field-aligned) configuration, are different from a numerical point of view, and different numerical schemes are usually employed for their resolution. To understand this difference better, we shall investigate in the present work in detail some numerical schemes for simplified versions of (2.4) and (2.6). We deliberately simplified these equations in order to be able to do a complete numerical analysis and to understand in all details the features of the here introduced AP-schemes.

### 2.2.1 First toy model - Polar, field-aligned configuration

Let us start from the Vlasov equation (2.4), assume here that  $\mathbf{E} \equiv 0$ ,  $\mathbf{B} = e_z$  and consider furthermore only the dynamics in the perpendicular plane  $(x, y)$ , *i.e.*

$$\partial_t f + \mathbf{v}_{\perp} \cdot \nabla_x f + \frac{1}{\epsilon} (\mathbf{v} \times \mathbf{B}) \cdot \nabla_v f = 0, \quad (2.7)$$

where  $\epsilon \ll 1$  accounts as usual for very strong magnetic fields. In order to simplify the computations, one often shifts to polar coordinates for the velocity, leading to

$$\partial_t F + r \cos \theta \partial_x F + r \sin \theta \partial_y F - \frac{1}{\epsilon} \partial_{\theta} F = 0, \quad (2.8)$$

where the unknown now is  $F(t, x, y, r, \theta)$ . We recognize thus a simple 3D anisotropic transport equation, the variable  $r$  being considered as a parameter in (2.8).

Choosing an initial condition  $F_{in}$  independent on the variable  $y$ , would even lead to a more simpler 2D transport model

$$\partial_t F + r \cos \theta \partial_x F - \frac{1}{\epsilon} \partial_{\theta} F = 0. \quad (2.9)$$

This problem represents the simplest example of an anisotropic advection equation, to be understood in detail before designing an efficient scheme for the resolution of the Vlasov equation in the gyrokinetic regime (2.4). It is sufficiently difficult in order to study the behavior of the various schemes we shall introduce, and shall be the starting point of Section 3.2.

## 2.2.2 Second toy model - Cartesian, not field-aligned configuration

In this second part, we shall differently simplify our Vlasov equation in order to study a different behavior. In particular, setting  $\mathbf{E} \equiv 0$ ,  $\mathbf{B} \equiv e_z$  and taking an initial condition independent on the space variable, yields the following 2D equation, in Cartesian coordinates

$$\partial_t f + \frac{1}{\epsilon} (\mathbf{v} \times \mathbf{B}) \cdot \nabla_v f = 0, \quad (2.10)$$

or equivalently

$$\partial_t f + \frac{v_y}{\epsilon} \partial_{v_x} f - \frac{v_x}{\epsilon} \partial_{v_y} f = 0. \quad (2.11)$$

The difference of this model to the previous one is that this time the characteristics are no more straight lines but curves, such that the numerical schemes will behave differently. As mentioned earlier, these two models correspond to simplified versions of a field-aligned, polar coordinate framework, as well as a not field-aligned, Cartesian framework, both associated to the Vlasov equation (2.4) in the gyro-kinetic regime.

## 2.2.3 Aim of the present paper

The main points we are interested in within this study are the following :

- design of AP-schemes for an efficient numerical resolution of anisotropic Vlasov equations of type (2.9), (2.10). Important properties we are asking from the schemes are : (a) stability independent on  $\epsilon$ ; (b) numerical diffusion/accuracy independent on  $\epsilon$ ; (c) discretization of the limit model as  $\epsilon \rightarrow 0$ ;
- show that taking the stiff term  $\frac{1}{\epsilon} (\mathbf{v} \times \mathbf{B}) \cdot \nabla_v f$  in (2.4) implicitly is not sufficient for having an AP-scheme, meaning that AP-schemes are more than taking “implicitly” the suitable terms. AP-schemes have to mimic at the discrete level the precise asymptotic behavior of the solution in the limit  $\epsilon \rightarrow 0$ ;
- perform a detailed numerical analysis of the presented schemes in the framework of the two simplified toy-models (2.9), (2.10) and identify exactly which are the particularities of each scheme and each equation;
- understand the difference between a field-aligned framework (2.9) and a Cartesian framework (2.10), and this from a numerical point of view;
- prepare the foundation for a future, more realistic work, dealing with the resolution of the initial Vlasov equation (2.4) in the gyro-kinetic regime.

Finally, let us say some words about *Asymptotic-Preserving* schemes. In general, inaccuracy in numerical simulations can result from applying unstable algorithms to well-conditioned problems or stable algorithms to ill-conditioned problems. Dealing with singularly-perturbed problems is a hard task, as they are ill-conditioned from the beginning. A standard, stable discretization (implicit in this case) often results in inaccurate results. The essence of AP-procedures is to replace singularly-perturbed problems by equivalent problems, which are regularly perturbed, well-conditioned problems, leading to uniformly accurate results, if stable algorithms are used (AP-approach).

## 2.3 First anisotropic Vlasov toy model and its mathematical study

Let us investigate now in detail the following simplified toy model, corresponding to a field-aligned anisotropic Vlasov equation

$$(V)_\epsilon \begin{cases} \partial_t f^\epsilon + a \partial_x f^\epsilon + \frac{b}{\epsilon} \partial_y f^\epsilon = 0, & \forall (t, x, y) \in [0, T] \times [0, L_x] \times [0, L_y], \\ f^\epsilon(0, x, y) = f_{in}(x, y), \end{cases} \quad (2.12)$$

where  $f_{in}$  is a given initial condition,  $a > 0$  and  $b > 0$  are for the moment constants and  $0 < \epsilon \ll 1$  is a parameter representing the strong anisotropy/stiffness of the problem. Our computational domain is a doubly periodic box  $\Omega := [0, L_x] \times [0, L_y]$ .

We shall review here some standard numerical schemes as well as introduce some new ones for the resolution of such a singularly perturbed problem and discuss finally their advantages and disadvantages. In particular, one is interested in numerical schemes capable to solve (2.12) uniformly accurate in  $\epsilon$ , so-called ‘‘Asymptotic-Preserving’’ schemes.

Let us however start with a detailed mathematical study of the behavior of (2.12).

### 2.3.1 Singularly perturbed problem

Equation (2.12) is a simple advection problem, whose exact solution is given by the characteristic method, *i.e.*

$$f^\epsilon(t, x, y) = f_{in}\left(x - at, y - \frac{b}{\epsilon} t\right), \quad \forall (t, x, y) \in [0, T] \times \Omega. \quad (2.13)$$

Remark that this function is  $L_x$ -periodic in the variable  $x$ ,  $L_y$ -periodic in the variable  $y$ . Concerning the time-variable, two time-scales are present in the problem, a slow time-scale  $t$  and a rapid one  $t/\epsilon$ .

The term  $\frac{b}{\epsilon} \partial_y f^\epsilon$  in (2.12) is the dominant term in the case where  $\epsilon \ll 1$ , such that passing formally to the limit  $\epsilon \rightarrow 0$ , yields

$$(R) \begin{cases} \partial_y f = 0, & \forall (t, x, y) \in [0, T] \times [0, L_x] \times [0, L_y], \\ f(0, x, y) = f_{in}(x, y). \end{cases} \quad (2.14)$$

This system, called ‘‘reduced system’’, is ill-posed. Depending on the initial condition, it can admit or an infinite number of solutions, namely if  $\partial_y f_{in} = 0$ , or no regular solution (if  $\partial_y f_{in} \neq 0$ ). From a numerical point of view, this ill-posedness in the limit is translated into the singularity of the matrix of the linear system obtained by discretization of this problem. In particular, trying to solve (2.12) in a standard manner will necessarily lead to a linear system which degenerates in the limit  $\epsilon \rightarrow 0$ . This shall induce sever numerical problems.

More adequate schemes are hence necessary for an efficient resolution of (2.12), as for example ‘‘Asymptotic-Preserving’’ schemes which are uniformly stable and accurate independently on the small parameter  $\epsilon$ , and are additionally able to capture the limit model as  $\epsilon \rightarrow 0$ .



### 2.3.2 Limit model

For a better comprehension of our singularly-perturbed problem as well as for the construction of efficient “Asymptotic-Preserving” schemes, we have to identify the limit problem  $(V)_0$  of (2.12) and its solution denoted by  $f^0$ . The information we get from the reduced model is that the limit-function  $f^0$  has to be  $y$ -independent. With this information we introduce now the average of the function  $f^\epsilon$  with respect to the direction  $y$

$$\bar{f}^\epsilon(t, x) := \frac{1}{L_y} \int_0^{L_y} f^\epsilon(t, x, y) dy.$$

Integration of the equation (2.12) with respect to  $y$  yields  $\partial_t \bar{f}^\epsilon + a \partial_x \bar{f}^\epsilon = 0$ , which is an  $\epsilon$ -independent problem. Passing then to the limit  $\epsilon \rightarrow 0$  leads to the advection equation

$$(V)_0 \begin{cases} \partial_t f^0 + a \partial_x f^0 = 0, & \forall (t, x) \in [0, T] \times [0, L_x], \\ f^0(0, x) = \bar{f}_{in}(x), & \forall x \in [0, L_x], \end{cases} \quad (2.15)$$

with solution

$$f^0(t, x) = \bar{f}_{in}(x - at), \quad \forall (t, x) \in [0, T] \times [0, L_x].$$

The system  $(V)_0$  is what we call “limit-system” of the anisotropic Vlasov equation  $(V)_\epsilon$ , as shall be proved in the next section.

### 2.3.3 Weak convergence

So far, we proved the existence of a unique solution  $f^\epsilon$  for the system  $(V)_\epsilon$  resp.  $f^0$  for the limit system  $(V)_0$ . The next step is now to show the weak-convergence of  $f^\epsilon$  towards  $f^0$  as  $\epsilon \rightarrow 0$ , and this in a certain sense. To define this sense, we have to introduce the right mathematical framework. In the sequel the symbol  $\sharp$  shall underline the periodicity of the considered space.

**Theorem 2.3.1** *Let the initial condition  $f_{in} \in H_{\sharp}^1(\Omega)$ . Then the unique solutions to  $(V)_\epsilon$  resp.  $(V)_0$  satisfy  $f^\epsilon \in W^{1,\infty}(0, T; L_{\sharp}^2(\Omega)) \cap L^\infty(0, T; H_{\sharp}^1(\Omega))$  resp.  $f^0 \in W^{1,\infty}(0, T; L_{\sharp}^2(0, L_x)) \cap L^\infty(0, T; H_{\sharp}^1(0, L_x))$ . Moreover, we have the weak- $\star$  limit*

$$f^\epsilon \xrightarrow[\epsilon \rightarrow 0]{*} f^0 \quad \text{in } L^\infty(0, T; L_{\sharp}^2(\Omega)). \quad (2.16)$$

**Proof 1** *To prove (2.16), which signifies*

$$\int_0^T \int_{\Omega} (f^\epsilon(t, x, y) - f^0(t, x)) \phi(t, x, y) dx dy dt \xrightarrow[\epsilon \rightarrow 0]{} 0 \quad \forall \phi \in L^1(0, T; L_{\sharp}^2(\Omega)),$$

*we shall introduce first a primitive of the function  $f_{in}(x, \cdot) - \bar{f}_{in}(x)$ , i.e.*

$$g(x, y) := \int_0^y (f_{in}(x, z) - \bar{f}_{in}(x)) dz.$$

It follows that the function  $g$  belongs to  $H_{\#}^1(\Omega)$  such that  $g^\epsilon(t, x, y) := g(x-at, y-bt/\epsilon)$  belongs to  $W^{1,\infty}(0, T; L_{\#}^2(\Omega)) \cap L^\infty(0, T; H_{\#}^1(\Omega))$ . The  $L_y$ -periodicity of  $g$  is seen by the simple computation

$$\begin{aligned} g(x, y + L_y) &= \int_0^{y+L_y} \left( f_{in}(x, z) - \bar{f}_{in}(x) \right) dz = \int_{-L_y}^y f_{in}(x, z) dz - \bar{f}_{in}(x)(y + L_y) \\ &= \int_0^y f_{in}(x, z) dz - \bar{f}_{in}(x)y + \int_0^{L_y} f_{in}(x, z) dz - \bar{f}_{in}(x)L_y = g(x, y). \end{aligned}$$

Taking now an arbitrary test function  $\phi \in C_0^1(0, T; L_{\#}^2(\Omega))$  and introducing for simplicity for each  $f, g \in L_{\#}^2(\Omega)$  the bracket  $\langle f, g \rangle := \int_{\Omega} f g dx dy$ , we have

$$\begin{aligned} \int_0^T \left\langle f_{in}\left(x-at, y-\frac{b}{\epsilon}t\right) - \bar{f}_{in}(x-at), \phi(t) \right\rangle dt &= \int_0^T \left\langle (\partial_y g)\left(x-at, y-\frac{b}{\epsilon}t\right), \phi(t) \right\rangle dt \\ &= -\frac{\epsilon}{b} \left[ \int_0^T \left\langle \partial_t \left[ g\left(x-at, y-\frac{b}{\epsilon}t\right) \right] + a(\partial_x g)\left(x-at, y-\frac{b}{\epsilon}t\right), \phi(t) \right\rangle dt \right] \\ &= \frac{\epsilon}{b} \int_0^T \left\langle g\left(x-at, y-\frac{b}{\epsilon}t\right), \phi'(t) \right\rangle dt - \frac{\epsilon a}{b} \int_0^T \left\langle (\partial_x g)\left(x-at, y-\frac{b}{\epsilon}t\right), \phi(t) \right\rangle dt. \end{aligned}$$

As  $g^\epsilon \in W^{1,\infty}(0, T; L_{\#}^2(\Omega)) \cap L^\infty(0, T; H_{\#}^1(\Omega))$ , we can estimate

$$\forall \phi \in C_0^1(0, T; L_{\#}^2(\Omega)), \quad \left| \int_0^T \left\langle \left[ f_{in}\left(x-at, y-\frac{b}{\epsilon}t\right) - \bar{f}_{in}(x-at) \right], \phi(t) \right\rangle dt \right| \leq C\epsilon,$$

where  $C > 0$  is a constant independent on  $\epsilon$ . Therefore,

$$\forall \phi \in C_0^1(0, T; L_{\#}^2(\Omega)), \quad \int_0^T \left\langle \left[ f_{in}\left(x-at, y-\frac{b}{\epsilon}t\right) - \bar{f}_{in}(x-at) \right], \phi(t) \right\rangle dt \xrightarrow{\epsilon \rightarrow 0} 0,$$

which concludes the proof due to the dense injection  $C_0^1(0, T; L_{\#}^2(\Omega)) \subset L^1(0, T; L_{\#}^2(\Omega))$ .

### 2.3.4 Two-scale convergence

Let us investigate the two-scale convergence of  $f^\epsilon$ , which is more appropriate to be used in the context of high frequency oscillations. Indeed, we remind that our problem is two-scale with respect to the time variable, fact represented by the parameter  $\epsilon$ .

The weak limit of a given oscillating function-sequence does not adequately capture the oscillating behaviour, it gives only information about the average. We need hence to generalize the weak convergence and introduce the so-called two-scale convergence which is more adequate in this context. In particular, this new concept is able to capture the oscillations in the solution  $f^\epsilon$  through its limit.

To define the two-scale convergence, let us denote by  $X$  an open subset of  $\mathbb{R}^n$ ,  $Y \subset \mathbb{R}^m$  a periodic domain and  $C_{\#}^\infty(Y)$  the set of functions  $f$  which belong to  $C^\infty(\mathbb{R}^m)$  and are  $Y$ -periodic.

**Definition 1** Let  $1 < p \leq \infty$ . A sequence of functions  $\{f^\epsilon\}_{\epsilon>0}$  in  $L^p(X)$  is said to be “two-scale” convergent towards  $F \in L^p(X; L_{\#}^p(Y)) \simeq L^p(X \times Y)$  if for every test function  $\psi \in L^q(X; C_{\#}(Y))$ , with  $1/p + 1/q = 1$ , we have :

$$\lim_{\epsilon \rightarrow 0} \int_X f^\epsilon(x) \psi\left(x, \frac{x}{\epsilon}\right) dx = \frac{1}{|Y|} \int_X \int_Y F(x, y) \psi(x, y) dy dx. \quad (2.17)$$

The two-scale convergence is denoted simply by  $f^\epsilon \xrightarrow[\epsilon \rightarrow 0]{2-s} F$ .

Note that the space of test functions  $L^q(X; C_{\#}(Y))$  can be replaced by  $C_0^\infty(X) \otimes C_{\#}^\infty(Y)$  or  $C(\bar{X}; L_{\#}^\infty(Y))$ , but never by  $L^q(X \times Y)$  [3, 4].

Observe that the test functions are  $Y$ -periodic and that the sequence oscillates at the same frequency as the considered test functions. If the oscillations of the sequence and of the considered test functions are not in resonance, the two-scale convergence will never see these oscillations and shall convert into a usual weak convergence.

Finally, let us also remark that the two-scale convergence implies the weak convergence and that the strong convergence implies two-scale convergence, however the converse implications are not true.

Now, we return to our problem and prove the following theorem :

**Theorem 2.3.2** *Let  $f_{in} \in H_{\#}^1(\Omega)$  and  $f^\epsilon \in W^{1,\infty}(0, T; L_{\#}^2(\Omega)) \cap L^\infty(0, T; H_{\#}^1(\Omega))$  the unique solutions to  $(V)_\epsilon$ . Then, there exist a function  $F \in L^\infty(0, T; L_{\#}^\infty(0, L_\tau; L_{\#}^2(\Omega)))$ ,  $L_\tau := L_y/b$ -periodic in the second variable  $\tau$ , such that*

$$f^\epsilon \xrightarrow[\epsilon \rightarrow 0]{2-s} F, \quad (2.18)$$

meaning that for each test function  $\psi \in L^1(0, T; C_{\#}(0, L_\tau; L_{\#}^2(\Omega)))$  one has

$$\int_0^T \int_{\Omega} f^\epsilon(t, x, y) \psi(t, t/\epsilon, x, y) dx dy dt \xrightarrow[\epsilon \rightarrow 0]{} \frac{1}{L_\tau} \int_0^T \int_{\Omega} \int_0^{L_\tau} F(t, \tau, x, y) \psi(t, \tau, x, y) d\tau dx dy dt.$$

Furthermore, we have the following relation between  $F$  and the weak-limit  $f^0$  of  $(V)_0$

$$f^0(t, x) = \frac{1}{L_\tau} \int_0^{L_\tau} F(t, \tau, x, y) d\tau.$$

**Proof 2** *Let us now demonstrate that  $f^\epsilon$  converges in the two-scale sense towards  $F$ , which is defined by :*

$$F(t, \tau, x, y) := f_{in}(x - at, y - b\tau), \quad \forall (t, \tau, x, y) \in [0, T] \times [0, L_\tau] \times [0, L_x] \times [0, L_y].$$

For this, let us choose test functions  $\psi \in C_0^1(0, T; C_{\#}([0, L_\tau]; L_{\#}^2(\Omega)))$ , denote as usual  $\langle f, g \rangle := \int_{\Omega} f g dx dy$  and show that

$$\lim_{\epsilon \rightarrow 0} \int_0^T \left\langle f_{in}\left(x - at, y - b \frac{t}{\epsilon}\right), \psi\left(t, \frac{t}{\epsilon}, x, y\right) \right\rangle dt = \frac{1}{L_\tau} \int_0^{L_\tau} \int_0^T \left\langle f_{in}(x - at, y - b\tau), \psi(t, \tau, x, y) \right\rangle dt d\tau.$$

First of all, we have,

$$\int_0^T \left\langle f_{in}\left(x - at, y - b \frac{t}{\epsilon}\right), \psi\left(t, \frac{t}{\epsilon}, x, y\right) \right\rangle dt = \epsilon \int_0^{\frac{T}{\epsilon}} \left\langle f_{in}(x - a\epsilon\tau, y - b\tau), \psi(\epsilon\tau, \tau, x, y) \right\rangle d\tau.$$

For each  $\epsilon > 0$  there exists an index  $k_\epsilon \in \mathbb{N}$  such that  $0 \leq \frac{T}{\epsilon} - k_\epsilon L_\tau < L_\tau$  and  $\lim_{\epsilon \rightarrow 0} k_\epsilon = +\infty$ . With this, we have then the following decomposition :

$$\epsilon \int_0^{\frac{T}{\epsilon}} \left\langle f_{in}(x - a\epsilon\tau, y - b\tau), \psi(\epsilon\tau, \tau, x, y) \right\rangle d\tau \quad (2.19)$$

$$= \epsilon \left[ \sum_{k=0}^{k_\epsilon-1} \int_{kL_\tau}^{(k+1)L_\tau} \left\langle f_{in}(x - a\epsilon\tau, y - b\tau), \psi(\epsilon\tau, \tau) \right\rangle d\tau + \int_{k_\epsilon L_\tau}^{\frac{T}{\epsilon}} \left\langle f_{in}(x - a\epsilon\tau, y - b\tau), \psi(\epsilon\tau, \tau) \right\rangle d\tau \right]. \quad (2.20)$$

The second term on the right hand side is converging towards zero, as  $\epsilon$  goes to zero. Indeed, one has the estimate

$$\epsilon \left| \int_{k_\epsilon L_\tau}^{\frac{T}{\epsilon}} \left\langle f_{in}(x - a\epsilon\tau, y - b\tau), \psi(\epsilon\tau, \tau) \right\rangle d\tau \right| \leq \epsilon \left( \frac{T}{\epsilon} - k_\epsilon L_\tau \right) \|f_{in}\|_{L^\infty(0, T; L_\#^2)} \|\psi\|_{L^\infty(L_\#^2)} \leq C\epsilon,$$

with a constant  $C > 0$  independent on  $\epsilon$ .

For the first term on the right hand side of (2.19), we recognize the expression of a Riemann sum. Indeed, we have

$$\begin{aligned} & \int_{kL_\tau}^{(k+1)L_\tau} \left\langle f_{in}(x - a\epsilon\tau, y - b\tau), \psi(\epsilon\tau, \tau) \right\rangle d\tau \\ &= \int_0^{L_\tau} \left\langle f_{in}(x - a\epsilon(s + kL_\tau), y - bs), \psi(\epsilon(s + kL_\tau), s) \right\rangle ds \\ &= \int_0^{L_\tau} \left\langle f_{in}(x - a\epsilon kL_\tau, y - bs), \psi(\epsilon kL_\tau, s) \right\rangle ds + \mathcal{O}(\epsilon), \end{aligned}$$

after using a Taylor expansion, since we have suitable properties for both  $f_{in}$  and  $\psi$ . This leads to

$$\begin{aligned} & \epsilon \sum_{k=0}^{k_\epsilon-1} \int_{kL_\tau}^{(k+1)L_\tau} \left\langle f_{in}(x - a\epsilon\tau, y - b\tau), \psi(\epsilon\tau, \tau) \right\rangle d\tau \\ &= \frac{\epsilon L_\tau}{L_\tau} \sum_{k=0}^{k_\epsilon-1} \int_0^{L_\tau} \left\langle f_{in}(x - a\epsilon kL_\tau, y - bs), \psi(\epsilon kL_\tau, s) \right\rangle ds + \mathcal{O}(\epsilon) \\ &\xrightarrow{\epsilon \rightarrow 0} \frac{1}{L_\tau} \int_0^{L_\tau} \int_0^T \left\langle f_{in}(x - at, y - bs), \psi(t, s, x, y) \right\rangle dt ds, \end{aligned}$$

using that  $k_\epsilon$  is of order  $\mathcal{O}(1/\epsilon)$ . This concludes the proof, as  $C_0^1(0, T; C_\#([0, L_\tau]; L_\#^2(\Omega)))$  is dense in  $L^1(0, T; C_\#([0, L_\tau]; L_\#^2(\Omega)))$ .

## 2.4 Numerical schemes for the anisotropic Vlasov equation

In this section we shall now introduce several numerical schemes for the resolution of (2.12) and examine them in more details. Firstly, different time semi-discretizations will be presented and then some words mentioned about a standard upwind space-discretization. The time-discretization is the most important step in the construction of AP-schemes.

For this, let us first introduce the following homogeneous discretizations of our time interval  $[0, T]$  as well as of our simulation domain  $\Omega = [0, L_x] \times [0, L_y]$ :

$$\begin{aligned} \Delta t &:= T/N_t, \quad N_t \in \mathbb{N}; & t_n &:= n * \Delta t, \quad n = 0, \dots, N_t \\ \Delta x &:= L_x/(N_x - 1), \quad N_x \in \mathbb{N}; & x_i &:= (i - 1) * \Delta x, \quad i = 1, \dots, N_x \\ \Delta y &:= L_y/(N_y - 1), \quad N_y \in \mathbb{N}; & y_j &:= (j - 1) * \Delta y, \quad j = 1, \dots, N_y. \end{aligned} \quad (2.21)$$

We denote by  $Q_h$  the index domain  $Q_h := [0, N_t] \times [1, N_x] \times [1, N_y] \subset \mathbb{N}^3$ . We shall denote further by  $f^{\epsilon, n}$  resp.  $f_{ij}^{\epsilon, n}$  the numerical approximation of  $f^\epsilon(t^n, x, y)$  resp.  $f^\epsilon(t^n, x_i, y_j)$ . Recall also that we consider a doubly-periodic framework, such that

$$f_{0,j}^{\epsilon, n} = f_{N_x-1,j}^{\epsilon, n}, \quad f_{1,j}^{\epsilon, n} = f_{N_x,j}^{\epsilon, n}, \quad f_{i,0}^{\epsilon, n} = f_{i,N_y-1}^{\epsilon, n}, \quad f_{i,1}^{\epsilon, n} = f_{i,N_y}^{\epsilon, n}, \quad \forall (n, i, j) \in Q_h.$$

## 2.4.1 Semi-discretization in time

### IMEX scheme

The first time semi-discretization we shall study will be the implicit-explicit (IMEX) Euler method, where the stiff term is taken implicitly, *i.e*

$$(IMEX)_\epsilon \quad \frac{f^{\epsilon, n+1} - f^{\epsilon, n}}{\Delta t} + a \partial_x f^{\epsilon, n} + \frac{b}{\epsilon} \partial_y f^{\epsilon, n+1} = 0, \quad \forall n \geq 0. \quad (2.22)$$

To study the behavior of this scheme, as  $\epsilon$  becomes smaller, let us formally let  $\epsilon$  go to zero in (2.22) and get

$$\partial_y f^{0, n+1}(x, y) = 0, \quad \forall (x, y) \in \Omega.$$

This equation admits an infinite amount of solutions, namely all periodic functions dependent only on  $x$ . This formal analysis permits hence to conclude that the IMEX scheme can not be an AP-scheme, as it does not capture correctly the asymptotic behavior of the problem, which is rather given by the limit problem  $(V)_0$ . This property shall be tested numerically in Section 2.5.

### Fourier method/Micro-Macro method

A different way to solve (2.12) is to use a partial Fourier transform in the variable  $y$ , which is possible here, as we are in a simplified periodic context with constant coefficients. Denoting indeed the Fourier coefficients by

$$\widehat{f}_k^\epsilon(t, x) := \frac{1}{L_y} \int_0^{L_y} f^\epsilon(t, x, y) e^{-i\omega_y k y} dy, \quad \forall k \in \mathbb{Z}, \quad \omega_y := \frac{2\pi}{L_y},$$

one has

$$f^\epsilon(t, x, y) = \sum_{k=-\infty}^{\infty} \widehat{f}_k^\epsilon(t, x) e^{i\omega_y k y}, \quad (2.23)$$

where the Fourier coefficients are solutions of the system

$$\begin{cases} \partial_t \widehat{f}_0^\epsilon + a \partial_x \widehat{f}_0^\epsilon = 0, & \forall (t, x) \in [0, T] \times [0, L_x] \\ \partial_t \widehat{f}_k^\epsilon + a \partial_x \widehat{f}_k^\epsilon + i\omega_y k \frac{b}{\epsilon} \widehat{f}_k^\epsilon = 0, & \forall k \neq 0, \forall (t, x) \in [0, T] \times [0, L_x]. \end{cases} \quad (2.24)$$

A simple discretization of this problem can be

$$(F)_\epsilon \quad \begin{cases} \frac{\widehat{f}_0^{\epsilon, n+1} - \widehat{f}_0^{\epsilon, n}}{\Delta t} + a \partial_x \widehat{f}_0^{\epsilon, n} = 0, & \forall n \geq 0 \\ \frac{\widehat{f}_k^{\epsilon, n+1} - \widehat{f}_k^{\epsilon, n}}{\Delta t} + a \partial_x \widehat{f}_k^{\epsilon, n} + i\omega_y k \frac{b}{\epsilon} \widehat{f}_k^{\epsilon, n+1} = 0, & \forall k \neq 0, \forall n \geq 0. \end{cases}$$

Solving this system and using the inverse Fourier transform (2.23) permits to get the desired result, *i.e.* the values of the unknowns  $f_{ij}^{\epsilon,n}$ , solution of (2.12).

Let us investigate now the behavior of this system when  $\epsilon \rightarrow 0$ . Formally we get

$$(F)_0 \quad \begin{cases} \frac{\widehat{f_0^{\epsilon,n+1}} - \widehat{f_0^{\epsilon,n}}}{\Delta t} + a \partial_x \widehat{f_0^{\epsilon,n}} = 0, & \forall n \geq 0 \\ \widehat{f_k^{\epsilon,n+1}} = 0, & \forall k \neq 0, \forall n \geq 0. \end{cases}$$

Therefore, we find a discretized version of the Vlasov limit problem  $(V)_0$ , signifying that this method will be ‘‘Asymptotic-Preserving’’.

The Fourier method is very nice, however it can be applied only in a simplified periodic framework with constant coefficients. As a sort of generalization one can think at the micro-macro method [31], which is based on the decomposition of each quantity in its mean part over the variable  $y$ , denoted by  $H^\epsilon$  or simply  $\bar{f}^\epsilon$ , and the fluctuation part  $h^\epsilon$  or simply  $(f^\epsilon)'$ , defined as follows

$$H^\epsilon(t, x) := \frac{1}{L_y} \int_0^{L_y} f^\epsilon(t, x, y) dy, \quad h^\epsilon(t, x, y) := f^\epsilon(t, x, y) - H^\epsilon(t, x), \quad \bar{h}^\epsilon = 0.$$

Taking now the average of the advection equation (2.12) over  $y$  and subtracting the resulting equation then from the initial one, yields a system to be solved for the unknowns  $(H^\epsilon, h^\epsilon)$ , *i.e.*

$$(MM)_\epsilon \quad \begin{cases} \partial_t H^\epsilon + a \partial_x H^\epsilon = 0, & \forall (t, x) \in [0, T] \times [0, L_x] \\ \partial_t h^\epsilon + a \partial_x h^\epsilon + \frac{b}{\epsilon} \partial_y h^\epsilon = 0, & \forall (t, x, y) \in [0, T] \times \Omega \\ \bar{h}^\epsilon = 0, & \forall (t, x) \in [0, T] \times [0, L_x]. \end{cases} \quad (2.25)$$

Let us study now the behavior of this system when  $\epsilon \rightarrow 0$ . We have formally

$$(MM)_0 \quad \begin{cases} \partial_t H^0 + a \partial_x H^0 = 0, & \forall (t, x) \in [0, T] \times [0, L_x] \\ \partial_y h^0 = 0, & \forall (t, x, y) \in [0, T] \times \Omega \\ \bar{h}^0 = 0, & \forall (t, x) \in [0, T] \times [0, L_x]. \end{cases} \quad (2.26)$$

The two last equations establish that  $h^0 \equiv 0$ . Hence the system  $(MM)_0$  is nothing else than the Vlasov limit system  $(V)_0$ . Again, we have created a scheme which is a regular perturbation of the asymptotic limit model, and shall be hence ‘‘Asymptotic-Preserving’’.

This method is rather similar to Fourier method, however more general, as it can be applied in rather broad contexts. To understand this similitude, remark that  $H^\epsilon$  is nothing else than the first Fourier coefficient  $\widehat{f_0^\epsilon}$  and the fluctuation  $h^\epsilon$  regroups the remaining Fourier modes. However, there is still a disadvantage or difficulty, namely the implementation of the constraint  $\bar{h}^\epsilon = 0$ , which is crucial for the passage to the limit  $\epsilon \rightarrow 0$ . It is this constraint which permits in the limit to get a unique  $h^0$  and to have thus a well-posed limit problem  $(MM)_0$ . But averaging along the anisotropy lines can be very difficult in more general contexts, for ex. when these lines are not aligned with the axes.

## Lagrange-multiplier method

The Lagrange-multiplier method is based on the idea to replace the stiff, dominant term  $\frac{b}{\epsilon} \partial_y f$  by a smoother one  $\partial_y q$ , yielding the system

$$(La)_\epsilon \begin{cases} \partial_t f^\epsilon + a \partial_x f^\epsilon + b \partial_y q^\epsilon = 0, & \forall (t, x, y) \in [0, T] \times \Omega \\ \partial_y f^\epsilon = \epsilon \partial_y q^\epsilon, & \forall (t, x, y) \in [0, T] \times \Omega \\ q^\epsilon|_{\Gamma_{in}} = 0, \end{cases} \quad (2.27)$$

where the inflow boundary is defined as  $\Gamma_{in} := \{(x, y) \in \partial\Omega / y = 0\}$ . In the limit  $\epsilon \rightarrow 0$  one remarks that  $q^\epsilon$  is a sort of Lagrange multiplier corresponding to the constraint  $\partial_y f^0 = 0$ , where the name of the method.

First, we will prove the equivalence between the system  $(La)_\epsilon$  and the Vlasov equation  $(V)_\epsilon$ , proving thus the well-posedness of the reformulation  $(La)_\epsilon$ . For this, let us first consider the unique solution  $f^\epsilon$  of  $(V)_\epsilon$  and prove the existence of a function  $q^\epsilon$  such that  $(f^\epsilon, q^\epsilon)$  solves  $(La)_\epsilon$ . Since  $f_{in} \in H_{\sharp}^1(\Omega)$ , we have  $f^\epsilon \in \mathcal{V} := W^{1,\infty}(0, T; L_{\sharp}^2(\Omega)) \cap L^\infty(0, T; H_{\sharp}^1(\Omega))$ . The kernel of the dominant operator  $\frac{b}{\epsilon} \partial_y f$ , denoted by  $\mathcal{G}$ , reads :

$$\mathcal{G} := \{f^\epsilon \in \mathcal{V}, \partial_y f^\epsilon = 0\}.$$

Then we shall decompose  $f^\epsilon$  in the following manner, which is somehow similar to a Hilbert Ansatz :

$$f^\epsilon = p^\epsilon + \epsilon q^\epsilon, \quad (2.28)$$

with  $(p^\epsilon, q^\epsilon) \in \mathcal{G} \times \mathcal{V}$ . To have a unique decomposition, we have to single out the  $\mathcal{G}$ -part of  $q^\epsilon$ , by fixing for example  $q^\epsilon$  on the inflow boundary  $\Gamma_{in}$ , choosing  $q^\epsilon \in \mathcal{Q}$  with

$$\mathcal{Q} := \{q^\epsilon \in \mathcal{V}, q^\epsilon|_{\Gamma_{in}} = 0\}.$$

Obviously, we have  $\mathcal{G} \cap \mathcal{Q} = \{0_{\mathcal{V}}\}$ , implying the uniqueness of the decomposition (2.28). Replacing now this decomposition in the system  $(V)_\epsilon$ , we obtain directly the system  $(La)_\epsilon$ , which proves the existence of a solution to  $(La)_\epsilon$ . The converse is trivial, meaning that for  $(f^\epsilon, q^\epsilon) \in \mathcal{V} \times \mathcal{Q}$  solution to  $(La)_\epsilon$ ,  $f^\epsilon$  solves  $(V)_\epsilon$ . Altogether, we have proved the equivalence between both systems.

Now let us consider the limit problem of  $(La)_\epsilon$ , obtained by letting formally  $\epsilon \rightarrow 0$  in (2.27)

$$(La)_0 \begin{cases} \partial_t f^0 + a \partial_x f^0 + b \partial_y q^0 = 0, & \forall (t, x, y) \in [0, T] \times \Omega \\ \partial_y f^0 = 0, & \forall (t, x, y) \in [0, T] \times \Omega \\ q^0|_{\Gamma_{in}} = 0. \end{cases} \quad (2.29)$$

The second equation leads to  $f^0 = \bar{f}^0$ . Then, averaging the first equation of (2.29) in the  $y$ -variable, yields

$$\partial_t \bar{f}^0 + a \partial_x \bar{f}^0 = 0, \quad (2.30)$$

where we used that  $q^0$  is  $L_y$ -periodic. This equation permits the determination of the limit function  $f^0$ . Furthermore, the remaining well-posed system

$$\begin{cases} b \partial_y q^0 = -\partial_t f^0 - a \partial_x f^0, & \forall (t, x, y) \in [0, T] \times \Omega \\ q^0|_{\Gamma_{in}} = 0, & \forall (t, x) \in \times [0, T] \times [0, L_x], \end{cases} \quad (2.31)$$

can be solved to assure finally the existence of the unique solution  $(f^0, q^0)$  for the limit problem  $(La)_0$ .

The Lagrangian scheme seems to be the most “far-reaching” AP-scheme . The only disadvantage of this method is that we have now two unknowns and hence two equations to be solved, meaning longer simulation times. However, we are no more forced to follow the anisotropy lines and can choose coarse Cartesian, not-field aligned grids.

## 2.4.2 Space discretization for the IMEX scheme

For any numerical scheme presented above, we decided to consider the standard upwind method to discretize the transport terms in the equation (2.12). The idea behind this choice is that the space-discretization is not the important step in the construction of an AP-scheme, such that we opted for a simple discretization, in order not to embroil the further numerical analysis as well as the understanding of the main ideas of our methods. The same arguments incited us to select only first order discretizations in time. A Runge-Kutta coupled to a second-order space-discretization would be naturally more accurate, changes however nothing in the essential concept of our AP-strategies. As mentioned earlier, in a forthcoming paper we shall be concerned with a realistic, fusion plasma situation, such that we shall adapt the most adequate of the here presented schemes to more accurate second order techniques, to gain in accuracy.

Now, let us recall the first-order upwind forms : for all  $(n, i, j) \in Q_h$ ,

$$a \partial_x f_{i,j}^{\epsilon,n} \approx a \frac{f_{i,j}^{\epsilon,n} - f_{i-1,j}^{\epsilon,n}}{\Delta x}, \text{ if } a > 0, \quad a \partial_x f_{i,j}^{\epsilon,n} \approx a \frac{f_{i+1,j}^{\epsilon,n} - f_{i,j}^{\epsilon,n}}{\Delta x}, \text{ if } a < 0.$$

We have analogous formulae for the partial derivative in the  $y$ -variable. Denoting now  $\alpha := \frac{a\Delta t}{\Delta x} > 0$  and  $\beta := \frac{b\Delta t}{\Delta y} > 0$  and using the periodicity, *i.e.*

$$f_{0,j}^{\epsilon,n} = f_{N_x-1,j}^{\epsilon,n}, \quad f_{1,j}^{\epsilon,n} = f_{N_x,j}^{\epsilon,n}, \quad f_{i,0}^{\epsilon,n} = f_{i,N_y-1}^{\epsilon,n}, \quad f_{i,1}^{\epsilon,n} = f_{i,N_y}^{\epsilon,n}, \quad \forall (n, i, j) \in Q_h.$$

the completely discretized IMEX scheme writes finally :

$$(IMEX)_\epsilon \quad (\epsilon + \beta) f_{i,j}^{\epsilon,n+1} - \beta f_{i,j-1}^{\epsilon,n+1} = \epsilon(1 - \alpha) f_{i,j}^{\epsilon,n} + \epsilon\alpha f_{i-1,j}^{\epsilon,n},$$

for all  $(n, i, j) \in [0, N_t - 1] \times [1, N_x - 1] \times [1, N_y - 1]$ . We remark that we can rewrite this scheme like a system of  $N_x - 1$  equations :

$$\mathcal{A} \mathcal{F}_i^{n+1} = \mathcal{B}_i^n, \quad \forall n \geq 0, \quad \forall i \in [1, N_x - 1], \quad (2.32)$$

where :

$$\mathcal{A} = \begin{pmatrix} (\epsilon + \beta) & 0 & \dots & 0 & -\beta \\ -\beta & \ddots & 0 & 0 & 0 \\ 0 & \ddots & \ddots & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -\beta & (\epsilon + \beta) \end{pmatrix}, \quad \mathcal{F}_i^{n+1} = \begin{pmatrix} f_{i,1}^{\epsilon,n+1} \\ f_{i,2}^{\epsilon,n+1} \\ \vdots \\ f_{i,N_y-2}^{\epsilon,n+1} \\ f_{i,N_y-1}^{\epsilon,n+1} \end{pmatrix},$$

$$\mathcal{B}_i^n = \begin{pmatrix} \epsilon(1 - \alpha) f_{i,1}^{\epsilon,n} + \epsilon\alpha f_{i-1,1}^{\epsilon,n} \\ \epsilon(1 - \alpha) f_{i,2}^{\epsilon,n} + \epsilon\alpha f_{i-1,2}^{\epsilon,n} \\ \vdots \\ \epsilon(1 - \alpha) f_{i,N_y-2}^{\epsilon,n} + \epsilon\alpha f_{i-1,N_y-2}^{\epsilon,n} \\ \epsilon(1 - \alpha) f_{i,N_y-1}^{\epsilon,n} + \epsilon\alpha f_{i-1,N_y-1}^{\epsilon,n} \end{pmatrix}.$$



At each time step, we resolve this system  $\forall i \in [1, N_x - 1]$ , to get the unknowns  $f_{i,j}^{\epsilon, n+1}$ . Remark that  $\mathcal{A} = \epsilon Id + \mathcal{C}_\beta$  is a regular perturbation of a singular, cyclic matrix  $\mathcal{C}_\beta$ .

## 2.5 Numerical simulations

In this part, we shall test numerically every scheme introduced in the previous Section for the resolution of the anisotropic Vlasov equation (2.12). The homogeneous time and phase-space discretization was previously introduced in (2.21) and we choose in the sequel the following parameters :  $T = 1$ ,  $L_x = 2\pi$ ,  $L_y = 2\pi$ ,  $N_t = 101$ ,  $N_x = N_y = 201$ ,  $a = 0.1$  and  $b = 1$ . Changes in these parameters shall be explicitly mentioned. The initial condition we adopt is given by :

$$f_{in}(x, y) := \sin(x) \left( \cos(2y) + 1 \right), \quad \forall (x, y) \in \Omega := [0, L_x] \times [0, L_y].$$

We recall that the exact solution of (2.12) is known and reads, for each  $\epsilon > 0$  :

$$f_{ex}^\epsilon(t, x, y) = \sin \left( x - at \right) \left[ \cos \left( 2 \left( y - \frac{b}{\epsilon} t \right) \right) + 1 \right], \quad \forall (t, x, y) \in [0, T] \times \Omega.$$

In Figure 4.3, we reveal two graphics which contain on the one hand  $f_{in}$  and on the other hand  $f_{ex}^\epsilon$  at the final time  $T = 1$ .

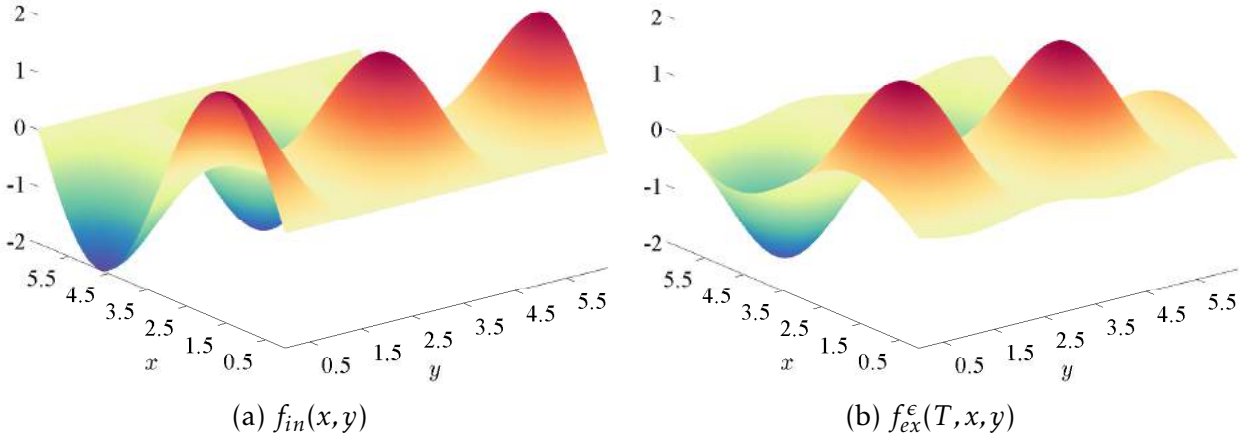


FIGURE 2.1 – Representation of the initial condition  $f_{in}$  (A) and the exact solution  $f_{ex}^\epsilon$  at the final time  $T = 1$  (B). Here  $\epsilon = 1$ .

Furthermore, in order to better figure out our problem, we plotted in Figure 4.4 the exact solution of the limiting Vlasov system (2.15) at the final time  $T$ , i.e.  $f_{ex}^0(T, x) = \bar{f}_{in}(x - aT)$ . Remark that this solution is homogeneous in the  $y$ -variable.

Finally, we show in Figure 2.3 the time-evolution of the exact solution  $f_{ex}^\epsilon$  at one point only, i.e.  $(x_{N_x-1}, y_{N_y-1})$ . We distinguish easily on the left plot (A) of Fig. 2.3 the two periods, one linked with the  $x$ -variable, and the other one corresponding to the  $y$ -variable. This last one is  $\epsilon$ -dependent and we see that more  $\epsilon$  is small, more the frequency of the time-oscillations becomes important. As the 2D situation is not so eloquent, we eliminate the  $x$ -variable in the problem and considered also a 1D problem, keeping only the term containing the parameter  $\epsilon$  (i.e.  $a = 0$ ). The time-evolution of the exact solution at the point

$y_{N_y-1}$  is now plotted in Fig. 2.3 (B). One observes here more easily that with smaller becoming  $\epsilon$ , the frequency of the time-oscillations is increasing. In the limit  $\epsilon \rightarrow 0$ ,  $f^\epsilon(t, y_{N_y-1})$  converges weakly towards the average, which is here the constant 1.

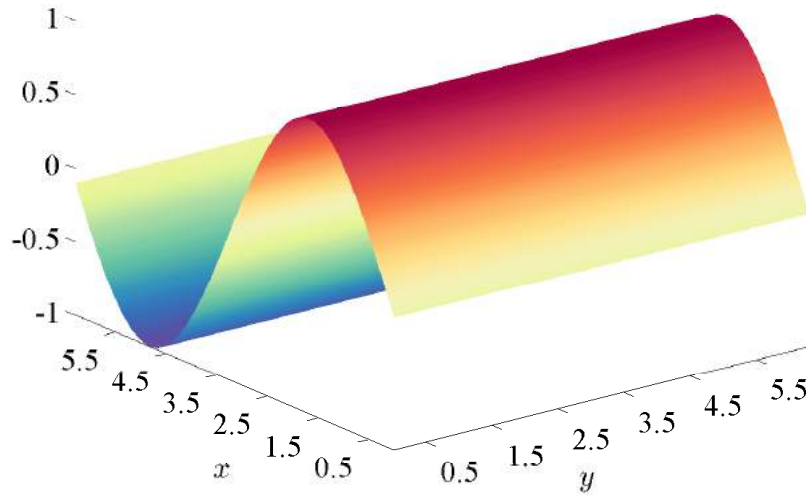
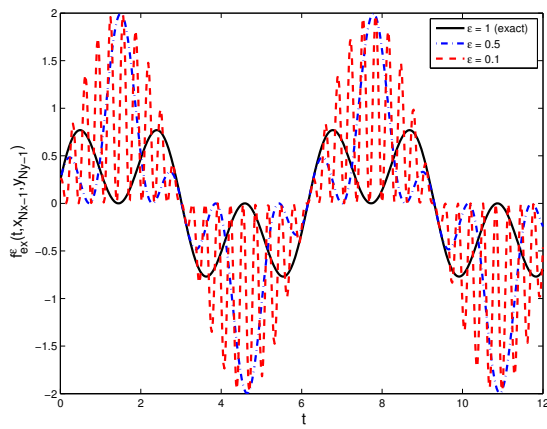
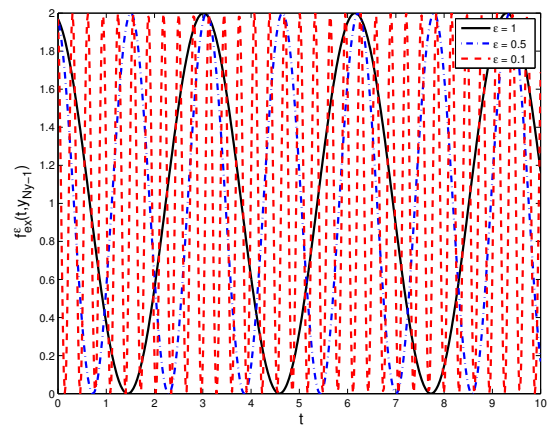


FIGURE 2.2 – Representation of the exact limit solution  $f_{ex}^0(t, x)$  at the final time  $T$ .



(a)  $f_{ex}^\epsilon(t, x_{N_x-1}, y_{N_y-1})$



(b)  $f_{ex}^\epsilon(t, y_{N_y-1})$

FIGURE 2.3 – Time-evolution of the exact solution at point  $(x_{N_x-1}, y_{N_y-1})$  in the two dimensional case (A) with  $T = 12$  and  $N_t = 501$ ; resp. at point  $y_{N_y-1}$  in the one dimensional case with  $T = 10$ ,  $a = 0$  and  $N_t = 501$  (B).

### 2.5.1 Some results obtained with our schemes

Now we examine how the different numerical schemes introduced above cope with such an asymptotic behavior.

## IMEX scheme

We start by first showing in Fig. 4.5 as well as in the left plot of Fig. 4.6 the numerical solution  $f^\epsilon$  via the IMEX-scheme, for three different values of  $\epsilon$ , namely  $\epsilon = 1$ ,  $\epsilon = 0.1$  and  $\epsilon = 10^{-10}$ , all of them at the final time  $T = 1$ .

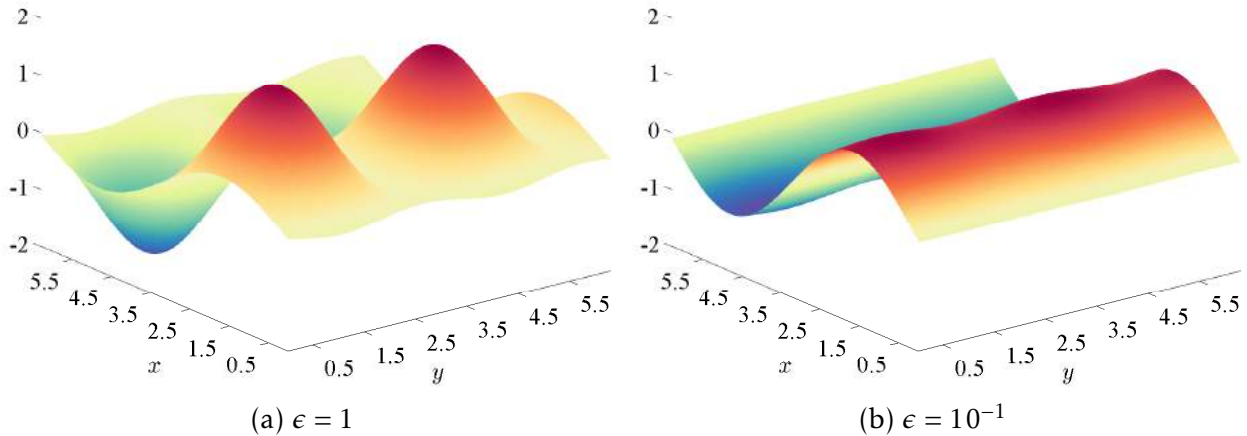


FIGURE 2.4 – Representation of the numerical solution  $f^\epsilon$  for two values of  $\epsilon$ , and at the final time  $T$ , corresponding to the IMEX scheme.

For  $\epsilon = 1$ , we recognize an approximation of the exact solution (see Figure 4.3) and for  $\epsilon = 10^{-10}$ , the limit solution is clearly obtained (see Figure 4.4). Briefly one can say that the numerical solution follows the weak- $\star$  convergence  $f^\epsilon \xrightarrow{\star} f^0$  as  $\epsilon$  becomes smaller and smaller. But, one can remark a numerical diffusion which leads to a loss of amplitude, especially visible in the non-limit case  $\epsilon = 1$  or  $\epsilon = 10^{-1}$ . To observe better this numerical diffusion, we show in the right plot of Fig. 4.6 the time-evolution of just one point of the numerical solution, corresponding again to a 1D situation as the one plotted on the right of Fig. 2.3, and this for several values of  $\epsilon$ . As one can observe, the damping is more and more

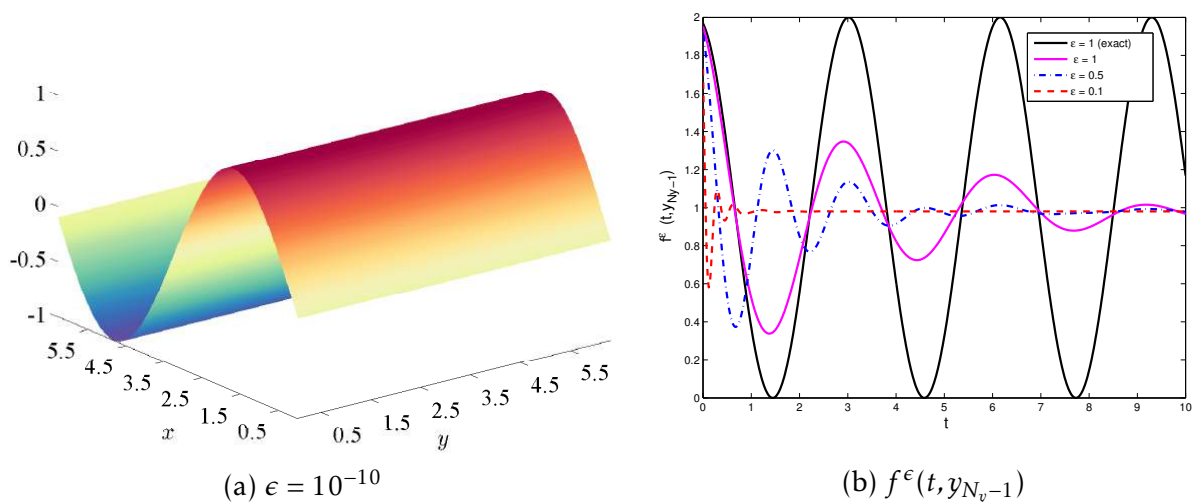


FIGURE 2.5 – Left (A) : Plot of the num. sol.  $f^\epsilon$  for  $\epsilon = 10^{-10}$ , at the final time  $T$ . Right (B) : Time-evolution of the IMEX scheme sol. at point  $y_{N_y-1}$  in the 1D case for  $T = 10$  and several  $\epsilon$ . We have added the exact solution for  $\epsilon = 1$ .

pronounced if  $\epsilon \rightarrow 0$ . For small  $\epsilon$ -values the numerical solution recovers quasi immediately the weak limit solution, here the constant 1. This damping phenomenon will be understood from the numerical analysis we shall fulfill in Section 2.6.

### Fourier, Micro-Macro and Lagrange-multiplier schemes

Let us now present analogous results for the remaining schemes, namely the Fourier, Micro-Macro and Lagrange-multiplier schemes. The 2D plots are rather similar to the ones presented for the IMEX-scheme (see Fig. 4.5-4.6). To examine the difference between these methods, we preferred to plot in Fig. 2.6 only the time-evolution of the numerical solution in the 1D-context again. We remark that the damping of the Fourier method is more slowly

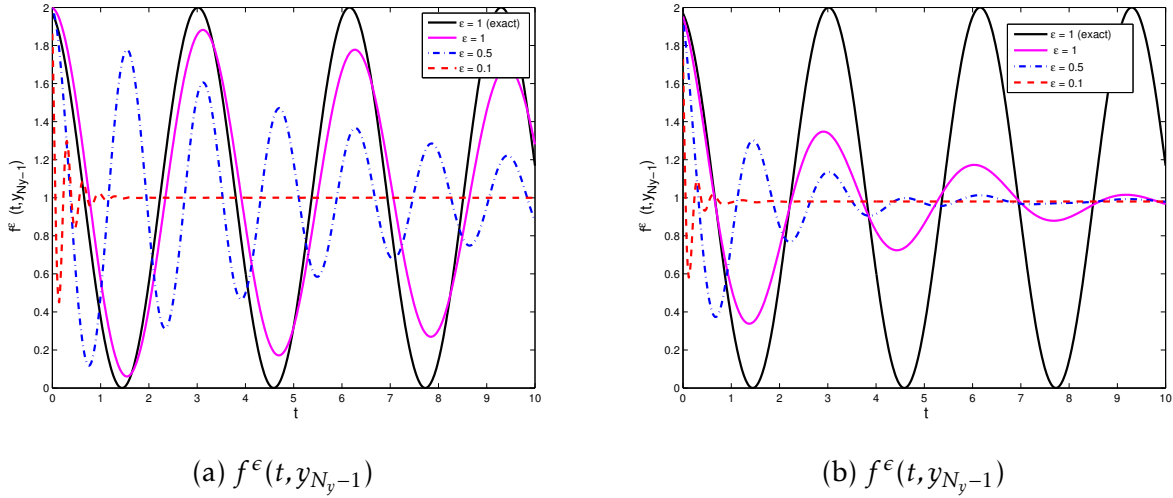


FIGURE 2.6 – Time-evolution of the solution via Fourier (A) and IMEX, MM- resp. Lagrange-multiplier schemes (B), at  $y_{N_y-1}$  in 1D with  $T = 10$ ,  $a = 0$ ,  $N_t = 501$ . We have added in both cases the exact solution for  $\epsilon = 1$ .

than the ones of the IMEX-scheme as well as Micro-Macro and Lagrange-multiplier scheme (which are completely overlapping). But, once again we observe that in the limit  $\epsilon/t \rightarrow 0$ , the fluctuations are completely damped out and we recover the weak limit solution.

### 2.5.2 Convergence of the schemes for fixed $\epsilon > 0$

Let us study now the convergence of the here presented schemes with respect to time and space, and this for fixed  $\epsilon > 0$ , permitting to show their validity in the large  $\epsilon$ -regime. For this, fix  $\epsilon > 0$  and consider the error between exact and numerical solutions as a function of the mesh-size, at the final time  $T$ . Firstly, concerning the convergence with respect to  $\Delta t$ , we choose small space steps ( $N_x = N_y = 501$ ) such that the space errors are much smaller than the time error and vary then the time step. Equally we apply the same strategy for the convergence with respect to  $\Delta x$  and  $\Delta y$ , by fixing a time step of  $N_t = 501$ . In all cases, the parameter  $\epsilon$  is fixed to 1. In Figure 2.7, we have plotted curves in log-log scale, showing the evolution of the errors as a function of  $\Delta x$ ,  $\Delta y$  and  $\Delta t$ , respectively.

As expected, we observe that all schemes are first order in time and space. Some comments are however necessary to understand Figure 2.7. First, the slope of the curves gets smaller than 1 in the small-grid ranges. This is due to the fact that the error to be investigated (for ex. in  $\Delta t$ ) becomes as small as the fixed error term (in  $\Delta x$ ,  $\Delta y$ ) and saturates. Secondly, the slope of the curves becomes also smaller in the large-grid ranges. This is

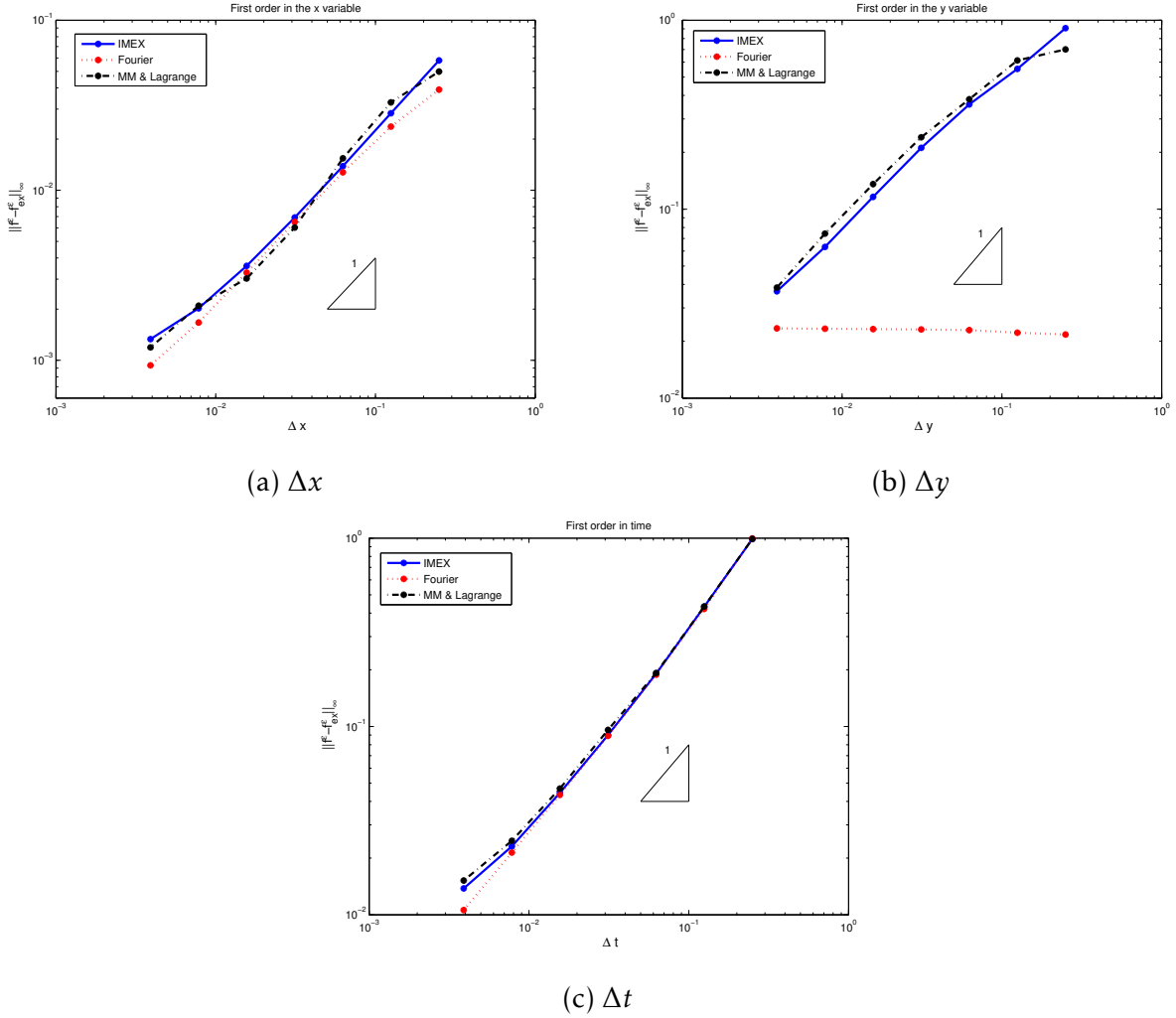


FIGURE 2.7 – Evolution of the  $L^\infty$ -error between  $f_{ex}^\epsilon(t, \cdot)$  and  $f^\epsilon(t, \cdot)$  at final time  $T = 1$  and for  $\epsilon = 1$ , as a function of  $\Delta x$  (with  $N_y = 15001$ ,  $N_t = 15001$ ),  $\Delta y$  (with  $N_x = 15001$ ,  $N_t = 15001$ ) and  $\Delta t$  (with  $N_x = N_y = 1001$ ).

usual, as for large discretization steps, the rest-terms in the Taylor series for the error analysis can no longer be neglected. Finally, we would like to draw the attention of the reader to the Fourier error curve, which has a constant slope in (B). This is completely natural, as the Fourier method has spectral accuracy.

### 2.5.3 Asymptotic behavior as $\epsilon \rightarrow 0$

To begin the study of the asymptotic behavior, we define the following two errors

$$\eta_\epsilon(t) = \max_{i,j} |f_{ex,i,j}^\epsilon - f_{num,i,j}^\epsilon|(t), \quad \gamma_\epsilon(t) = \max_{i,j} |f_{num,i,j}^\epsilon - f_{ex,i,j}^0|(t),$$

where  $\eta_\epsilon(t)$  represents the  $L^\infty$ -error between the exact and the numerical solution at instant  $t$ , for fixed  $\epsilon > 0$ , whereas  $\gamma_\epsilon(t)$  denotes the  $L^\infty$ -error at instant  $t$  between the numerical solution  $f_{num}^\epsilon$  and the exact limit solution  $f_{ex}^0$ .

We are interested in the evolution of these two errors at the final time  $T$  as functions of  $\epsilon$ . The curves corresponding to the different schemes are plotted in Figure 4.12. As expected, we observe a decrease of  $\eta_\epsilon(T)$  and an increase of  $\gamma_\epsilon(T)$  when  $\epsilon \rightarrow 1$ . For  $\epsilon \rightarrow 0$  the

converse behavior is observed. This plot shows that each scheme approximates well either the exact solution  $f_{ex}^\epsilon$  for large  $\epsilon$ , or the exact limit solution  $f_{ex}^0$  for small  $\epsilon$ .

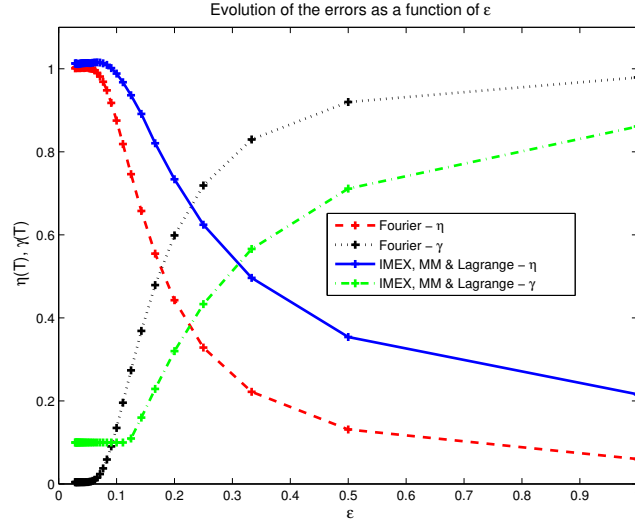


FIGURE 2.8 – Evolution of  $\eta_\epsilon(T)$  and  $\gamma_\epsilon(T)$  as a function of  $\epsilon$  for each scheme.

What can be said as a conclusion, is that all schemes seem to have the right asymptotic behavior in this simple test case. Indeed, for fixed  $\epsilon > 0$ , each numerical solution  $f_{num}^\epsilon$  converges to the expected solution  $f_{ex}^\epsilon$  as long as the grid is refined (Fig. 2.7). For fixed discretization steps, all numerical solutions  $f_{num}^\epsilon$  converge towards the limit solution  $f^0$  when  $\epsilon$  becomes smaller and smaller, underlying the AP property of our methods.

It is worth mentioning however that the IMEX-scheme is no more working for  $\epsilon$  smaller than  $10^{-14}$ , the matrix  $\mathcal{A}$  of the IMEX linear-system (2.32), namely

$$\mathcal{A} \mathcal{F}_i^{n+1} = \mathcal{B}_i^n, \quad \mathcal{A} = \epsilon Id + \mathcal{C}_\beta, \quad \det \mathcal{C}_\beta = 0,$$

is becoming numerically singular in the limit  $\epsilon \rightarrow 0$ . This is not the case for the Micro-Macro as well as Lagrange-multiplier schemes, which give accurate results even for  $\epsilon = 0$ . This difference in the behavior can be observed also from the study of the condition-number of the discretization matrices, paying attention especially on the  $\epsilon$ -dependence. Remark here that an “Asymptotic-Preserving scheme” must have an  $\epsilon$ -independent condition number, depending merely on the discretization parameters  $\Delta x, \Delta y$ .

In Fig. 2.9 we plotted thus the matrix condition-number  $cond(\mathcal{A}) := \|\mathcal{A}^{-1}\|_2 \|\mathcal{A}\|_2$  corresponding to the three schemes (IMEX, Micro-Macro and Lagrange-multiplier) as a function of  $\epsilon$ . What can be observed is that for the Micro-Macro and Lagrange-multiplier scheme, the condition-number is  $\epsilon$ -independent (for  $\epsilon \leq 10^{-2}$ ), which is a hint of the well-posedness of these two problems in the limit  $\epsilon \rightarrow 0$ , namely of  $(MM)_0$  resp.  $(La)_0$ . On the other hand, for the IMEX-scheme  $cond(\mathcal{A})$  is proportional to  $1/\epsilon$  (slope of the curve is approx.  $-1$ ). This circumstance is the translation on the discrete level of the fact that the reduced model (3.2), obtained on the continuous level by letting formally  $\epsilon \rightarrow 0$  in the IMEX time-discretization, is ill-posed, admitting an infinite amount of solutions.

However, even if these arguments show clearly that the IMEX-method should behave badly for very small  $\epsilon$ -values, it is not the case in our simplified toy model, in particular it does not seem to be affected by the bad condition number. This will no more be the case in our second toy-model. To understand in detail what happens, a more refined error

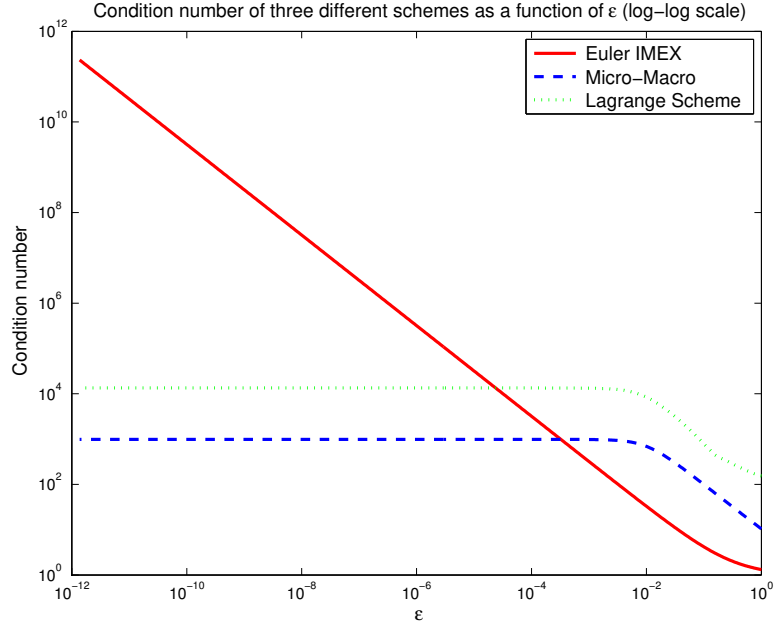


FIGURE 2.9 – Condition number  $cond(A)$  as a function of  $\epsilon$  in log-log scale. The three curves correspond to the IMEX, Micro-Macro and Lagrange-multiplier schemes.

study could be profitable and shall be done in the next section. The final interpretation is postponed to Section 2.6.3 after having estimated the truncation error. One can only say here that the functioning of the IMEX-scheme is due to the fact that the investigated problem is very simple and specifically the anisotropy is aligned with the Cartesian mesh.

## 2.6 Numerical analysis

Let us now perform a numerical analysis study of our schemes introduced for the resolution of (2.12), permitting to understand in detail the behavior observed in the last section. In particular we shall detail only the error-analysis of the standard IMEX-scheme and the *Asymptotic-Preserving* Lagrange-multiplier scheme. The error study of the other schemes is very similar. See [75, 95] for more details on this analysis part.

### 2.6.1 IMEX scheme

We begin by recalling the full discretized form of the IMEX scheme for all  $(n, i, j) \in Q_h$  :

$$(IMEX)_\epsilon \quad \frac{f_{i,j}^{\epsilon,n+1} - f_{i,j}^{\epsilon,n}}{\Delta t} + a \frac{f_{i,j}^{\epsilon,n} - f_{i-1,j}^{\epsilon,n}}{\Delta x} + \frac{b}{\epsilon} \frac{f_{i,j}^{\epsilon,n+1} - f_{i,j-1}^{\epsilon,n+1}}{\Delta y} = 0. \quad (2.33)$$

**Theorem 2.6.1** *The IMEX scheme (2.33) is consistent with the Vlasov equation (2.12), and first order accurate in space and time. Furthermore, the local truncation error writes*

$$\mathcal{T}_I(t_n, x_i, y_j, \Delta t, \Delta x, \Delta y) = -\nabla \cdot (D_I \nabla f^\epsilon) + \mathcal{O}(\Delta t^2) + \mathcal{O}(\Delta x^2) + \mathcal{O}(\Delta y^2),$$

with

$$D_I := \begin{pmatrix} \frac{a\Delta x}{2}(1-\alpha) & 0 \\ 0 & \frac{b\Delta y}{2\epsilon}\left(1+\frac{\beta}{\epsilon}\right) \end{pmatrix}, \quad \alpha := \frac{a\Delta t}{\Delta x}, \quad \beta := \frac{b\Delta t}{\Delta y}.$$

Finally, we observe that the IMEX scheme (2.33) is a second-order scheme for the modified Vlasov equation

$$\partial_t g^\epsilon + a \partial_x g^\epsilon + \frac{b}{\epsilon} \partial_y g^\epsilon - \frac{a\Delta x}{2}(1-\alpha) \partial_{xx} g^\epsilon - \frac{b\Delta y}{2\epsilon}\left(1+\frac{\beta}{\epsilon}\right) \partial_{yy} g^\epsilon = 0. \quad (2.34)$$

**Proof 3** The local truncation error of the method (2.33) is defined by

$$\begin{aligned} & \mathcal{T}_I(t, x, y, \Delta t, \Delta x, \Delta y) \\ &= \frac{f^\epsilon(t + \Delta t, x, y) - f^\epsilon(t, x, y)}{\Delta t} + a \frac{f^\epsilon(t, x, y) - f^\epsilon(t, x - \Delta x, y)}{\Delta x} \\ &+ \frac{b}{\epsilon} \frac{f^\epsilon(t + \Delta t, x, y) - f^\epsilon(t + \Delta t, x, y - \Delta y)}{\Delta y}. \end{aligned}$$

Supposing that  $f^\epsilon$  is sufficiently smooth in order to apply a Taylor development, we find

$$\begin{aligned} \mathcal{T}_I(t_n, x_i, y_j, \Delta t, \Delta x, \Delta y) &= \partial_t f^\epsilon + \frac{\Delta t}{2} \partial_{tt} f^\epsilon + \frac{b\Delta t}{\epsilon} \partial_{yt} f^\epsilon + a \partial_x f^\epsilon - \frac{a\Delta x}{2} \partial_{xx} f^\epsilon \\ &+ \frac{b}{\epsilon} \partial_y f^\epsilon - \frac{b\Delta y}{2\epsilon} \partial_{yy} f^\epsilon + \mathcal{O}(\Delta t^2) + \mathcal{O}(\Delta x^2) + \mathcal{O}(\Delta y^2), \end{aligned}$$

where  $f^\epsilon$  is taken in  $(t_n, x_i, y_j)$ . Since  $f^\epsilon$  satisfies the Vlasov equation (2.12), the  $\mathcal{O}(1)$  terms drop out. Moreover, by differentiating the Vlasov equation along  $t$ ,  $y$  and  $x$ , we express the partial derivatives  $\partial_{tt} f$  and  $\partial_{ty} f$  as functions of  $\partial_{xx} f$  and  $\partial_{yy} f$ . We find thus

$$\partial_{tt} f^\epsilon = a^2 \partial_{xx} f^\epsilon + 2\frac{ab}{\epsilon} \partial_{xy} f^\epsilon + \frac{b^2}{\epsilon^2} \partial_{yy} f^\epsilon, \quad \partial_{ty} f^\epsilon = -a \partial_{xy} f^\epsilon - \frac{b}{\epsilon} \partial_{yy} f^\epsilon.$$

The local truncation error writes finally

$$\begin{aligned} \mathcal{T}_I(t_n, x_i, y_j, \Delta t, \Delta x, \Delta y) &= -\frac{a\Delta x}{2}(1-\alpha) \partial_{xx} f^\epsilon - \frac{b\Delta y}{2\epsilon}\left(1+\frac{\beta}{\epsilon}\right) \partial_{yy} f^\epsilon \\ &+ \mathcal{O}(\Delta t^2, \Delta x^2, \Delta y^2). \end{aligned}$$

**Remark 1** The modified equation (2.34) is an advection/diffusion equation. Note that the diffusion is stronger in the  $y$ -direction due to the term  $1/\epsilon$ . These diffusion terms are responsible for the damping that we observed in the numerical simulations (see Fig. 4.6 (B)), damping which tends towards infinity in the  $y$ -direction, as  $\epsilon \rightarrow 0$ . Note also that the diffusion coefficient is positive if  $\alpha \leq 1$ . This is precisely the stability condition of the upwind scheme, as we will see afterwards. If this condition is not respected, the diffusion becomes negative, leading to an ill-posed problem with exponentially growing solutions.

**Theorem 2.6.2** The IMEX scheme is stable in the Von Neumann sense if and only if the CFL-condition  $\frac{a\Delta t}{\Delta x} \leq 1$  is satisfied.



**Proof 4** To study the stability of our scheme, let us inject in (2.33) for fixed  $n \in \mathbb{N}$  a plane wave of the form

$$f_{i,j}^{\epsilon,n} = e^{ikx_i} e^{ily_j} \quad \forall (i,j),$$

with  $k, l \in \mathbb{Z}$  two arbitrary modes, and look how it evolves from one time-step to the other. Let us denote by  $\xi_I$  the amplification factor for this passage  $t_n \rightarrow t_{n+1}$ , meaning

$$f_{i,j}^{\epsilon,n+1} = \xi_I f_{i,j}^{\epsilon,n} = \xi_I e^{ikx_i} e^{ily_j}, \quad \forall (i,j).$$

Inserting now these terms in the discretized equation (2.33), yields, after simplification

$$\xi_I \left[ 1 + \frac{b\Delta t}{\epsilon\Delta y} (1 - e^{-il\Delta y}) \right] = \left[ 1 - \frac{a\Delta t}{\Delta x} (1 - e^{-ik\Delta x}) \right].$$

A scheme is said to be stable in the Von Neumann sense, if the amplification factor satisfies  $|\xi_I| \leq 1$ , such that the modes are not amplified from one time-step to the other. Straightforward computations yield

$$|\xi_I| = \epsilon \sqrt{\frac{1 - 4\alpha(1 - \alpha) \sin^2\left(\frac{k\Delta x}{2}\right)}{\epsilon^2 + 4\beta(\epsilon + \beta) \sin^2\left(\frac{l\Delta y}{2}\right)}}, \quad \forall k, l \in \mathbb{Z}.$$

Then, a necessary and sufficient condition to have the Von Neumann stability is :

$$\frac{a\Delta t}{\Delta x} \leq 1.$$

**Proof 5** Note that in the case  $l \neq 0$ , when  $\epsilon$  tends towards 0, the amplification factor converges towards 0. This means that for injected waves with mode  $l \neq 0$ , the scheme becomes more and more diffusive and attenuates completely the oscillations.

## 2.6.2 Lagrange-multiplier scheme

We do now the same work for the Lagrange-multiplier scheme, i.e.

$$(La)_\epsilon \begin{cases} \frac{f_{i,j}^{\epsilon,n+1} - f_{i,j}^{\epsilon,n}}{\Delta t} + a \frac{f_{i,j}^{\epsilon,n} - f_{i-1,j}^{\epsilon,n}}{\Delta x} + b \frac{q_{i,j}^{\epsilon,n+1} - q_{i,j-1}^{\epsilon,n+1}}{\Delta y} = 0, \quad \forall (n,i,j) \in Q_h \\ \frac{f_{i,j}^{\epsilon,n+1} - f_{i,j-1}^{\epsilon,n+1}}{\Delta y} = \epsilon \frac{q_{i,j}^{\epsilon,n+1} - q_{i,j-1}^{\epsilon,n+1}}{\Delta y}, \quad \forall (n,i,j) \in Q_h \\ q_{i,1}^{\epsilon,n} = 0, \quad \forall (n,i) \in [0, N_t] \times [0, N_x]. \end{cases} \quad (2.35)$$

**Theorem 2.6.3** The Lagrange-multiplier scheme (2.35) is consistent with the Vlasov equation (2.12), and first order accurate in space and time. Furthermore, the local truncation error writes

$$\mathcal{T}_L(t_n, x_i, y_j, \Delta t, \Delta x, \Delta y) = -\nabla \cdot (D_L \nabla f^\epsilon) + \mathcal{O}(\Delta t^2) + \mathcal{O}(\Delta x^2) + \mathcal{O}(\Delta y^2),$$

with

$$D_L := \begin{pmatrix} \frac{a\Delta x}{2}(1 - \alpha) & 0 \\ 0 & \frac{b\Delta y}{2\epsilon} \left(1 + \frac{\beta}{\epsilon}\right) \end{pmatrix}, \quad \alpha := \frac{a\Delta t}{\Delta x}, \quad \beta := \frac{b\Delta t}{\Delta y}.$$

**Proof 6** In order to prove the result, we write the local truncation error of the first equation. We find that

$$\begin{aligned} \mathcal{T}_L(t_n, x_i, y_j, \Delta t, \Delta x, \Delta y) &= \frac{\Delta t}{2} \partial_{tt} f^\epsilon - a \frac{\Delta x}{2} \partial_{xx} f^\epsilon - b \frac{\Delta y}{2} \partial_{yy} q^\epsilon + b \Delta t \partial_{yt} q^\epsilon \\ &\quad + \mathcal{O}(\Delta t^2) + \mathcal{O}(\Delta x^2) + \mathcal{O}(\Delta y^2). \end{aligned}$$

Since the first equation of (2.35) is verified by  $(f^\epsilon, q^\epsilon)$ , we have

$$\partial_{tt} f^\epsilon = -a \partial_{xt} f^\epsilon - b \partial_{yt} q^\epsilon, \quad \partial_{xt} f^\epsilon = -a \partial_{xx} f^\epsilon - b \partial_{xy} q^\epsilon, \quad \partial_{ty} f^\epsilon = -a \partial_{xy} f^\epsilon - b \partial_{yy} q^\epsilon.$$

Then,

$$\begin{aligned} \mathcal{T}_L(t_n, x_i, y_j, \Delta t, \Delta x, \Delta y) &= \frac{\Delta t}{2} (a^2 \partial_{xx} f^\epsilon + ab \partial_{xy} q^\epsilon) - a \frac{\Delta x}{2} \partial_{xx} f^\epsilon - b \frac{\Delta y}{2} \partial_{yy} q^\epsilon \\ &\quad + b \frac{\Delta t}{2} \partial_{ty} q^\epsilon + \mathcal{O}(\Delta t^2) + \mathcal{O}(\Delta x^2) + \mathcal{O}(\Delta y^2). \end{aligned}$$

Since the second equation of (2.35) is verified, we have

$$\partial_{ty} q^\epsilon = \frac{1}{\epsilon} \partial_{ty} f^\epsilon, \quad \partial_{yy} q^\epsilon = \frac{1}{\epsilon} \partial_{yy} f^\epsilon, \quad \partial_{xy} q^\epsilon = \frac{1}{\epsilon} \partial_{xy} f^\epsilon,$$

such that we find the same expression as for the IMEX scheme, i.e.

$$\mathcal{T}_L(t_n, x_i, y_j, \Delta t, \Delta x, \Delta y) = -\nabla \cdot (D_L \nabla f^\epsilon) + \mathcal{O}(\Delta t^2) + \mathcal{O}(\Delta x^2) + \mathcal{O}(\Delta y^2).$$

The just proved result confirms what we have seen on the numerical plots. Indeed, the IMEX and Lagrange-multiplier schemes have the same behavior when regarding the convergence and the asymptotic behavior.

**Theorem 2.6.4** *The Lagrange-multiplier scheme is stable in the Von Neumann sense if and only if the CFL condition  $\frac{a\Delta t}{\Delta x} \leq 1$  is satisfied.*

**Proof 7** Here, we have two unknown functions  $f^\epsilon$  and  $q^\epsilon$ . To study the Von Neumann stability, we write

$$q_{i,j}^{\epsilon,n+1} = \xi_q q_{i,j}^{\epsilon,n}, \quad f_{i,j}^{\epsilon,n+1} = \xi_f f_{i,j}^{\epsilon,n},$$

with the two amplification factors  $\xi_q$  and  $\xi_f$ . As usual, we insert these expressions in the discretized Lagrange-multiplier equations. We obtain a linear system where the unknowns are  $\xi_q$  and  $\xi_f$ . This system writes

$$\begin{pmatrix} 1 & \beta(1 - e^{-ik_m \Delta y}) \\ 1 & -\epsilon \end{pmatrix} \begin{pmatrix} \xi_f \\ \xi_q \end{pmatrix} = \begin{pmatrix} \alpha(1 - e^{-ik_n \Delta x}) \\ 0 \end{pmatrix},$$

and is easy to invert. Computing the amplification factor  $\xi_f$ , we remark that it is identical to the one calculated for the IMEX scheme.

### 2.6.3 AP-properties

We are now able to explain the numerical results obtained in Section 2.5, in particular to explain why the IMEX-scheme, even if being not an AP-scheme, gives in this simple field-aligned test case, good results up to a value of  $\epsilon = 10^{-14}$ . For this, let us recall that two types of errors arise during a numerical resolution of the Vlasov equation (2.12). First of all we have the truncation errors, estimated in the last subsections, and secondly one has also the round-off errors, arising at each elementary computation. To be more precise, one has to consider the three linear systems, corresponding to (2.32) :

$$\mathcal{A} \mathcal{F}_{ex} = \mathcal{B} + \Delta t \mathcal{T}, \quad \mathcal{A} \mathcal{F} = \mathcal{B}, \quad (\mathcal{A} + \delta \mathcal{A}) \mathcal{F}_{num} = \mathcal{B} + \delta \mathcal{B},$$

where to simplify notation we omitted all the time and space indices. Here we denoted by  $\mathcal{F}_{ex}$  the exact solution of the Vlasov equation (2.12), satisfying the linear system (2.32) up to a truncation error  $\mathcal{T}$ ,  $\mathcal{F}$  is the exact solution of the linear system (2.32), supposing exact arithmetics, and finally  $\mathcal{F}_{num}$  is the solution to the linear system (2.32) obtained via a computer, hence contaminated with round-off errors. The error we are interested in, can be estimated as follows

$$\|\mathcal{F}_{ex} - \mathcal{F}_{num}\| \leq \|\mathcal{F}_{ex} - \mathcal{F}\| + \|\mathcal{F} - \mathcal{F}_{num}\|.$$

Stability and consistency permit to show that the first error term is of the order of the truncation error. For the estimate of the second error term, we have to take into account the condition number of the matrix, in particular one has the estimate [126]

$$\frac{\|\mathcal{F} - \mathcal{F}_{num}\|}{\|\mathcal{F}\|} \leq \frac{\text{cond}(\mathcal{A})}{1 - \|\mathcal{A}^{-1}\| \|\delta \mathcal{A}\|} \left( \frac{\|\delta \mathcal{A}\|}{\|\mathcal{A}\|} + \frac{\|\delta \mathcal{B}\|}{\|\mathcal{B}\|} \right).$$

Performing our computations in double precision (machine accuracy of  $10^{-16}$ ), and as long as the condition number is not exceeding a value of  $10^{12}$  (see Fig. 2.9), the second error term is not so dangerous. For larger condition numbers, this term can give rise to erroneous results. In our test case, it is however rather the first error-term which leads to trouble, as the truncation error is  $1/\epsilon$ -dependent. In the first toy-model (2.12), the large truncation error impacts only the  $y$ -direction, leading to a large diffusion along the axes-aligned anisotropy and hence to the limit-model. We shall see a drastic difference in the second, not-field aligned toy-model.

## 2.7 Second Vlasov toy-model with variable coefficients

Finally, let us come now in this section to the second Vlasov toy model, given by :

$$\partial_t f^\epsilon + \frac{1}{\epsilon} (\mathbf{v} \times \mathbf{B}) \cdot \nabla_v f^\epsilon = 0, \quad (2.36)$$

with  $\epsilon \ll 1$  and the magnetic field  $\mathbf{B} = \mathbf{e}_z$ . This model is a simplified version of the anisotropic Vlasov equation (2.4) in not-field aligned Cartesian coordinates. Denoting, for notational simplicity, the velocity-variable as  $\mathbf{v} = (x, y, z)$ , we have  $\mathbf{v} \times \mathbf{B} = (y, -x, 0)^T$ , such that the previous equation writes :

$$(G)_\epsilon \begin{cases} \partial_t f^\epsilon + \frac{y}{\epsilon} \partial_x f^\epsilon - \frac{x}{\epsilon} \partial_y f^\epsilon = 0, & \forall (t, x, y) \in [0, T] \times \Omega. \\ f^\epsilon(0, x, y) = f_{in}(x, y), & \forall (x, y) \in \Omega, \end{cases} \quad (2.37)$$

where this time our velocity-domain is given by  $\Omega := [-L_x, L_x] \times [-L_y, L_y]$ . Again we will consider a doubly-periodic framework.

### 2.7.1 Exact solution by the characteristic method

The exact solution of the equation (2.37) is simply determined via the characteristic method. The characteristic curve  $C_\epsilon^{x,y}(s) := (X(s), Y(s))$  passing at instant  $t$  through  $(x, y)$ , solves the ODE :

$$\begin{cases} \dot{X}(s) = \frac{Y(s)}{\epsilon}, \\ \dot{Y}(s) = -\frac{X(s)}{\epsilon}, \end{cases} \quad (X(t), Y(t)) = (x, y).$$

We can write this system under matrix form :

$$\begin{pmatrix} \dot{X} \\ \dot{Y} \end{pmatrix} = \frac{1}{\epsilon} A \begin{pmatrix} X \\ Y \end{pmatrix}, \quad \text{with } A := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

leading to

$$C_\epsilon^{x,y}(s) := \begin{pmatrix} X \\ Y \end{pmatrix}(s) = e^{A \frac{s-t}{\epsilon}} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Denoting the rotation matrix by  $\mathcal{R}_\epsilon(y) := e^{A \frac{y}{\epsilon}}$ , one has

$$\mathcal{R}_\epsilon(s-t) = e^{A \frac{s-t}{\epsilon}} = \begin{pmatrix} \cos\left(\frac{s-t}{\epsilon}\right) & \sin\left(\frac{s-t}{\epsilon}\right) \\ -\sin\left(\frac{s-t}{\epsilon}\right) & \cos\left(\frac{s-t}{\epsilon}\right) \end{pmatrix}.$$

We can easily verify that the characteristic curves are nothing else than spirals in the  $(t, x, y)$ -phase-space. All characteristics are  $2\pi\epsilon$ -periodic (in  $s$ ).

The exact solution  $f^\epsilon$  of (2.37) is now simply the advection of the initial condition along these characteristic curves, such that

$$\begin{aligned} f^\epsilon(t, x, y) &= f_{in}(X(0, t, x, y), Y(0, t, x, y)) \\ &= f_{in}\left(\cos\left(\frac{t}{\epsilon}\right)x - \sin\left(\frac{t}{\epsilon}\right)y, \sin\left(\frac{t}{\epsilon}\right)x + \cos\left(\frac{t}{\epsilon}\right)y\right). \end{aligned}$$

### 2.7.2 Limit solution of the problem

The next step is to obtain the limit solution of the problem (2.37), as  $\epsilon \rightarrow 0$ . Keeping in mind that  $f^\epsilon$  is constant along the characteristic curves, we integrate (2.37) along  $C_\epsilon^{x,y}$ , to get

$$\partial_t \int_{C_\epsilon^{x,y}} f^\epsilon d\sigma + \frac{1}{\epsilon} \int_{C_\epsilon^{x,y}} (y, -x)^T \cdot \nabla f^\epsilon(t, x, y) d\sigma = 0,$$

leading to

$$\partial_t \int_{C_\epsilon^{x,y}} f^\epsilon d\sigma + \frac{1}{\epsilon} \int_t^{t+2\pi\epsilon} (Y(s), -X(s))^T \cdot \nabla f^\epsilon(t, X(s), Y(s)) \frac{\sqrt{x^2 + y^2}}{\epsilon} ds = 0.$$

Furthermore, as

$$\int_t^{t+2\pi\epsilon} (Y(s), -X(s))^T \cdot \nabla f^\epsilon(t, X(s), Y(s)) ds = \int_t^{t+2\pi\epsilon} \frac{d}{ds} \left[ f^\epsilon \left( t, X(s), Y(s) \right) \right] ds = 0,$$

which comes from the periodicity of the characteristics, and denoting the average along a curve by  $\langle f^\epsilon \rangle := \frac{1}{|C_\epsilon^{x,y}|} \int_{C_\epsilon^{x,y}} f^\epsilon d\sigma$ , with  $|C_\epsilon^{x,y}| = 2\pi\epsilon$ , we have :

$$\partial_t \langle f^\epsilon \rangle = 0.$$

Letting now formally  $\epsilon \rightarrow 0$ , we obtain the following limit problem associated to (2.37) :

$$(G)_0 \quad \langle f^0 \rangle = \langle f_{in} \rangle. \quad (2.38)$$

### 2.7.3 Numerical schemes for the second Vlasov toy model

Let us now discretize the second Vlasov toy model (2.37) via the IMP (fully implicit scheme this time) and Lagrange-multiplier schemes. The time semi-discretizations read

$$(IMP)_\epsilon \quad \frac{f^{\epsilon,n+1} - f^{\epsilon,n}}{\Delta t} + \frac{y}{\epsilon} \partial_x f^{\epsilon,n+1} - \frac{x}{\epsilon} \partial_y f^{\epsilon,n+1} = 0, \quad \forall n \geq 0, \quad \forall (x,y) \in \Omega, \quad (2.39)$$

as well as

$$(La)_\epsilon \quad \begin{cases} \frac{f^{\epsilon,n+1} - f^{\epsilon,n}}{\Delta t} + y \partial_x q^{\epsilon,n+1} - x \partial_y q^{\epsilon,n+1} = 0, \\ \forall n \geq 0, \\ -y \partial_x f^{\epsilon,n+1} + x \partial_y f^{\epsilon,n+1} - \epsilon \left( -y \partial_x q^{\epsilon,n+1} + x \partial_y q^{\epsilon,n+1} \right) - (\Delta x \Delta y)^\gamma q^{\epsilon,n+1} = 0. \end{cases} \quad (2.40)$$

The term  $(\Delta x \Delta y)^\gamma q^{\epsilon,n}$  in (2.40) is a stabilization term permitting to have the uniqueness of the solution  $(f^\epsilon, q^\epsilon)$ . In the former ‘‘field-aligned’’ example, we fixed  $q^\epsilon$  on the anisotropy lines by setting  $q^\epsilon|_{\Gamma_{in}} = 0$ , but here it is more arduous from a numerical point of view. The stabilization aims equally to fix  $q^\epsilon$ , however in a different manner. It is very delicate to choose the magnitude of this term, in order not to destroy the problem, in particular we took here  $\gamma = 0.91$ . First it is a small perturbation of the equation, of the order of the truncation error. Secondly, averaging the second equation of the Lagrange-multiplier scheme along the anisotropy lines, permits to obtain

$$(\Delta x \Delta y)^\gamma \langle q^{\epsilon,n+1} \rangle = 0,$$

which means that  $q^\epsilon$  is unique, by having zero average along the field lines. A more detailed study of this stabilization technique was performed in [98] for the elliptic framework.

For the spatial discretization, we use again an upwind scheme, observing that this time the equation has no more constant coefficients. Thus, we define for all  $(i,j) \in \mathbb{N}^2$  :

$$x_i^+ := \max(x_i, 0), \quad x_i^- := \min(0, x_i), \quad y_j^+ := \max(y_j, 0), \quad y_j^- := \min(0, y_j).$$

The full discretization of the IMP scheme writes now

$$(IMP)_\epsilon \quad f_{i,j}^{\epsilon,n+1} + \frac{1}{\epsilon} \left[ \left( r_x (y_j^+ - y_j^-) + r_y (x_i^+ - x_i^-) \right) f_{i,j}^{\epsilon,n+1} - r_x (y_j^+ f_{i-1,j}^{\epsilon,n+1} - y_j^- f_{i+1,j}^{\epsilon,n+1}) \right. \\ \left. - r_y (x_i^+ f_{i,j+1}^{\epsilon,n+1} - x_i^- f_{i,j-1}^{\epsilon,n+1}) \right] = f_{i,j}^{\epsilon,n}, \quad \forall (n,i,j) \in Q_h,$$

with  $r_x = \frac{\Delta t}{\Delta x}$  and  $r_y = \frac{\Delta t}{\Delta y}$ . And for the Lagrange-multiplier scheme, we have for all  $(n,i,j) \in Q_h$  :

$$(La)_\epsilon \left\{ \begin{array}{l} f_{i,j}^{\epsilon,n+1} + \left[ (r_x(y_j^+ - y_j^-) + r_y(x_i^+ - x_i^-)) q_{i,j}^{\epsilon,n+1} - r_x(y_j^+ q_{i-1,j}^{\epsilon,n+1} - y_j^- q_{i+1,j}^{\epsilon,n+1}) \right. \\ \left. - r_y(x_i^+ q_{i,j+1}^{\epsilon,n+1} - x_i^- q_{i,j-1}^{\epsilon,n+1}) \right] = f_{i,j}^{\epsilon,n}, \quad \forall (n, i, j) \in Q_h, \\ \\ \frac{1}{\Delta t} \left[ (r_x(y_j^+ - y_j^-) + r_y(x_i^+ - x_i^-)) f_{i,j}^{\epsilon,n+1} - r_x(y_j^+ f_{i+1,j}^{\epsilon,n+1} - y_j^- f_{i-1,j}^{\epsilon,n+1}) - \right. \\ \left. r_y(x_i^+ f_{i,j+1}^{\epsilon,n+1} - x_i^- f_{i,j-1}^{\epsilon,n+1}) \right] - \frac{\epsilon}{\Delta t} \left[ (r_x(y_j^+ - y_j^-) + r_y(x_i^+ - x_i^-)) q_{i,j}^{\epsilon,n+1} - \right. \\ \left. r_x(y_j^+ q_{i+1,j}^{\epsilon,n+1} - y_j^- q_{i-1,j}^{\epsilon,n+1}) - r_y(x_i^+ q_{i,j+1}^{\epsilon,n+1} - x_i^- q_{i,j-1}^{\epsilon,n+1}) \right] - (\Delta x \Delta y)^\gamma q_{i,j}^{\epsilon,n+1} = 0. \end{array} \right.$$

## 2.7.4 Numerical simulations

Here we present our simulations corresponding to both numerical schemes. We consider  $\Omega = [-3, 3]^2$ ,  $T = 1$  and the discretization parameters  $N_t = 64$  and  $N_x = N_y = 160$ . The initial data is defined by a Gaussian function :

$$f_{in}(x, y) = \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right), \quad \sigma = 0.5, \quad \forall (x, y) \in \Omega.$$

As we showed before, the exact solution is known thanks to the characteristic method. In the present simple test case, one can easily prove that

$$f_{ex}^\epsilon(t, x, y) = f_{in}(x, y) = \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right), \quad (2.41)$$

in other words, the exact solution is a stationary solution, independent of  $\epsilon$ , the initial condition being constant along the anisotropy field. This simple test case permits in a very simple way to compare both methods with respect to the  $\epsilon$ -dependence of the results, in particular to show that the IMP-scheme is not an *Asymptotic-Preserving* scheme. We shall investigate in the Chapter 3 a more involved, physical test-case, where we shall adapt the here introduced Lagrange-multiplier-method, which seems to be the most appropriate method for our singularly-perturbed Vlasov problem (2.4), to second-order schemes and test more thoroughly its AP-properties.

In Figure 2.10 we first plot the condition-number  $\text{cond}(\mathcal{A}) := \|\mathcal{A}^{-1}\|_2 \|\mathcal{A}\|_2$  associated to the two schemes. As for the first toy-model, one remarks the  $\epsilon$ -independent condition-number of the Lagrange-multiplier-scheme, whereas, as expected, the IMP scheme has an  $1/\epsilon$ -dependent condition-number.

Then, in Figure 4.15, we show the numerical solution  $f^\epsilon$  at the final time  $T$  and computed for several values of  $\epsilon$ , with both IMP and Lagrange-multiplier schemes. For  $\epsilon = 1$ , we do not distinguish any difference. However for smaller  $\epsilon$  values, the solution obtained with the Lagrange-multiplier scheme seems to be  $\epsilon$ -independent, contrary to the IMP scheme, which diffuses more and more as  $\epsilon \rightarrow 0$ . Indeed, the IMP solution is completely damped as  $\epsilon \rightarrow 0$  and leads towards the zero-solution, whereas the Lagrange-multiplier scheme keeps the form of the Gaussian, with a usual  $\epsilon$ -independent  $(\Delta x, \Delta y)$ -diffusion. This permits to conclude that the Lagrange-multiplier scheme is an AP-scheme contrary to the IMP scheme.

In order to distinguish much better this AP-property, we plot on Figure 2.11 a cut of the previous curves at the point  $x = 0$ . We observe clearly the diffusion in the IMP scheme which depends of  $\epsilon$  contrary to the Lagrange-multiplier scheme.

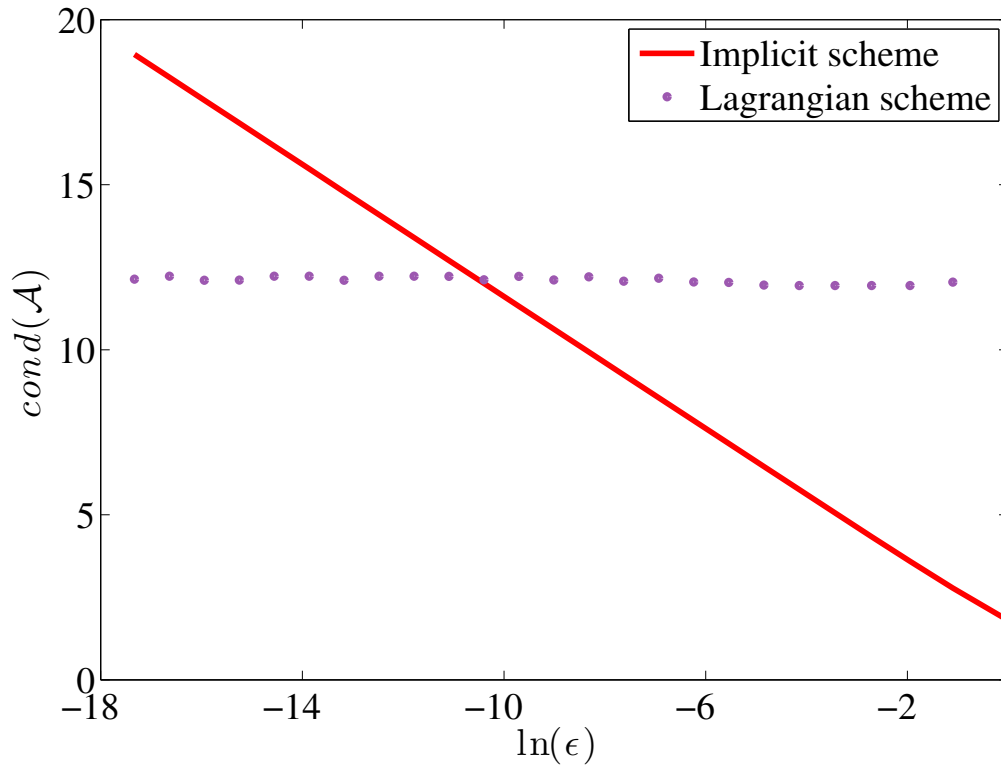


FIGURE 2.10 – Condition number  $\text{cond}(A)$  as a function of  $\epsilon$  in log-log scale. The two curves correspond to the IMP and Lagrange-multiplier schemes.

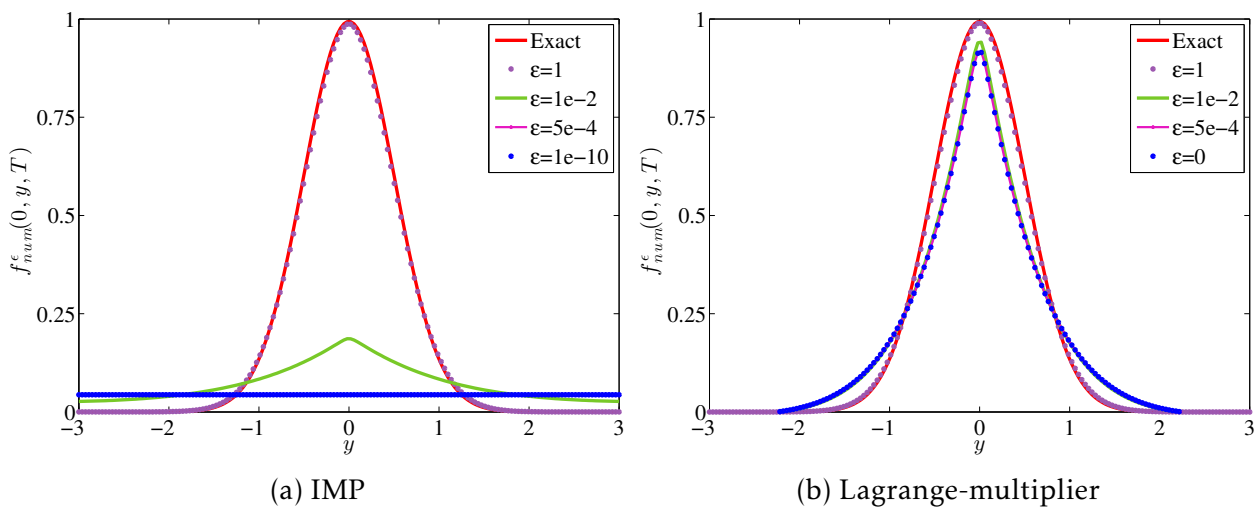


FIGURE 2.11 – Representation of a cut at  $x = 0$  of  $f_{num}^\epsilon$  at the final time  $T$  for the IMP and Lagrange-multiplier schemes, and several values of  $\epsilon$ .

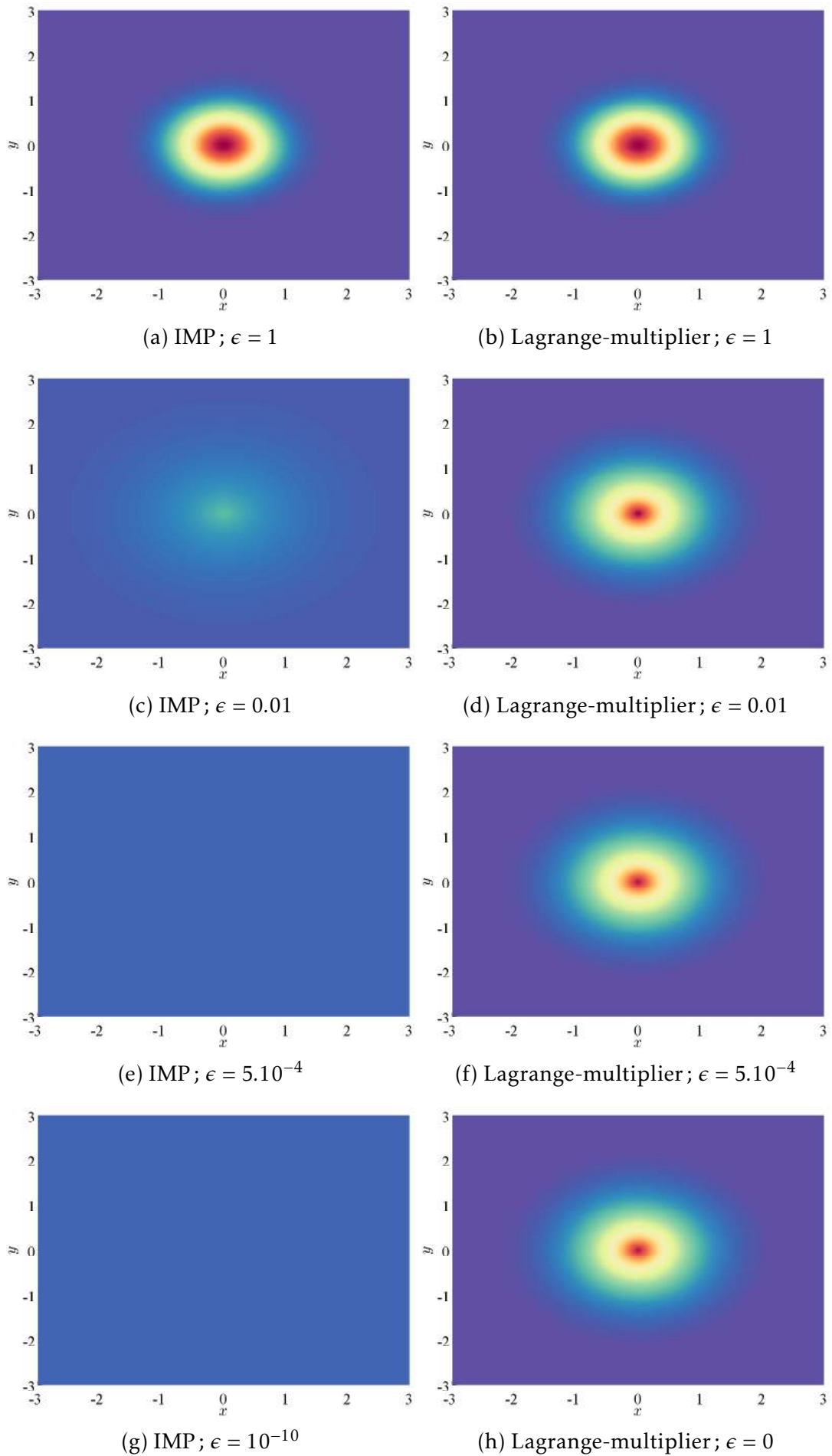


FIGURE 2.12 – Representation of the function  $f_9^\epsilon$  at the final time  $T$  for the IMP and Lagrange-multiplier scheme, with several values of  $\epsilon$ .



## 2.7.5 Numerical analysis

The aim of this section is to explain the plots presented before. In particular we will investigate why the IMP scheme does not work for small  $\epsilon$ -values, whereas the Lagrange-multiplier scheme preserves the asymptotics. First of all, we compute the local truncation error of both schemes. We shall consider only the case  $x \geq 0$  and  $y \geq 0$ , the remaining cases changing nothing in the following reasoning.

### IMP scheme

We begin by recalling the expression of the full discretized expression of this scheme for all  $(n, i, j) \in Q_h$  :

$$(IMP)_\epsilon \frac{f_{i,j}^{\epsilon,n+1} - f_{i,j}^{\epsilon,n}}{\Delta t} + \frac{y_j}{\epsilon} \frac{f_{i,j}^{\epsilon,n+1} - f_{i-1,j}^{\epsilon,n+1}}{\Delta x} - \frac{x_i}{\epsilon} \frac{f_{i,j+1}^{\epsilon,n+1} - f_{i,j}^{\epsilon,n+1}}{\Delta y} = 0. \quad (2.42)$$

**Theorem 2.7.1** *The IMP scheme (2.42) is consistent with the second Vlasov problem (2.36), first order accurate in time and in space. Moreover, the local truncation error writes*

$$\mathcal{T}_I(t_n, x_i, y_j, \Delta t, \Delta x, \Delta y) = -\nabla \cdot \left[ D_I \nabla f^\epsilon \right] + \mathcal{O}(\Delta t^2) + \mathcal{O}(\Delta x^2) + \mathcal{O}(\Delta y^2).$$

where

$$D_I := \frac{1}{\epsilon} \begin{pmatrix} \frac{y_j \Delta x}{2} \left( 1 + \frac{\alpha_j}{\epsilon} \right) & \frac{-x_i y_j \Delta t}{2\epsilon} \\ \frac{-x_i y_j \Delta t}{2\epsilon} & \frac{x_i \Delta y}{2} \left( 1 + \frac{\beta_i}{\epsilon} \right) \end{pmatrix}, \quad \alpha_j := \frac{y_j \Delta t}{\Delta x}, \quad \beta_i := \frac{x_i \Delta t}{\Delta y}.$$

**Proof 8** *This proof is very similar to the proof of Theorem 2.6.1.*

**Remark 2** *Contrary to the first toy-model, where the diffusion was  $1/\epsilon$ -dependent only in the anisotropy-direction, which was aligned with the coordinate system, in the present case, the diffusion-matrix is scaled by a  $1/\epsilon$  factor, meaning that this time we have a very strong  $1/\epsilon$ -dependent diffusion in all directions. This large diffusion leads rapidly to a damping of the solution towards zero, as  $\epsilon$  becomes smaller, and leads thus to completely erroneous results.*

### Lagrange-multiplier scheme

We use the same reasoning for the Lagrange-multiplier scheme

$$(La)_\epsilon \begin{cases} \partial_t f^\epsilon + y \partial_x q^\epsilon - x \partial_y q^\epsilon = 0, \\ -y \partial_x f^\epsilon + x \partial_y f^\epsilon - \epsilon \left( -y \partial_x q^\epsilon + x \partial_y q^\epsilon \right) - (\Delta x \Delta y)^\nu q^\epsilon = 0. \end{cases} \quad (2.43)$$

Supposing  $y \geq 0$  and  $x \geq 0$ , we have the full discretization of  $(La)_\epsilon$

$$(La)_\epsilon \begin{cases} \frac{f_{i,j}^{\epsilon,n+1} - f_{i,j}^{\epsilon,n}}{\Delta t} + y_j \frac{q_{i,j}^{\epsilon,n+1} - q_{i-1,j}^{\epsilon,n+1}}{\Delta x} - x_i \frac{q_{i,j+1}^{\epsilon,n+1} - q_{i,j}^{\epsilon,n+1}}{\Delta y} = 0, \\ -y_j \frac{f_{i+1,j}^{\epsilon,n+1} - f_{i,j}^{\epsilon,n+1}}{\Delta x} + x_i \frac{f_{i,j}^{\epsilon,n+1} - f_{i,j-1}^{\epsilon,n+1}}{\Delta y} - \epsilon \left( -y_j \frac{q_{i+1,j}^{\epsilon,n+1} - q_{i,j}^{\epsilon,n+1}}{\Delta x} \right. \\ \left. + x_i \frac{q_{i,j}^{\epsilon,n+1} - q_{i,j-1}^{\epsilon,n+1}}{\Delta y} \right) - (\Delta x \Delta y)^\nu q_{i,j}^{\epsilon,n+1} = 0. \end{cases} \quad (2.44)$$

**Theorem 2.7.2** *The Lagrange-multiplier scheme (2.44) is consistent with the second Vlasov model (2.36) and first order accurate in time and in space. Furthermore, the local truncation error writes*

$$\begin{aligned} \begin{pmatrix} \mathcal{T}_{L1} \\ \mathcal{T}_{L2} \end{pmatrix} &= - \begin{pmatrix} \nabla \cdot & 0 \\ 0 & \nabla \cdot \end{pmatrix} \begin{pmatrix} 0 & D_{L1} \\ D_{L2} & -\epsilon D_{L2} \end{pmatrix} \begin{pmatrix} \nabla f^\epsilon \\ \nabla q^\epsilon \end{pmatrix} + \mathcal{O}(\Delta t^2, \Delta x^2, \Delta y^2) \\ &= - \begin{pmatrix} \nabla \cdot (D_{L1} \nabla q^\epsilon) \\ \nabla \cdot (D_{L2} \nabla f^\epsilon) - \epsilon \nabla \cdot (D_{L2} \nabla q^\epsilon) \end{pmatrix} + \mathcal{O}(\Delta t^2, \Delta x^2, \Delta y^2) \end{aligned}$$

where

$$D_{L1} := \begin{pmatrix} \frac{y_j \Delta x}{2} \left(1 + \frac{\alpha_j}{\epsilon}\right) & \frac{-x_i y_j \Delta t}{2\epsilon} \\ \frac{-x_i y_j \Delta t}{2\epsilon} & \frac{x_i \Delta y}{2} \left(1 + \frac{\beta_i}{\epsilon}\right) \end{pmatrix}, \quad D_{L2} = \begin{pmatrix} \frac{y_j \Delta x}{2} & 0 \\ 0 & \frac{x_i \Delta y}{2} \end{pmatrix}.$$

**Proof 9** *We begin by the computation of the  $\mathcal{T}_{L1}$  term. Supposing sufficient regularity for the functions  $f^\epsilon$  and  $q^\epsilon$ , we use Taylor series expansion to get*

$$\begin{aligned} \mathcal{T}_{L1}(t_n, x_i, y_j, \Delta t, \Delta x, \Delta y) &= \frac{\Delta t}{2} \partial_{tt} f^\epsilon + y_j \Delta t \partial_{xt} q^\epsilon - \frac{y_j \Delta x}{2} \partial_{xx} q^\epsilon - \frac{x_i \Delta y}{2} \partial_{yy} q^\epsilon \\ &\quad - x_i \Delta t \partial_{ty} q^\epsilon + \mathcal{O}(\Delta t^2) + \mathcal{O}(\Delta x^2) + \mathcal{O}(\Delta y^2). \end{aligned}$$

Since the first equation (2.43) is verified, we can write

$$\partial_{tt} f^\epsilon = -y \partial_{xt} q^\epsilon + x \partial_{yt} q^\epsilon.$$

Differentiating in time the second equation of (2.43) we obtain

$$\begin{aligned} \mathcal{T}_{L1}(t_n, x_i, y_j, \Delta t, \Delta x, \Delta y) &= \frac{\Delta t}{2\epsilon} \partial_t (y_j \partial_x f^\epsilon - x_i \partial_y f^\epsilon) - \frac{y_j \Delta x}{2} \partial_{xx} q^\epsilon - \frac{x_i \Delta y}{2} \partial_{yy} q^\epsilon \\ &\quad + \mathcal{O}(\Delta t^2) + \mathcal{O}(\Delta x^2) + \mathcal{O}(\Delta y^2). \end{aligned}$$

As

$$\partial_{tx} f^\epsilon = x \partial_{yx} q^\epsilon - y \partial_{xx} q^\epsilon + \partial_y q^\epsilon, \quad \partial_{ty} f^\epsilon = x \partial_{yy} q^\epsilon - \partial_x q^\epsilon - y \partial_{xy} q^\epsilon,$$

the local truncation error writes finally

$$\begin{aligned} \mathcal{T}_{L1}(t_n, x_i, y_j, \Delta t, \Delta x, \Delta y) &= \frac{\Delta t}{2\epsilon} (y_j \partial_y q^\epsilon + x_i \partial_x q^\epsilon + 2x_i y_j \partial_{xy} q^\epsilon) \\ &\quad - \frac{y_j \Delta x}{2} \left(1 + \frac{\alpha_j}{\epsilon}\right) \partial_{xx} q^\epsilon - \frac{x_i \Delta y}{2} \left(1 + \frac{\beta_i}{\epsilon}\right) \partial_{yy} q^\epsilon \\ &\quad + \mathcal{O}(\Delta t^2) + \mathcal{O}(\Delta x^2) + \mathcal{O}(\Delta y^2). \end{aligned}$$

With an analogous reasoning, we compute the truncation error of the second equation :

$$\begin{aligned} \mathcal{T}_{L2}(t_n, x_i, y_j, \Delta t, \Delta x, \Delta y) &= - \frac{y_j \Delta x}{2} \partial_{xx} f^\epsilon - \frac{x_i \Delta y}{2} \partial_{yy} f^\epsilon + \epsilon \left( \frac{y_j \Delta x}{2} \partial_{xx} q^\epsilon \right. \\ &\quad \left. + \frac{x_i \Delta y}{2} \partial_{yy} q^\epsilon \right) + \mathcal{O}(\Delta t^2) + \mathcal{O}(\Delta x^2) + \mathcal{O}(\Delta y^2). \end{aligned}$$

**Remark 3** *In contrast to the first Vlasov toy-model (2.12), the IMP and Lagrange-multiplier schemes do not have the same behavior with respect to the local truncation error. More particularly, the dependence on  $\epsilon$  is very different. The IMP-scheme is diffusing in all directions, diffusion proportional to  $1/\epsilon$ . The only  $1/\epsilon$ -dependent diffusion in the Lagrange-multiplier scheme arises in relation with the auxiliary unknown  $q^\epsilon$ , i.e. in the term  $\nabla \cdot (D_{L_1} \nabla q^\epsilon)$ . And one can immediately verify that the  $1/\epsilon$ -dependence arises only aligned with the anisotropy field lines, and not perpendicular to them. Indeed, one gets immediately for the diffusion along resp. perpendicular to the field lines :*

$$\begin{aligned} (y, -x) D_{L_1} (y, -x)^T &= \frac{y^3 \Delta x}{2} + \frac{x^3 \Delta y}{2} + \frac{\Delta t}{2\epsilon} (x^2 + y^2)^2 \\ (x, y) D_{L_1} (x, y)^T &= \frac{xy}{2} [x \Delta x + y \Delta y]. \end{aligned}$$

## 2.8 Concluding remarks

To conclude, let us summarize here the knowledge we acquired about the resolution of anisotropic Vlasov equations of the type (2.4) arising in fusion plasma modelisation. Two types of techniques can be adopted from the beginning. One can decide to pass directly to polar coordinates in velocity and get hence a field-aligned formulation as for ex. (2.8). In this case, a simple IMEX-scheme is the most appropriate scheme to be used, being simple enough and giving rise to accurate results up a sufficiently small  $\epsilon$ -value. However, the disadvantage is that one has to change coordinate system, which can be rather cumbersome if the magnetic field is variable, in time and space.

The second technique is rather simple, as it avoids to pass to field-aligned coordinates and remains in a nice Cartesian framework. The drawback is that in this case it is no more sufficient to implicit the stiff term and take the other terms explicitly. Indeed, for small  $\epsilon$ -values (already  $\epsilon = 10^{-4}$ ), meaning strong magnetic fields as in tokamak plasmas, an IMEX scheme would lead to erroneous results. An Asymptotic-Preserving reformulation like our "Lagrange-multiplier-method" is more adequate, leading in the limit  $\epsilon \rightarrow 0$  towards the right Limit-problem. This Lagrange-multiplier-method is indeed usable for all  $\epsilon \geq 0$  and gives accurate and stable results independently on  $\epsilon$ . However there is a disadvantage, namely the fact that it is more time-consuming, as it involves an additional unknown  $q^\epsilon$ .

Solving an anisotropic Vlasov equation of the type (2.4) needs hence an *a priori* decision, which of these two techniques to follow. The first technique is at the moment the basis of several codes. The second technique has not be tested up to now, and its rigorous validation and comparison with the first one will be the aim of a forthcoming paper, in a more physical context.



---

---

## CHAPITRE 3

---

# ASYMPTOTIC-PRESERVING SCHEME FOR THE RESOLUTION OF EVOLUTION EQUATIONS WITH STIFF TRANSPORT TERMS

Le contenu de ce chapitre correspond à un article écrit avec C. Negulescu et S. Possanner. Il a donné lieu à la publication suivante :

B. Fedele, C. Negulescu and S. Possanner, **Asymptotic-Preserving scheme for the resolution of evolution equations with stiff transport terms.**, *SIAM, Multiscale Model. Simul.*, 17 (2019), 307–343.

### Abstract.

Dans ce chapitre, on développe une méthode préservant l'asymptotique afin de résoudre des équations avec un terme de transport raide. Ce schéma numérique est basé sur une décomposition Micro-Macro de l'inconnue, couplé avec une procédure de stabilisation. Cette méthode numérique sera appliquée à une équation Vlasov simplifiée dans le régime gyrocinétique et au modèle de Vlasov-Poisson  $1D1V$ . Ces modèles interviennent en physique des plasmas. Les propriétés préservant l'asymptotique permettront d'étudier le comportement en temps longs de ces modèles. En effet, les approches numériques classiques entraînent une accumulation d'erreur importante sur de telles asymptotiques de temps, alors que notre méthode de préservation asymptotiques permet de s'affranchir de cet inconvénient.

*La suite de ce chapitre est écrite en anglais.*

### 3.1 Introduction

The main objective of this work is to introduce and subsequently investigate an efficient numerical scheme for the resolution of evolution equations containing stiff transport terms, namely

$$\partial_t f^\epsilon + \mathcal{L}f^\epsilon + \frac{\mathbf{b}}{\epsilon} \cdot \nabla f^\epsilon = 0, \quad t \in \mathbb{R}^+, \mathbf{x} \in \Omega \subset \mathbb{R}^d, \quad (3.1)$$

where  $\mathbf{b} : \mathbb{R}^+ \times \Omega \rightarrow \mathbb{R}^d$  is a known (passive, linear transport model) or self-consistently computed (active, nonlinear transport model) vector-field satisfying  $\nabla \cdot \mathbf{b} = 0$ , and  $\mathcal{L}$  is a given operator (for ex. transport or diffusion operator). The small parameter  $\epsilon \ll 1$  represents the stiffness of the problem and signifies that we have to cope with a very strong vector-field  $\mathbf{b}$ . It brings up the main difficulties in the numerical resolution of (3.1), due to the introduction of multiple scales in the problem. Indeed, the dynamics along the  $\mathbf{b}$ -field is very rapid, as compared to its perpendicular evolution. In the formal limit  $\epsilon \rightarrow 0$ , the problem reduces to the constraint

$$\mathbf{b} \cdot \nabla f^0 = 0, \quad (3.2)$$

which signifies that the unknown  $f^0$  is constant along the field-lines of  $\mathbf{b}$ . However, in general, problem (3.2) does not permit to determine these different constants, for example when  $\mathbf{b}$  has closed field lines in the considered domain. Thus the reduced problem (3.2) is ill-posed, information has been lost while setting formally  $\epsilon = 0$  in (3.1). This feature is typical for singularly-perturbed problems or multi-scale problems (see [23, 129]).

The study of multi-scale problems is challenging from a mathematical as well a numerical point of view. Standard explicit numerical schemes require very small time steps, dependent on the  $\epsilon$ -parameter, in order to accurately account for the microscopic information (living at the  $\epsilon$ -scale). This procedure, even if accurate, has however the big disadvantage of being numerically very costly in simulation time and memory. Fully implicit schemes or IMEX-schemes are also not of use for  $\epsilon \ll 1$ , due to the ill-conditioned reduced model (3.2). Alternative methodologies are thus required, taking into account for the various scales present in the problem. Asymptotic analysis will be one of the mathematical tools used in this paper, permitting to recover the microscopic information lost in the reduced model (3.2). The numerical scheme presented here is based on such developments.

Evolution equations of the type (3.1) arise often in applications coming from fluid dynamics (see [102]) and plasma physics (see [26, 64, 73]). To mention only some examples, in thermonuclear fusion plasmas, the evolution of ions is described via the non-dimensional Vlasov ( $\eta = 0$ ) or Fokker-Planck ( $\eta > 0$ ) equation

$$\partial_t f_i^\epsilon + \mathbf{v} \cdot \nabla_{\mathbf{x}} f_i^\epsilon + \left( \mathbf{E} + \frac{1}{\epsilon} \mathbf{v} \times \mathbf{B} \right) \cdot \nabla_{\mathbf{v}} f_i^\epsilon = \eta \nabla_{\mathbf{v}} \cdot \left[ \mathbf{v} f_i^\epsilon + \nabla_{\mathbf{v}} f_i^\epsilon \right], \quad (3.3)$$

where  $f_i^\epsilon(t, \mathbf{x}, \mathbf{v})$  represents the ion distribution, dependent on time, space and velocity. This equation is coupled via the electromagnetic fields ( $\mathbf{E}(t, \mathbf{x}), \mathbf{B}(t, \mathbf{x})$ ) to an equation describing the electron evolution. The coupling is done by means of Maxwell's equations or Poisson's equation in the electrostatic case. The magnetic field is strong in tokamak experiments in the aim to confine the plasma and to render the fusion possible. This feature is translated in (3.3) in the magnitude of the scaling parameter  $\epsilon \ll 1$ .

The further example concerns the long-time asymptotic study of the electron 1D1V Vlasov-Poisson system

$$\begin{cases} \partial_t f_e + v \partial_x f_e - E(t, x) \partial_v f_e = 0, & \forall t \in \mathbb{R}^+, \forall (x, v) \in \Omega \subset \mathbb{R}^2 \\ -\partial_{xx} \varphi = 1 - n_e, & n_e(t, x) = \int_{\mathbb{R}} f_e(t, x, v) dv, \quad E = -\partial_x \varphi. \end{cases} \quad (3.4)$$

Introducing the field  $\mathbf{u} := (v, -E(t, x))^t$  and the stream function  $\Psi := \frac{1}{2}|v|^2 - \varphi(t, x)$ , one has  $\mathbf{u} = {}^\perp \nabla \Psi$ , where  ${}^\perp \nabla := (\partial_v, -\partial_x)$ . Considering additionally long-time scales, the Vlasov-Poisson system (3.4) transforms into the nonlinear, coupled system

$$\begin{cases} \partial_t f_e^\epsilon + \frac{\mathbf{u}^\epsilon}{\epsilon} \cdot \nabla_{x,v} f_e^\epsilon = 0, & \forall (t, x, v) \in \mathbb{R}^+ \times \Omega \\ -\Delta_{x,v} \Psi^\epsilon = n_e^\epsilon - 2, & n_e^\epsilon(t, x) = \int_{\mathbb{R}} f_e^\epsilon(t, x, v) dv, \quad \mathbf{u}^\epsilon = {}^\perp \nabla \Psi^\epsilon. \end{cases} \quad (3.5)$$

Finally a last example comes from fluid mechanics : consider the incompressible Euler equations in the long-time scaling, describing a bi-dimensional, inviscid flow with velocity  $\mathbf{u} := (u_1, u_2, 0)$  and pressure  $p$

$$\begin{cases} \epsilon \partial_t \mathbf{u}^\epsilon + (\mathbf{u}^\epsilon \cdot \nabla) \mathbf{u}^\epsilon + \nabla p^\epsilon = 0, & \forall (t, \mathbf{x}) \in \mathbb{R}^+ \times \Omega, \\ \nabla \cdot \mathbf{u}^\epsilon = 0. \end{cases} \quad (3.6)$$

Introducing the vorticity  $\omega^\epsilon := \nabla \times \mathbf{u}^\epsilon$ , the Euler system leads to the following nonlinear, coupled system

$$\begin{cases} \partial_t \omega^\epsilon + \frac{\mathbf{u}^\epsilon}{\epsilon} \cdot \nabla \omega^\epsilon = 0, \\ -\Delta \Psi^\epsilon = \omega^\epsilon, \quad \mathbf{u}^\epsilon = {}^\perp \nabla \Psi^\epsilon, \end{cases} \quad (3.7)$$

constituted of a transport equation for the vorticity, which is self-consistently coupled with a Poisson equation for the determination of the stream-function  $\Psi^\epsilon$ , result of the divergence-free constraint of  $\mathbf{u}^\epsilon$ . When adding in (4.1) a small viscosity term  $\nu \Delta \omega^\epsilon$ ,  $\nu$  being the reciprocal of the Reynolds number, one obtains an equation related to the incompressible Navier-Stokes equations.

The goal of this work is now to develop an efficient, uniformly accurate and stable (w.r.t.  $\epsilon$ ) numerical scheme for the resolution of the following linear, stiff transport problem

$$(V)^\epsilon \begin{cases} \partial_t f^\epsilon + \frac{\mathbf{b}}{\epsilon} \cdot \nabla f^\epsilon = 0, & \forall t \in (0, T), \forall \mathbf{x} = (x, y) \in \Omega \subset \mathbb{R}^2, \\ f^\epsilon(0, \mathbf{x}) = f_{in}(\mathbf{x}) & \forall \mathbf{x} \in \Omega, \end{cases} \quad (3.8)$$

with given, smooth and time-independent vector-field  $\mathbf{b} : \Omega \rightarrow \mathbb{R}^2$ , satisfying  $\nabla \cdot \mathbf{b} = 0$ . This simplified transport equation contains the numerical difficulties arising, due to stiffness, in equation (3.1). Given an efficient numerical algorithm for the resolution of (3.8), the above examples can be managed as follows : the nonlinear coupling can be treated iteratively, as shall be shown in Section 3.6 for the Vlasov-Poisson test case, and the discretization of the general non-stiff term  $\mathcal{L}f^\epsilon$  of (3.1) can be done via standard schemes suited for this particular operator. The scheme we propose in this paper shall be verified in two test cases : firstly in a simplified version of the gyrokinetic scaling (3.3), containing only the

stiff magnetic term  $\frac{1}{\epsilon}(\mathbf{v} \times \mathbf{B}) \cdot \nabla_{\mathbf{v}} f_i^\epsilon$ , and secondly in the long-time asymptotics of the Vlasov-Poisson system (3.5).

Due to the divergence constraint of  $\mathbf{b}$ , there exists a stream-function  $\Psi$  such that  $\mathbf{b} = \perp \nabla \Psi$ . Using the Poisson-bracket notation for two functions  $\chi, \theta$ , namely

$$\{\chi, \theta\} := \partial_x \chi \partial_y \theta - \partial_y \chi \partial_x \theta = \nabla \chi \cdot \perp \nabla \theta,$$

the transport equation (3.8) can be simply rewritten as

$$(V)^\epsilon \begin{cases} \partial_t f^\epsilon + \frac{1}{\epsilon} \{f^\epsilon, \Psi\} = 0, & t \in (0, T), \mathbf{x} \in \Omega \subset \mathbb{R}^2, \\ f^\epsilon(0, \cdot) = f_{in}, \end{cases} \quad (3.9)$$

and shall be completed with adequate boundary conditions, depending on the shape of the domain  $\Omega$  and on the vector-field  $\mathbf{b}$ . In order to recover the examples presented above, we shall investigate two different cases, resumed in the following Hypothesis.

**Hypothesis A** : *The domain  $\Omega$  will be either the whole  $\mathbb{R}^2$  (simplified case (3.3)) or an infinite strip  $(L_1, L_2) \times \mathbb{R}$  (case (3.5)) of the  $(x, y)$ -plane. In the second case, we shall assume periodic boundary conditions in  $x$  and the field  $\mathbf{b} : \Omega \mapsto \mathbb{R}^2$  is supposed to be also periodic in  $x$ . Furthermore  $\mathbf{b}$  is supposed to be sufficiently smooth, time-independent and divergence-free  $\nabla \cdot \mathbf{b} = 0$ .*

With regard to the numerical solution of (3.8), we aim to :

- design a simple and robust numerical scheme, working on a Cartesian grid ;
- design a scheme which enjoys the Asymptotic-Preserving properties (AP-scheme), in the sense that it has to be uniformly stable and accurate w.r.t..  $\epsilon$  ;
- give a detailed explanation why the proposed AP-method behaves better than standard methods (explicit, implicit, IMEX) ;
- design a scheme which is simply “generalizable” to more dimensions and various advection fields.

Let us underline at this point one important fact. We are interested in designing a scheme working on a Cartesian grid. One can imagine that for stiff problems of the type (3.8) (or more generally (3.1)), it could be better to adapt the coordinate system, choosing field-aligned variables, and transforming thus the problem into an evolution problem with a strong anisotropy aligned with one coordinate axis, which is much simpler to solve (via IMEX schemes for ex., see [51]). However, our aim is to avoid a coordinate transformation and to design a simple scheme based on a Cartesian grid. The advantage is that the numerical treatment becomes simpler, the disadvantage will be mentioned in Section 3.2, namely the introduction of a second, auxiliary unknown. The scheme we are going to propose is hence an alternative to the existing schemes for such evolution problems with stiff transport terms, and marries at the same time simplicity and Asymptotic-Preserving property.

Several AP-schemes were designed in the last years for various types of problems, including aniso-tropic elliptic [35, 36] or parabolic [107] equations, Vlasov equations in the hydrodynamic regime [55], drift-diffusion regime [30, 84] or the high-field limit [31, 81], Euler system in the low-Mach regime [37, 42]. Briefly, an AP-scheme is a numerical scheme specially designed for singularly-perturbed problems  $P^\epsilon$ , containing some small parameter  $\epsilon \ll 1$ , and which enjoy the following properties (see commutative diagram 4.2) :

- for fixed  $\epsilon > 0$ , the AP-scheme, denoted in this diagram  $P^{\epsilon, h}$ , is a consistent discretization of the continuous problem  $P^\epsilon$ , where  $h$  denotes the grid parameters ;



- the stability condition is independent of  $\epsilon$  ;
- for fixed discretization parameters  $h$ , the AP-scheme  $P^{\epsilon,h}$  provides in the limit  $\epsilon \rightarrow 0$  a consistent discretization of the limit problem  $P^0$ .

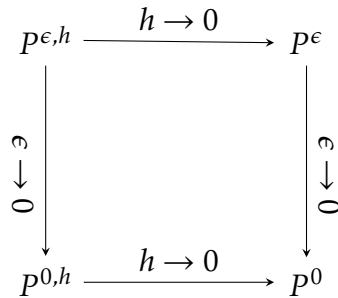


FIGURE 3.1 – Properties of AP-schemes.

One can put these schemes in the category of multi-scale numerical methods. At the end, let us also remark here that standard schemes for the resolution of (3.1) exist in literature, based on Galerkin methods [74], IMEX-techniques [18] or spectral methods [57].

The outline of this paper is the following. In Section 3.2, the asymptotic-preserving reformulation of the singularly-perturbed advection equation (3.8) is detailed. The reformulation is based on a micro-macro decomposition and a stabilization procedure. Section 3.3 deals with some mathematical aspects of the AP-reformulation, in order to show its well-posedness. Section 3.4 presents the numerical discretization of our asymptotic-preserving procedure. Section 3.5 focuses on a mathematical test case and its numerical resolution by our AP-scheme. In particular, we study the stabilization of the numerical scheme. Finally, Section 3.6 is dedicated to the numerical resolution of the Vlasov-Poisson 1D1V system. We focus notably in this part on the long-time behavior of the two-stream instability, leading to BGK-like equilibria. The last section concludes the paper with some remarks and perspectives.

## 3.2 Asymptotic-Preserving reformulation

We shall present in this section an AP-reformulation of the singularly-perturbed advection problem (3.8) completed with adequate boundary conditions, explicated in *Hypothesis A*, scheme which shall behave better (regularly) in the limit  $\epsilon \rightarrow 0$ . For this, the well-posed limit-model has firstly to be identified by investigating the asymptotic behaviour of the solutions  $f^\epsilon$ , as  $\epsilon \ll 1$ . We underline here that  $\mathbf{b}$  is time-independent in the following, if not explicitly mentioned, as in Section 3.6.

### 3.2.1 Identification of the limit model

As mentioned in the introduction, letting formally  $\epsilon \rightarrow 0$  in (3.8), leads to an ill-posed problem, which does not permit to compute in a unique manner the limit solution  $f^0(t, x, y)$ . The only information we get is that  $f^0$  is constant along the field-lines of  $\mathbf{b}$ .

In order to establish the limit model  $(V)^0$  corresponding to (3.8), let us suppose that  $f^\epsilon$  admits the following Hilbert expansion

$$f^\epsilon = f^0 + \epsilon f^1 + \epsilon^2 f^2 + \dots \quad (3.10)$$

Injecting this Ansatz in (3.8) leads to the infinite hierarchy of equations

$$\mathbf{b} \cdot \nabla f^0 = 0, \quad (3.11)$$

$$\partial_t f^0 + \mathbf{b} \cdot \nabla f^1 = 0, \quad (3.12)$$

$$\partial_t f^1 + \mathbf{b} \cdot \nabla f^2 = 0, \quad (3.13)$$

⋮

Equation (3.11) reveals that  $f^0$  belongs to the kernel of the dominant operator  $\mathcal{T} := \mathbf{b} \cdot \nabla$ . However, this information is not enough to determine completely  $f^0$ . It is necessary to use the next equation (3.12), to get the missing information. To eliminate  $f^1$  from this equation, one projects (3.12) on the kernel of  $\mathcal{T}$ . This projection is nothing else than the average of a quantity  $q$  along the field lines of  $\mathbf{b}$  and will be denoted by  $\langle q \rangle$ . Briefly, if  $Z(s; \mathbf{x})$  is the characteristic flow associated to the field  $\mathbf{b}$ , *i.e.*

$$\begin{cases} \frac{d}{ds} Z(s; \mathbf{x}) = \mathbf{b}(Z(s; \mathbf{x})), \\ Z(0; \mathbf{x}) = \mathbf{x}, \end{cases}$$

the average of a function  $q \in L^2(\Omega)$  over the field lines of  $\mathbf{b}$  is defined as

$$\langle q \rangle(\mathbf{x}) := \lim_{S \rightarrow \infty} \frac{1}{S} \int_0^S q(Z(s; \mathbf{x})) ds \quad \forall \mathbf{x} \in \Omega.$$

One can show (after some hypothesis on the regularity of  $\mathbf{b}$ , see [19]) that  $\langle \cdot \rangle$  is a well-defined application, furthermore that  $\langle q \rangle$  is constant along the field lines of  $\mathbf{b}$  and that one has  $\langle \mathbf{b} \cdot \nabla q \rangle = 0$ . The above mentioned procedure permits then to obtain a well-posed limit model for  $f^0$ . We already know that  $f^0$  belongs to the kernel of  $\mathcal{T}$ , meaning  $f^0 = \langle f^0 \rangle$ , such that the limit model  $(V)^0$  writes

$$(V)^0 \begin{cases} \partial_t f^0 = 0, & \mathbf{b} \cdot \nabla f^0 = 0, & \forall (t, \mathbf{x}) \in (0, T) \times \Omega, \\ f^0(0, \mathbf{x}) = \langle f_{in}(\mathbf{x}) \rangle \in \ker \mathcal{T}, & \mathbf{x} \in \Omega. \end{cases} \quad (3.14)$$

The following theorem proves rigorously the convergence of the solution  $f^\epsilon$  of (3.8) towards the solution  $f^0$  of the limit model (3.14), as  $\epsilon \rightarrow 0$ .

**Theorem 3.2.1** [19] *Consider a subset  $\Omega$  of  $\mathbb{R}^2$  satisfying Hypothesis A. Assume that  $\mathbf{b} \in W_{loc}^{1,\infty}(\mathbb{R}^2)$  (where in the case  $\Omega$  is a strip, we extend  $\mathbf{b}$  periodically to the whole  $\mathbb{R}^2$ ) satisfying  $\nabla \cdot \mathbf{b} = 0$  as well as the growth condition*

$$\exists C > 0 \text{ s.t. } |\mathbf{b}(\mathbf{x})| \leq C(1 + |\mathbf{x}|), \quad \forall \mathbf{x} \in \Omega.$$

*Suppose furthermore that  $f_{in} \in L^2(\Omega)$ . Then (3.8) resp. (3.14) admit unique weak solutions  $f^\epsilon, f^0 \in L^\infty(0, T; L^2(\Omega))$  and one has  $f^\epsilon \rightharpoonup f^0$ , weakly- $\star$  in  $L^\infty(0, T; L^2(\Omega))$ . If the initial conditions are well prepared in the sense that  $f_{in}^\epsilon$  is smooth enough and satisfies  $f_{in}^\epsilon \xrightarrow{\epsilon \rightarrow 0} f_{in}^0 \in \ker \mathcal{T}$  in  $L^2(\Omega)$ , then one has even  $f^\epsilon \xrightarrow{\epsilon \rightarrow 0} f^0$  in  $L^\infty(0, T; L^2(\Omega))$ .*

### 3.2.2 Micro-Macro reformulation

The design of a multiscale numerical procedure for the resolution of problem (3.8) is now inspired by the asymptotic study performed in Section 3.2.1. To recover the missing microscopic information in the reduced model (3.2), we shall decompose  $f^\epsilon$  into a macroscopic and a microscopic part, as follows

$$f^\epsilon = p^\epsilon + \epsilon q^\epsilon, \quad \text{with } \mathbf{b} \cdot \nabla f^\epsilon = \epsilon \mathbf{b} \cdot \nabla q^\epsilon. \quad (3.15)$$

This signifies that  $p^\epsilon$  belongs to the kernel of the dominant operator  $\mathcal{T} = \mathbf{b} \cdot \nabla$  and is considered as the macroscopic part. This decomposition is not unique as one has still to fix the values of  $p^\epsilon$  (or equivalently  $q^\epsilon$ ) on the field-lines, fact which shall be done in the next subsections.

Plugging for the moment (3.15) into (3.8) leads to the following augmented system for the two unknowns  $(f^\epsilon, q^\epsilon)$

$$\begin{cases} \partial_t f^\epsilon + \mathbf{b} \cdot \nabla q^\epsilon = 0, & \forall (t, \mathbf{x}) \in (0, T) \times \Omega, \\ \mathbf{b} \cdot \nabla f^\epsilon = \epsilon \mathbf{b} \cdot \nabla q^\epsilon, & \forall (t, \mathbf{x}) \in (0, T) \times \Omega, \end{cases} \quad (3.16)$$

associated with the initial condition  $f^\epsilon(0, \cdot) = f_{in}$  and adequate boundary conditions (*Hypothesis A*). Now several possibilities are conceivable to fix the values of  $q^\epsilon$  on the field-lines, rendering thus the decomposition (3.15) unique. Let us observe here that the values of  $q^\epsilon$  on these lines are of no importance for the computation of our physical unknown  $f^\epsilon$ , as only  $\mathbf{b} \cdot \nabla q^\epsilon$  is occurring in the system (3.16). Thus any arbitrary choice could do the work.

### 3.2.3 Zero mean value

From a purely mathematical point of view, one first idea is to fix the average of  $q^\epsilon$  along the field lines of  $\mathbf{b}$ , by enforcing zero mean, *i.e.*

$$\langle q^\epsilon \rangle = 0. \quad (3.17)$$

Imposing (3.17) can be done by slightly changing the system, adding an additional “subtle” term  $\sigma \langle q^\epsilon \rangle$ , with  $\sigma \in \mathbb{R}$  an arbitrary constant, namely

$$\begin{cases} \partial_t f^\epsilon + \mathbf{b} \cdot \nabla q^\epsilon = 0, \\ \mathbf{b} \cdot \nabla f^\epsilon = \epsilon \mathbf{b} \cdot \nabla q^\epsilon - \sigma \langle q^\epsilon \rangle. \end{cases} \quad (3.18)$$

Indeed, one can remark immediately that taking the average of the second equation over the field-lines yields automatically the constraint  $\langle q^\epsilon \rangle = 0$ , since  $\langle \mathbf{b} \cdot \nabla \theta \rangle = 0$ , for all  $\theta$ . The new introduced term is hence a tricky zero, rendering  $q^\epsilon$  unique by fixing its average values along the  $\mathbf{b}$ -lines to zero. One can show now that (3.18) is completely equivalent to (3.8), for each  $\epsilon > 0$ . Indeed, two ingredients help to prove this equivalence between both formulations. On one hand, for given  $f^\epsilon$  and  $\epsilon > 0$ , the equation

$$\begin{cases} \mathbf{b} \cdot \nabla q^\epsilon = \frac{1}{\epsilon} \mathbf{b} \cdot \nabla f^\epsilon, \\ \langle q^\epsilon \rangle = 0, \end{cases}$$

has a unique solution  $q^\epsilon$ . On the other hand, the second equation in (3.18) yields immediately, as mentioned above, the constraint  $\langle q^\epsilon \rangle = 0$ .

This idea is very nice from a mathematical point of view, however, if one is thinking at the numerical implementation, one has to average over the field lines of  $\mathbf{b}$ , in order to discretize the new term  $\sigma \langle q^\epsilon \rangle$  in the second equation of (3.18). This procedure is rather hard (we are working on Cartesian grids with not-aligned fields  $\mathbf{b}$ ) and can introduce moreover  $\epsilon$ -dependent error terms in the results. Thus we shall leave this idea behind, and search for a more practical one.

### 3.2.4 Regularization

In order to render  $q^\epsilon$  unique in (3.16), one can imagine to use a regularization technique. Regularization is a very broad field in mathematics, and is devoted to the design and analysis of methods for obtaining stable solutions of ill-posed problems. In particular, the usual regularization technique consists in replacing the ill-posed problem by a nearby (slightly-perturbed) well-posed problem, whose resolution poses no difficulties (uniqueness, stability of the solution). The original solution is recovered only in the limit of vanishing regularization/perturbation parameter. The choice of the perturbation term as well as the strength of the perturbation parameter is essential and constitutes the key point of the method. There is a rich literature on regularization techniques, we refer the interested reader to the references [16, 22, 24, 48].

Coming now to our problem, one can regularize (3.16) either by adding in the second equation a term of the type “ $\alpha \partial_t q^\epsilon$ ” or of the form “ $\sigma q^\epsilon$ ”. These two regularizations permit to get a unique  $q^\epsilon$ , however the corresponding solutions behave very differently in the desired asymptotic limit  $\epsilon \rightarrow 0$ . To see this difference, let us simplify by putting formally  $\epsilon = 0$  in (3.16) and take a look at both regularizations, namely

$$(W) \begin{cases} \partial_t f^0 + \mathbf{b} \cdot \nabla q^0 = 0, \\ \alpha \partial_t q^0 + \mathbf{b} \cdot \nabla f^0 = 0, \end{cases} \quad (P) \begin{cases} \partial_t f^0 + \mathbf{b} \cdot \nabla q^0 = 0, \\ \mathbf{b} \cdot \nabla f^0 + \sigma q^0 = 0, \end{cases} \quad (3.19)$$

where the constants  $\alpha > 0$  resp.  $\sigma > 0$  have to be sufficiently small in order not to perturb too much the original problem. Now, one can eliminate in both systems the auxiliary unknown  $q^0$  and get an equation involving only  $f^0$ , which reads

$$(W) \partial_{tt} f^0 - \frac{1}{\alpha} \nabla \cdot [(\mathbf{b} \otimes \mathbf{b}) \nabla f^{0,\sigma}] = 0; \quad (P) \partial_t f^0 - \frac{1}{\sigma} \nabla \cdot [(\mathbf{b} \otimes \mathbf{b}) \nabla f^{0,\sigma}] = 0.$$

As one can observe, regularizing (3.16) by adding a term of the form “ $\alpha \partial_t q^\epsilon$ ” leads in the limit  $\epsilon \rightarrow 0$  to a wave-equation, whereas the regularization by adding “ $\sigma q^\epsilon$ ” leads to a parabolic equation.

Which one is better suited for our anisotropic transport problem can be understood by remembering the asymptotic behaviour of the unique solution  $f^\epsilon$  of the original advection equation (3.8) as  $\epsilon$  becomes smaller and smaller. As shown in Section 3.2.1, one gets in the (weak) limit  $\epsilon \rightarrow 0$  a function  $f^0$  which is constant along the field lines of  $\mathbf{b}$ . This gives us a hint that the regularization (P) is better suited for our problem, as the corresponding limit problem is a diffusion problem, with strong diffusivity along the field lines. Hence one is expecting to get a better approximation of  $f^0$  via (P) than via a wave-equation of the form (W), which describes rather a very rapid wave-motion.

### 3.2.5 The stabilized AP-reformulation

To summarize, our Asymptotic-Preserving scheme for an efficient resolution of the anisotropic transport equation (3.8) is based on the resolution of the following well-posed reformulated system

$$(MM)_\epsilon^\sigma \begin{cases} \partial_t f^{\epsilon,\sigma} + \mathbf{b} \cdot \nabla q^{\epsilon,\sigma} = 0, & \forall (t, \mathbf{x}) \in (0, T) \times \Omega, \\ \mathbf{b} \cdot \nabla f^{\epsilon,\sigma} = \epsilon \mathbf{b} \cdot \nabla q^{\epsilon,\sigma} - \sigma q^{\epsilon,\sigma}, & \forall (t, \mathbf{x}) \in (0, T) \times \Omega, \end{cases} \quad (3.20)$$

with  $\sigma > 0$  a small parameter to be fixed later on. This system is completed by an initial condition  $f^{\epsilon,\sigma}(0, \cdot) = f_{in}(\cdot)$  and adequate boundary conditions (*Hypothesis A*). Remark that no initial condition is needed for the auxiliary variable  $q^\epsilon$ . Let us underline here the difference between (3.18) and (3.20). Both procedures are fixing the value of the auxiliary variable  $q$  on the field lines of  $\mathbf{b}$  by imposing zero mean  $\langle q \rangle = 0$ . To see this in (3.20), it is enough to take the average  $\langle \cdot \rangle$  of the second equation. However, while (3.18) is completely equivalent to the starting model (3.8), the system (3.20) introduces an error, as the supplementary term we introduced,  $\sigma q^{\epsilon,\sigma}$ , is no more zero but contains also the non-zero fluctuation part of  $q^{\epsilon,\sigma}$ . The big advantage of (3.20) with respect to (3.18) is that a discretization of the average operator  $\langle \cdot \rangle$  is not necessary anymore.

The  $\epsilon$ -regularity of the system (3.20) allows now to pass directly to the  $\epsilon \rightarrow 0$  limit in (3.20) to get the corresponding limit model, *i.e.*

$$(MM)_0^\sigma \begin{cases} \partial_t f^{0,\sigma} + \mathbf{b} \cdot \nabla q^{0,\sigma} = 0, \\ \mathbf{b} \cdot \nabla f^{0,\sigma} + \sigma q^{0,\sigma} = 0. \end{cases} \quad (3.21)$$

Eliminating  $q^{0,\sigma}$  from this system, yields the degenerate diffusion equation with fixed  $\sigma > 0$ )

$$\partial_t f^{0,\sigma} - \frac{1}{\sigma} \nabla \cdot [(\mathbf{b} \otimes \mathbf{b}) \nabla f^{0,\sigma}] = 0, \quad (3.22)$$

which shows clearly what the regularization term is doing in the limit  $\epsilon \rightarrow 0$ . Let us also remark here that (3.22) is a well-posed problem (for  $\sigma > 0$ ), due to the standard parabolic theory associated with the evolution triple  $Q \hookrightarrow V \equiv V^\star \hookrightarrow Q^\star$ , where the spaces are defined in the next section. Remark that the positive, symmetric bilinear form  $m(u, v) := \frac{1}{\sigma} (\mathbf{b} \cdot \nabla u, \mathbf{b} \cdot \nabla v)_V$  satisfies Garding's inequality (see [131] for more details). For future numerical discretizations, it will be more convenient to rewrite these systems by using the Poisson bracket. Introducing the stream function  $\Psi$  such that  $\mathbf{b} = (\partial_y \Psi, -\partial_x \Psi) = {}^\perp \nabla \Psi$ , the previous Micro-Macro system (3.20) reads

$$(MM)_\epsilon^\sigma \begin{cases} \partial_t f^{\epsilon,\sigma} + \{q^{\epsilon,\sigma}, \Psi\} = 0, \\ \{f^{\epsilon,\sigma}, \Psi\} = \epsilon \{q^{\epsilon,\sigma}, \Psi\} - \sigma q^{\epsilon,\sigma}. \end{cases} \quad (3.23)$$

Before introducing the numerical discretization of the AP-reformulation (3.23) in Section 3.4, let us however do in the next section some mathematical investigations.

## 3.3 Some mathematical observations

The rigorous mathematical study of the existence and uniqueness of a solution to the AP-reformulation (3.20) along with the rigorous study of its limit towards (3.21) is a delicate question and shall be treated in a supplementary, more mathematical work. To give

however some ideas about the well-posedness of this model, and underline the difficulties of the mathematical study, we shall concentrate in this paper only on the existence and uniqueness questions corresponding to the implicit time semi-discretization of (3.20), namely

$$(MM)_\epsilon^{\sigma,n} \begin{cases} f^{\epsilon,\sigma,n+1} + \Delta t \mathbf{b} \cdot \nabla q^{\epsilon,\sigma,n+1} = f^{\epsilon,\sigma,n}, \\ \mathbf{b} \cdot \nabla f^{\epsilon,\sigma,n+1} = \epsilon \mathbf{b} \cdot \nabla q^{\epsilon,\sigma,n+1} - \sigma q^{\epsilon,\sigma,n+1}. \end{cases} \quad (3.24)$$

Here, we discretized the time interval  $[0, T]$  with fixed  $T > 0$ , as follows

$$t^n := n\Delta t, \quad \Delta t := T/N_t, \quad n \in \llbracket 0, N_t \rrbracket, \quad N_t \in \mathbb{N},$$

and denoted by  $f^{\epsilon,\sigma,n}$  an approximation of  $f^{\epsilon,\sigma}(t^n, \cdot)$ . This system is associated with boundary conditions for  $(f^{\epsilon,\sigma,n+1}, q^{\epsilon,\sigma,n+1})$  following *Hypothesis A*.

Let us now specify the mathematical framework of problem (3.24). For this, choose firstly the Hilbert spaces  $V = L^2(\Omega)$  and  $Q = \{v \in L^2(\Omega) \mid \mathbf{b} \cdot \nabla v \in L^2(\Omega), v \text{ satisfying the boundary conditions precised in } \textit{Hypothesis A}\}$ , associated with the standard  $L^2$  scalar-product for  $V$  and  $(u, v)_Q := (u, v)_{L^2} + (\mathbf{b} \cdot \nabla u, \mathbf{b} \cdot \nabla v)_{L^2}$  for  $Q$ . We introduce now the following bi-linear forms  $\mathcal{A}, \mathcal{B}, \mathcal{C}_{\epsilon,\sigma}$  :

$$\begin{aligned} \mathcal{A}(u, v) &:= \int_{\Omega} uv \, dx, & \mathcal{A} : V \times V &\rightarrow \mathbb{R}, \\ \mathcal{B}(v, r) &:= \int_{\Omega} v(\mathbf{b} \cdot \nabla r) \, dx, & \mathcal{B} : V \times Q &\rightarrow \mathbb{R}, \\ \mathcal{C}_{\epsilon,\sigma}(r, s) &:= -\epsilon \mathcal{B}(s, r) + \sigma (r, s)_{L^2}, & \mathcal{C}_{\epsilon,\sigma} : Q \times Q &\rightarrow \mathbb{R}, \end{aligned} \quad (3.25)$$

and their associated linear operators  $A, B, C_{\epsilon,\sigma}$  :

$$\begin{aligned} A : V &\rightarrow V^*, \quad B : V \rightarrow Q^*, \quad C_{\epsilon,\sigma} : Q \rightarrow Q^*, \\ \langle Au, v \rangle_{V^*, V} &:= \mathcal{A}(u, v), \quad \langle Bv, r \rangle_{Q^*, Q} := \mathcal{B}(v, r), \quad \langle C_{\epsilon,\sigma} r, s \rangle_{Q^*, Q} := \mathcal{C}_{\epsilon,\sigma}(r, s). \end{aligned}$$

**Remark 4** The bi-linear form  $\mathcal{B}$  defines also the adjoint linear operator  $B^* : Q \rightarrow V^*$  via  $\mathcal{B}(v, r) = \langle Bv, r \rangle_{Q^*, Q} = \langle v, B^* r \rangle_{V, V^*}$  for all  $(v, r) \in V \times Q$ . Observe also that  $B^* r = \mathbf{b} \cdot \nabla r$  for all  $r \in Q$ , whereas in the distributional sense  $Bv = -\mathbf{b} \cdot \nabla v$  for all  $v \in V$ .

With these definitions, the variational formulation of the previous problem (3.24) writes now : for fixed  $\epsilon \geq 0, \sigma > 0, \Delta t > 0$  and  $f^{\epsilon,\sigma,n} \in V^*$ , find  $(f^{\epsilon,\sigma,n+1}, q^{\epsilon,\sigma,n+1}) \in V \times Q$ , such that :

$$\begin{cases} \mathcal{A}(f^{\epsilon,\sigma,n+1}, \theta) + \Delta t \mathcal{B}(\theta, q^{\epsilon,\sigma,n+1}) = \langle f^{\epsilon,\sigma,n}, \theta \rangle_{V^* \times V}, & \forall \theta \in V, \\ \mathcal{B}(f^{\epsilon,\sigma,n+1}, \chi) - \mathcal{C}_{\epsilon,\sigma}(q^{\epsilon,\sigma,n+1}, \chi) = 0, & \forall \chi \in Q. \end{cases} \quad (3.26)$$

To prove the existence and uniqueness of a weak solution to (3.26), we shall need some properties of these operators.

**Lemma 1** Let *Hypothesis A* be satisfied. Then,  $\mathcal{A}(\cdot, \cdot), \mathcal{B}(\cdot, \cdot)$  resp.  $\mathcal{C}_{\epsilon,\sigma}(\cdot, \cdot)$  are continuous bi-linear forms on  $V \times V, V \times Q$  resp.  $Q \times Q$ . Furthermore,  $\mathcal{A}(\cdot, \cdot)$  is coercive on  $V \times V$  and  $\mathcal{C}_{\epsilon,\sigma}(\cdot, \cdot)$  is positive semi-definite on  $Q \times Q$ .

**Remark 5** Let us remark here that  $\mathcal{C}_{\epsilon,\sigma}(\cdot, \cdot)$  is not coercive on  $Q \times Q$ . However, as we will see later, this hypothesis is not crucial for both existence and uniqueness of a solution of the variational formulation (3.26).

Without any other hypothesis on  $\mathcal{B}(\cdot, \cdot)$  (as for example an inf-sup condition) we are now able to present the following existence/uniqueness result of a solution to (3.26), and this due to the presence of the regularization term  $\sigma q$ .

**Theorem 3.3.1** *Let Hypothesis A be satisfied and let  $\mathcal{A}(\cdot, \cdot)$ ,  $\mathcal{B}(\cdot, \cdot)$  and  $\mathcal{C}_{\epsilon, \sigma}(\cdot, \cdot)$  be the continuous bi-linear forms defined in (3.25). Then, for every given  $f^{\epsilon, \sigma, n} \in V^*$  the problem (3.26) has for each fixed  $\epsilon \geq 0$ ,  $\sigma > 0$ ,  $\Delta t > 0$  and  $n \in \llbracket 0, N_t \rrbracket$ , a unique weak solution  $(f^{\epsilon, \sigma, n+1}, q^{\epsilon, \sigma, n+1})$  in  $V \times Q$ .*

*This solution satisfies the following estimates, for all  $\epsilon \geq 0$ ,  $\sigma > 0$ ,  $\Delta t > 0$ ,  $n \in \llbracket 0, N_t \rrbracket$  :*

$$\|f^{\epsilon, \sigma, n+1}\|_V \leq \|f^{\epsilon, \sigma, n}\|_{V^*}, \quad (3.27)$$

$$\|q^{\epsilon, \sigma, n+1}\|_V^2 \leq \frac{1}{\sigma \Delta t} \|f^{\epsilon, \sigma, n}\|_{V^*}^2, \quad \|\mathbf{b} \cdot \nabla q^{\epsilon, \sigma, n+1}\|_V \leq \frac{2}{\Delta t} \|f^{\epsilon, \sigma, n}\|_{V^*}. \quad (3.28)$$

**Proof 10** *Due to the lack of coercivity of  $\mathcal{C}_{\epsilon, \sigma}$ , we shall start by considering the regularized problem : for  $\alpha > 0$ , find  $(f_\alpha^{\epsilon, \sigma, n+1} \in V, q_\alpha^{\epsilon, \sigma, n+1}) \in V \times Q$ , solution of :*

$$\begin{cases} \mathcal{A}(f_\alpha^{\epsilon, \sigma, n+1}, \theta) + \Delta t \mathcal{B}(\theta, q_\alpha^{\epsilon, \sigma, n+1}) = \langle f^{\epsilon, \sigma, n}, \theta \rangle_{V^* \times V}, & \forall \theta \in V, \\ \mathcal{B}(f_\alpha^{\epsilon, \sigma, n+1}, \chi) - \alpha (q_\alpha^{\epsilon, \sigma, n+1}, \chi)_Q - \mathcal{C}_{\epsilon, \sigma}(q_\alpha^{\epsilon, \sigma, n+1}, \chi) = 0, & \forall \chi \in Q. \end{cases} \quad (3.29)$$

*Multiplying the second equation by  $\Delta t$  and subtracting both equations, one can show via Lax-Milgram theorem that (3.29) admits, for each fixed  $\alpha > 0$ ,  $\epsilon \geq 0$ ,  $\sigma > 0$ ,  $\Delta t > 0$  and  $n \in \llbracket 0, N_t \rrbracket$ , a unique weak solution. Indeed, thanks to the term  $\alpha (q_\alpha^{\epsilon, \sigma, n+1}, \chi)_Q$  in the second equation of (3.29), the regularized operator  $\mathcal{C}_{\epsilon, \sigma}(\cdot, \cdot) + \alpha(\cdot, \cdot)_Q$  is now coercive on  $Q \times Q$ .*

*Our aim is now to bound  $f_\alpha^{\epsilon, \sigma, n+1}$  and  $q_\alpha^{\epsilon, \sigma, n+1}$  uniformly in  $\alpha$ . Then, passing to the limit  $\alpha \rightarrow 0$  in (3.29) would permit to conclude about both existence and uniqueness of a weak solution of (3.26).*

*First, let us choose  $\theta = f_\alpha^{\epsilon, \sigma, n+1}$  and  $\chi = q_\alpha^{\epsilon, \sigma, n+1}$  in (3.29), multiply the second equation with  $\Delta t$  and subtract both equations to get*

$$\mathcal{A}(f_\alpha^{\epsilon, \sigma, n+1}, f_\alpha^{\epsilon, \sigma, n+1}) + \alpha \Delta t \|q_\alpha^{\epsilon, \sigma, n+1}\|_Q^2 + \Delta t \mathcal{C}_{\epsilon, \sigma}(q_\alpha^{\epsilon, \sigma, n+1}, q_\alpha^{\epsilon, \sigma, n+1}) = \langle f^{\epsilon, \sigma, n}, f_\alpha^{\epsilon, \sigma, n+1} \rangle_{V^* \times V}.$$

*By using the coercivity of  $\mathcal{A}(\cdot, \cdot)$  and the fact that  $\mathcal{C}_{\epsilon, \sigma}(q, q) = \sigma \|q\|_{L^2}^2$  (due to the boundary conditions), we get*

$$\|f_\alpha^{\epsilon, \sigma, n+1}\|_V^2 + \alpha \Delta t \|q_\alpha^{\epsilon, \sigma, n+1}\|_Q^2 + \Delta t \sigma \|q_\alpha^{\epsilon, \sigma, n+1}\|_V^2 \leq \|f^{\epsilon, \sigma, n}\|_{V^*} \|f_\alpha^{\epsilon, \sigma, n+1}\|_V,$$

*leading to*

$$\|f_\alpha^{\epsilon, \sigma, n+1}\|_V \leq \|f^{\epsilon, \sigma, n}\|_{V^*}, \quad \|q_\alpha^{\epsilon, \sigma, n+1}\|_V^2 \leq \frac{1}{\Delta t \sigma} \|f^{\epsilon, \sigma, n}\|_{V^*}^2. \quad (3.30)$$

*Now, thanks to the first equation of (3.29), i.e.*

$$\Delta t \mathcal{B}(\theta, q_\alpha^{\epsilon, \sigma, n+1}) = -\mathcal{A}(f_\alpha^{\epsilon, \sigma, n+1}, \theta) + \langle f^{\epsilon, \sigma, n}, \theta \rangle_{V^* \times V}, \quad \forall \theta \in V,$$

*we have*

$$\|\mathbf{b} \cdot \nabla q_\alpha^{\epsilon, \sigma, n+1}\|_{L^2(\Omega)} = \sup_{\theta \in L^2(\Omega)} \frac{(\mathbf{b} \cdot \nabla q_\alpha^{\epsilon, \sigma, n+1}, \theta)_{L^2(\Omega)}}{\|\theta\|_{L^2(\Omega)}} \leq \frac{2}{\Delta t} \|f^{\epsilon, \sigma, n}\|_{V^*}. \quad (3.31)$$

*The estimates (3.30) as well as (3.31) are independent of  $\alpha > 0$  and thus one can extract weakly convergent subsequences and pass to the limit  $\alpha \rightarrow 0$  in the variational formulation (3.29) to conclude the proof.*

**Remark 6** One can also observe from the  $\epsilon$ -independent estimates (3.27)-(3.28) that up to extracting a subsequence of  $\{f^{\epsilon,\sigma,n+1}, q^{\epsilon,\sigma,n+1}\}_{\epsilon>0}$  we have the  $\epsilon$ -convergences

$$f^{\epsilon,\sigma,n+1} \underset{\epsilon \rightarrow 0}{\rightharpoonup} f^{0,\sigma,n+1} \quad \text{in } V, \quad q^{\epsilon,\sigma,n+1} \underset{\epsilon \rightarrow 0}{\rightharpoonup} q^{0,\sigma,n+1} \quad \text{in } Q,$$

which underlines the fact that the AP-reformulation (3.20) is regular and tends towards the limit-model (3.21) as  $\epsilon$  goes to zero.

The  $\sigma$ -convergences are more delicate. One has only

$$f^{\epsilon,\sigma,n+1} \underset{\sigma \rightarrow 0}{\rightharpoonup} f^{\epsilon,0,n+1} \quad \text{in } V,$$

however there is no convergence for  $q^{\epsilon,\sigma,n+1}$ . This again was to be expected as in the limit  $\sigma \rightarrow 0$  one loses the uniqueness of  $q$ , the term  $\sigma q$  being a regularization term.

## 3.4 Numerical discretization

Let us come now to the numerical part of our work and introduce here a numerical discretization of our reformulated system :

$$(MM)_\epsilon^\sigma \begin{cases} \partial_t f^{\epsilon,\sigma} + \{q^{\epsilon,\sigma}, \Psi^{\epsilon,\sigma}\} = 0, \\ \{f^{\epsilon,\sigma}, \Psi^{\epsilon,\sigma}\} = \epsilon \{q^{\epsilon,\sigma}, \Psi^{\epsilon,\sigma}\} - \sigma q^{\epsilon,\sigma}, \end{cases} \quad (3.32)$$

where the stream function  $\Psi^{\epsilon,\sigma}$  is supposed to be given in this section, linked to the given time-independent vector field  $\mathbf{b}$  through  $\mathbf{b} := \nabla^\perp \Psi^{\epsilon,\sigma}$ .

### 3.4.1 Discretization parameters

In what follows, one assumes a bounded simulation domain  $\Omega_S := (-L_x, L_x) \times (-L_y, L_y)$ . Concerning the time interval  $[0, T]$ ,  $T > 0$ , we employ the discretization :

$$t^n := n \Delta t, \quad \Delta t := T/N_t, \quad n \in \llbracket 0, N_t \rrbracket, \quad N_t \in \mathbb{N}.$$

Similarly, let us supply the domain  $\Omega_S$  with a uniform spatial discretization :

$$x_i := (i-1)\Delta x - L_x, \quad y_j := (j-1)\Delta y - L_y, \quad \Delta x := 2L_x/N_x, \quad \Delta y := 2L_y/N_y,$$

with  $i \in \llbracket 1, N_x + 1 \rrbracket$ , and  $j \in \llbracket 1, N_y + 1 \rrbracket$ . For any function  $f : [0, T] \times \Omega_S \rightarrow \mathbb{R}$ ,  $f_{i,j}^n$  refers to the numerical approximation of  $f(t^n, x_i, y_j)$ , and  $f_h^n$  shall simply denote the discrete grid-function  $(f_{i,j}^n)_{i,j}$ .

The domain  $\Omega_S$  is a truncation of the real physical domain  $\Omega = \mathbb{R}^2$  or  $\Omega = (L_1, L_2) \times \mathbb{R}$ . To be close to the physical reality, we took in our simulations for the variable belonging to  $\mathbb{R}$  a sufficiently large bounded domain and supposed that the distribution function  $f^{\epsilon,\sigma}$  is vanishing on that border, whereas periodic boundary conditions are imposed for the bounded variable. To be more precise, we imposed for the truncation of the physical domain  $\Omega = \mathbb{R}^2$  that

$$f_{i,j}^n = 0 \quad \text{for } i = 1; j = 1; i = N_x + 1; j = N_y + 1,$$

whereas for the truncation of the physical domain  $\Omega = (L_1, L_2) \times \mathbb{R}$  we imposed

$$f_{i,j}^n = 0 \quad \forall i \text{ and } j = 1; j = N_y + 1 \quad \text{as well as} \quad f_{1,j}^n = f_{N_x+1,j}^n \quad \forall j.$$



### 3.4.2 Space semi-discretization

For the Poisson brackets appearing in the Micro-Macro reformulation (3.32), let us adopt the second order Arakawa discretization [6]. For two functions  $u, v : \Omega_S \rightarrow \mathbb{R}$ , the discrete version of the Poisson bracket  $\{u, v\}$  calculated at the point  $(x_i, y_j)$  is expressed by :

$$[u_h, v_h]_{i,j} := \frac{1}{12\Delta x \Delta y} \left( u_{i+1,j} \mathcal{A}_{i,j} + u_{i-1,j} \mathcal{B}_{i,j} + u_{i,j+1} \mathcal{C}_{i,j} + u_{i,j-1} \mathcal{D}_{i,j} \right. \\ \left. + u_{i+1,j+1} \mathcal{E}_{i,j} + u_{i-1,j-1} \mathcal{F}_{i,j} + u_{i-1,j+1} \mathcal{G}_{i,j} + u_{i+1,j-1} \mathcal{H}_{i,j} \right).$$

where the coefficients write

$$\begin{aligned} \mathcal{A}_{i,j} &:= v_{i,j+1} - v_{i,j-1} + v_{i+1,j+1} - v_{i+1,j-1}, & \mathcal{E}_{i,j} &:= v_{i,j+1} - v_{i+1,j}, \\ \mathcal{B}_{i,j} &:= v_{i,j-1} - v_{i,j+1} - v_{i-1,j+1} + v_{i-1,j-1}, & \mathcal{F}_{i,j} &:= v_{i,j-1} - v_{i-1,j}, \\ \mathcal{C}_{i,j} &:= v_{i-1,j} - v_{i+1,j} - v_{i+1,j+1} + v_{i-1,j+1}, & \mathcal{G}_{i,j} &:= v_{i-1,j} - v_{i,j+1}, \\ \mathcal{D}_{i,j} &:= v_{i+1,j} - v_{i-1,j} + v_{i+1,j-1} - v_{i-1,j-1}, & \mathcal{H}_{i,j} &:= v_{i+1,j} - v_{i,j-1}. \end{aligned}$$

Thus, the semi-discretization in space of the Micro-Macro problem (3.32) reads :

$$(MM)_{\epsilon,h}^\sigma \begin{cases} \partial_t f_{i,j}^{\epsilon,\sigma} + [q_h^{\epsilon,\sigma}, \Psi_h^{\epsilon,\sigma}]_{i,j} = 0, \\ [f_h^{\epsilon,\sigma}, \Psi_h^{\epsilon,\sigma}]_{i,j} - \epsilon [q_h^{\epsilon,\sigma}, \Psi_h^{\epsilon,\sigma}]_{i,j} + \sigma q_{i,j}^{\epsilon,\sigma} = 0. \end{cases} \quad (3.33)$$

### 3.4.3 Time discretization

We shall use a DIRK (diagonally-implicit Runge-Kutta) approach in order to achieve second-order accuracy in time for the problem (3.32). The general form of a RK-method is recalled here for the following equation

$$\partial_t u = Lu,$$

where  $L$  refers to some differential operator. An  $r$ -stage Runge-Kutta approach is determined by its Butcher table

$$\begin{array}{c|ccc} c_1 & a_{11} & \dots & a_{1r} \\ \vdots & \vdots & & \vdots \\ c_r & a_{r1} & \dots & a_{rr} \\ \hline & b_1 & \dots & b_r \end{array}$$

For a given  $u^n$ , the subsequent  $u^{n+1}$  is defined by the formula

$$u^{n+1} = u^n + \Delta t \sum_{j=1}^r b_j u_j,$$

where each  $u_i$  is defined by

$$u_i = u^n + \Delta t \sum_{j=1}^r a_{ij} Lu_j.$$

In the case where  $b_j = a_{rj}$  for  $j = 1, \dots, r$ , then  $u^{n+1}$  is equal to the last stage of the method, namely  $u_r$ . For the Micro-Macro problem (3.32), we consider the following 2-stage Butcher table

$$\begin{array}{c|cc} \lambda & \lambda & 0 \\ 1 & 1-\lambda & \lambda \\ \hline & 1-\lambda & \lambda \end{array}$$

For  $\lambda := 1 - 1/\sqrt{2}$ , the method is  $L$ -stable. For all  $n \in \llbracket 0, N_t \rrbracket$ , the full discretization of the Micro-Macro problem (3.32) writes now

$$(MM)_{\epsilon, h}^{\sigma, n} \left\{ \begin{array}{l} \text{Stage 1 :} \\ f_{1,i,j}^{\epsilon, \sigma} + \lambda \Delta t [q_{1,h}^{\epsilon, \sigma}, \Psi_h^{\epsilon, \sigma}]_{i,j} = f_{i,j}^{\epsilon, \sigma, n}, \\ [f_{1,h}^{\epsilon, \sigma}, \Psi_h^{\epsilon, \sigma}]_{i,j} - \epsilon [q_{1,h}^{\epsilon, \sigma}, \Psi_h^{\epsilon, \sigma}]_{i,j} + \sigma q_{1,i,j}^{\epsilon, \sigma} = 0. \\ \text{Stage 2 :} \\ f_{2,i,j}^{\epsilon, \sigma} + \lambda \Delta t [q_{2,h}^{\epsilon, \sigma}, \Psi_h^{\epsilon, \sigma}]_{i,j} = f_{i,j}^{\epsilon, \sigma, n} + \frac{1-\lambda}{\lambda} (f_{1,i,j}^{\epsilon, \sigma} - f_{i,j}^{\epsilon, \sigma, n}), \\ [f_{2,h}^{\epsilon, \sigma}, \Psi_h^{\epsilon, \sigma}]_{i,j} - \epsilon [q_{2,h}^{\epsilon, \sigma}, \Psi_h^{\epsilon, \sigma}]_{i,j} + \sigma q_{2,i,j}^{\epsilon, \sigma} = 0. \\ \text{Final Stage :} \\ (f_{i,j}^{\epsilon, \sigma, n+1}, q_{i,j}^{\epsilon, \sigma, n+1}) = (f_{2,i,j}^{\epsilon, \sigma}, q_{2,i,j}^{\epsilon, \sigma}). \end{array} \right. \quad (3.34)$$

**Remark 7** In the following, we shall simply call our Micro-Macro scheme (3.34), obtained with Arakawa space discretization and DIRK time discretization, the (DAMM)-scheme.

## 3.5 Verification of the AP-scheme in a mathematical test case

In this section we investigate the numerical properties of our asymptotic-preserving (DAMM)-scheme (3.34) for the resolution of (3.32) with given field  $\mathbf{b}$ . The section is devoted to the investigation of a linear case where the stream function  $\Psi$  does not depend on  $f^\epsilon$  and is static. In mind, we have as an application the Vlasov/Fokker-Planck equation with strong given magnetic field (3.3). Convergence results regarding the discretization parameters and numerical study of the asymptotic limit  $\epsilon \rightarrow 0$  are presented. Moreover, the influence of the stabilization parameter  $\sigma$  on the numerical results is discussed.

### 3.5.1 Analytical solution for both $\epsilon$ -regimes

Let us choose in this section the stationary stream function  $\Psi(x, y) = \frac{1}{2}(x^2 + y^2)$ , corresponding to  $\mathbf{b} = (y, -x)^T$ . In this case, we can compute explicitly the characteristics corresponding to the transport equation (3.8). Indeed, the characteristic curve  $\mathcal{C}_\epsilon^{t,x,y}(s) := (X(s), Y(s))$  passing at instant  $t$  through  $(x, y)$ , solves the ODE

$$\left\{ \begin{array}{l} \dot{X}(s) = \frac{Y(s)}{\epsilon}, \\ \dot{Y}(s) = -\frac{X(s)}{\epsilon}, \end{array} \right. \quad (X(t), Y(t)) = (x, y),$$

and has the explicit form

$$\mathcal{C}_\epsilon^{t,x,y}(s) := \begin{pmatrix} X \\ Y \end{pmatrix}(s; t, x, y) = \mathcal{R}\left(\frac{s-t}{\epsilon}\right) \begin{pmatrix} x \\ y \end{pmatrix}.$$

with the rotation matrix given by

$$\mathcal{R}(\theta) := \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}.$$

These characteristic curves are nothing else than spirals in the  $(t, x, y)$ -phase-space. All characteristics are  $2\pi\epsilon$ -periodic in  $s$ . The solution  $f^\epsilon$  of the advection equation (3.8) is now constant along these curves, such that

$$f^\epsilon(t, x, y) = f_{in}(X(0; t, x, y), Y(0; t, x, y)), \quad \forall (t, x, y) \in [0, T] \times \Omega,$$

which leads to the following analytic expression of the unique solution to (3.8)

$$f^\epsilon(t, x, y) = f_{in}\left(\cos\left(\frac{t}{\epsilon}\right)x - \sin\left(\frac{t}{\epsilon}\right)y, \sin\left(\frac{t}{\epsilon}\right)x + \cos\left(\frac{t}{\epsilon}\right)y\right). \quad (3.35)$$

Finally, the  $\epsilon \rightarrow 0$  limit solution  $f^0$  solves the problem (3.14), leading to

$$f^0 = \langle f_{in} \rangle = \frac{1}{2\pi} \int_0^{2\pi} f_{in}(R \cos(s), R \sin(s)) ds, \quad (3.36)$$

where  $R := \sqrt{x^2 + y^2}$ . Let us choose now as initial data a Gaussian peak not centered in the origin, i.e.

$$f_{in}(x, y) = \exp\left(-\frac{(x-0.5)^2 + (y-0.5)^2}{2\eta^2}\right), \quad \eta = 0.05, \quad L_x = L_y = 1,$$

and investigate how the numerical scheme is rendering its movement.

In the following, we shall denote by  $\Pi_h(f^\epsilon)$  resp.  $f_h^{\epsilon, \sigma}$  the exact solution (3.35) calculated on the grid mesh resp. the numerical solution of (3.32) obtained with our (DAMM)-scheme, and  $\Pi_h(f^0)$  refers to the exact limit solution (3.36) calculated on the grid mesh. Let  $\mathcal{Q} := (0, T) \times \Omega_S$ . We introduce also the numerical errors

$$\mathcal{L}_{\epsilon, \mathcal{X}}^p := \|\Pi_h(f^\epsilon) - f_h^{\epsilon, \sigma}\|_{L_h^p(\mathcal{X})}, \quad \mathcal{L}_{0, \mathcal{X}}^p := \|\Pi_h(f^0) - f_h^{\epsilon, \sigma}\|_{L_h^p(\mathcal{X})}, \quad (3.37)$$

where  $p \in \{1, 2, \dots, \infty\}$ ,  $\mathcal{X}$  stands for  $\mathcal{Q}$  or  $\Omega_S$ , and  $L_h^p$  denotes the discrete  $L^p$ -norm.

Figure 3.2 shows now the numerical solution obtained with the (DAMM)-scheme and the corresponding exact solution, in the non-limit ( $\epsilon = 1$ ) regime. The solutions related to the limit regime ( $\epsilon = 0$ ) are plotted in Figure 3.3. One observes that the numerical scheme we propose in this paper approximates well the exact solutions in both extreme regimes. In the next paragraphs we will try to prove more rigorously this visible correspondence and justify the choice of the stabilization parameter  $\sigma$ .

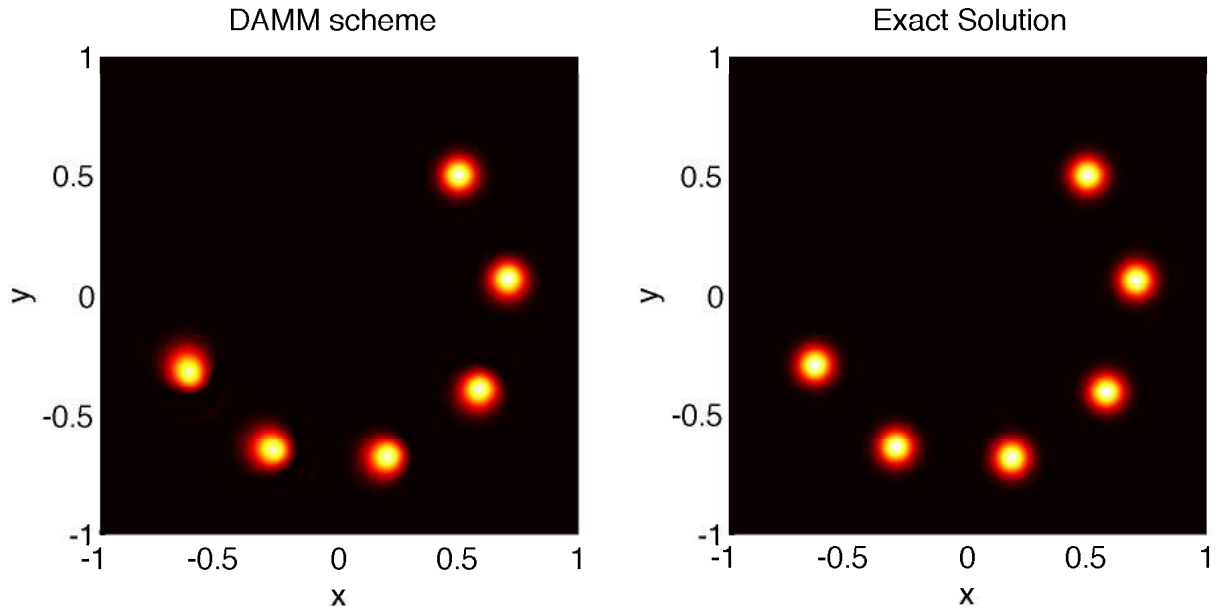


FIGURE 3.2 – (Non-limit case  $\epsilon = 1$ ). Superposition of the distribution function  $f_h^{\epsilon, \sigma}$  at several time steps  $t^n$  (left-panel) compared to the exact solution  $f^\epsilon$  (right-panel), with  $N_t = 2000$ ,  $T = 3.5$ ,  $N_x = N_y = 200$ ,  $\sigma = \Delta x^2$ .

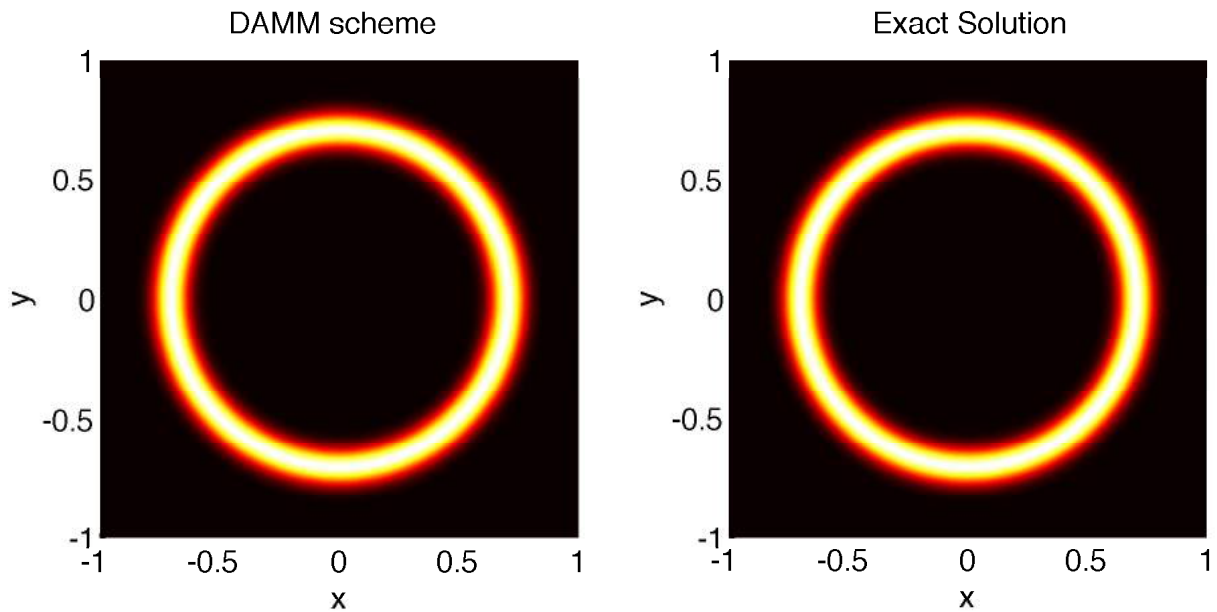


FIGURE 3.3 – (Limit case  $\epsilon = 0$ ). Plot of  $f_h^{0, \sigma}$  at final time  $T = 1$  with  $N_t = 200$  and  $\sigma = \Delta x$  (left-panel), compared to the limit exact solution  $f^0$  (right-panel). Mesh size :  $N_x = N_y = 200$ .

### 3.5.2 AP-property

Let us start by discussing the AP-property of our scheme. As a reference scheme, we decided to take a fully implicit DIRK-scheme with Arakawa space discretization, solving (3.8). In Figure 4.5 (a) we plot the condition number of the system matrix of the (DAMM)-

scheme and of the implicit reference scheme as a function of  $\epsilon$ . One observes that the condition number of the (DAMM)-scheme is bounded uniformly in  $\epsilon$ , whereas the implicit scheme is ill-conditioned in the limit  $\epsilon \rightarrow 0$ . This underlines one important advantage of our (DAMM)-scheme when compared with standard schemes, namely the fact that the (DAMM)-scheme does not degenerate in the limit  $\epsilon \rightarrow 0$ . From the right panel (b), however, it is evident that the condition number of the (DAMM)-scheme depends on the stabilization parameter  $\sigma$ . This reflects the fact that in the limit  $\sigma \rightarrow 0$ , the solution  $(f^{\epsilon,0,n}, q^{\epsilon,0,n})$  of (3.24) is not unique, and therefore the problem becomes ill-posed. As mentioned later on,  $\sigma$  has to be chosen not too small, such to have a reasonable condition number, and not too large, in order not to modify too much the problem. Concerning the numerical analysis, we do not present any result in this work, since we have proved in a previous publication (see [51]) that the Micro-Macro procedure coupled with a stabilization method permits to avoid an amplification of the truncation error with decreasing  $\epsilon$ .

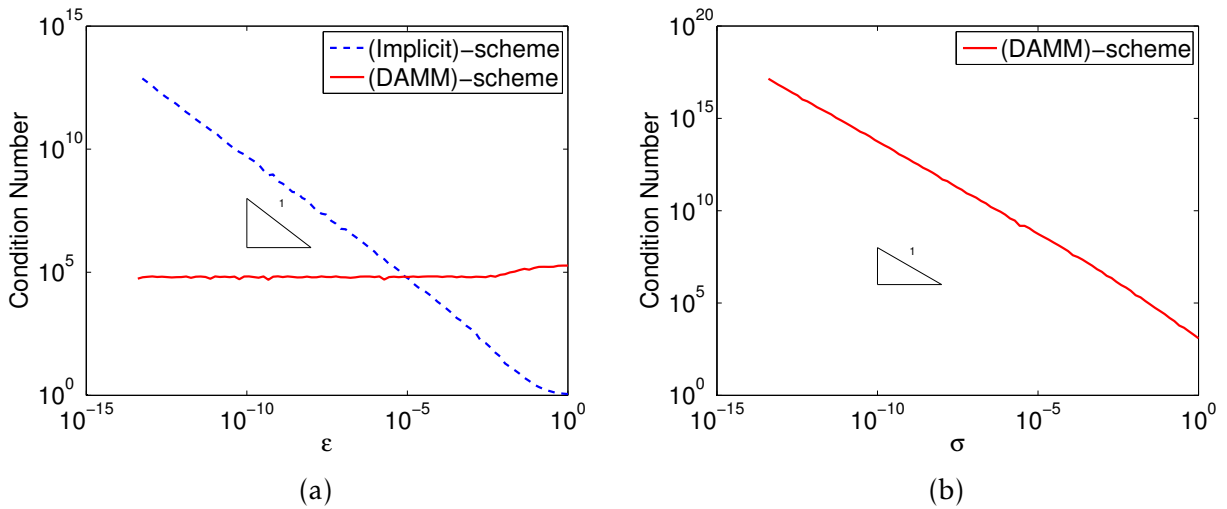


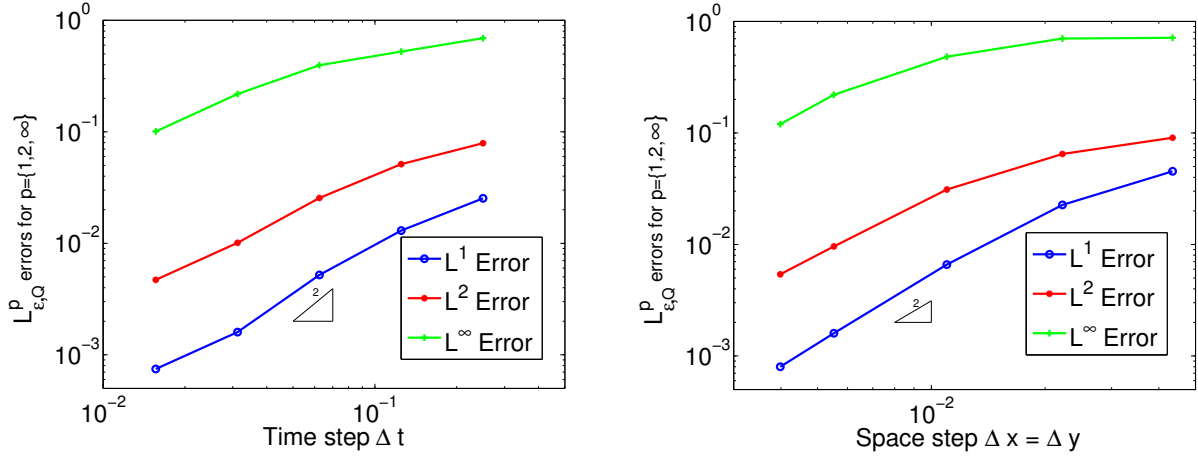
FIGURE 3.4 – Condition numbers as a function of  $\epsilon$  with  $\sigma = \Delta x^2$  for both (DAMM) and fully implicit schemes (a), and as a function of  $\sigma$  with  $\epsilon = 1$  for the (DAMM)-scheme (b). Here  $T = 1$ ,  $N_t = 200$ ,  $N_x = N_y = 50$ .

### 3.5.3 Convergence property

Next we study the convergence properties of the (DAMM)- scheme. Figure 4.6 displays the convergence rates in  $\Delta t$  and in  $\Delta x = \Delta y$ , obtained by comparison with the exact solutions (3.35) for  $\epsilon = 1$  and  $\sigma = \Delta x^2$ . In panel (a), one observes the expected second-order convergence rates with respect to time. In panel (b), the second-order convergence in space due to the Arakawa discretization of the Poisson brackets is evident. We observe from Table 3.1 that the convergence rate in space is even better for  $\epsilon = 0$ . Indeed, in this simple test-case, in the limit  $\epsilon \rightarrow 0$ , the distribution function is radial, namely  $f^{0,\sigma} = g^\sigma(x^2 + y^2)$ . Following [6], we observe that the leading terms of the truncation error, corresponding to the discretization of the Poisson bracket, vanish for that case.

### 3.5.4 Influence of the stabilization

Our next objective is to investigate the influence of the stabilization parameter  $\sigma$  on the quality of the numerical solution in the different  $\epsilon$ -regimes. Let us start the discussion with the limit regime  $\epsilon = 0$ . In Figure 3.6 (a), we plot the  $\mathcal{L}_{0,\Omega_\sigma}^1$  error (w.r.t. the exact limit sol.)



(a) Convergence w.r.t.  $\Delta t$ . Fixed mesh size :  $N_x = N_y = 350$ . (b) Convergence w.r.t.  $\Delta x = \Delta y$ . Fixed number of iterations :  $N_t = 700$ .

FIGURE 3.5 – Convergence studies for the (DAMM)-scheme : errors  $\mathcal{L}_{\epsilon, \Omega}^1$ ,  $\mathcal{L}_{\epsilon, \Omega}^2$ , and  $\mathcal{L}_{\epsilon, \Omega}^\infty$  as a function of  $\Delta t$  (a) and  $\Delta x = \Delta y$  (b), respectively. Parameters were  $T = 1$ ,  $\epsilon = 1$  and  $\sigma = \Delta x^2$ .

over time  $t$ , for two different choices of  $\sigma$ , namely  $\sigma = \Delta x$  and  $\sigma = \Delta x^2$ , and for several choices of  $\Delta x$ . One observes two phases : at first the numerical solution relaxes towards the (weak) analytic limit solution, until a plateau is reached and the error remains constant over time. This relaxation is faster for smaller values of  $\sigma$ , fact which can be explained by taking a look at the degenerate diffusion equation (3.22) we are effectively solving in the limit  $\epsilon \rightarrow 0$ . Smaller  $\sigma$  means stronger diffusion along the field lines of  $\mathbf{b}$ , which means that the number of iterations  $n_{eq}$  to reach the equilibrium plateau decreases with  $\Delta x$ , see Table 3.1. Observe also that the error in the equilibrium phase is the same for each  $\sigma = \Delta x^r$  with  $r \geq 1$ , only the relaxation rate strongly depends on the choice of  $r$ .

Let us mention briefly the computational time (CPU time) one needs for reaching the equilibrium plateau for the different cases studied in Figure 3.6 (a). The problem is that the condition number of the system matrix is inversely proportional to  $\sigma$ , as already demonstrated in Figure 4.5 (b). This bad conditioning would lead necessarily to an increase in CPU time for very small  $\sigma$  which has to be evaluated. For example, regarding the case  $\Delta x = 2/80$  (third curve in Figure 3.6 (a) and third line in Table 3.1), one obtains  $t_{CPU}(\sigma = \Delta x) = 53$  s in contrast to  $t_{CPU}(\sigma = \Delta x^2) = 104$  s. Thus, even if Figure 3.6 (a) suggests that a higher  $r$  would be more suitable to attain quickly the equilibrium plateau, this previous study about the CPU time advices us to be more careful and choose  $\sigma$  not too small.

In the regime  $\epsilon = 1$ , displayed in Figure 3.6 (b), the  $\mathcal{L}_{\epsilon, \Omega_S}^1$  error increases linearly with time for all choices of  $\Delta x$ . There is a very weak  $\sigma$  dependence in this regime as shown by the quasi superposition of the curves. The linear increase of the error is due to the dispersive character of the Arakawa discretization which leads to errors in the phase velocities. The not-dependence on  $\sigma$  is due to the fact that the term  $\epsilon \mathbf{b} \cdot \nabla q^\epsilon$  in the second equation of (3.20) is larger for  $\epsilon = 1$  than the term  $\sigma q^\epsilon$ .

Finally, we shall discuss an intermediate  $\epsilon$ -regime given by  $\epsilon = 0.1$ . We simulate  $n_{final} = 500$  time steps with  $\Delta t = 0.01$ , thus a final time of  $T = 5$  for  $\epsilon = 0.1$ , and choose  $\sigma = \Delta x$  with  $\Delta x \in \{2/40, 2/80, 2/160\}$ . In Figure 3.7, one observes the two previously discussed regimes for the  $\mathcal{L}^1$  errors. The dashed lines refer to the error  $\mathcal{L}_{\epsilon, \Omega_S}^1$  with respect to the exact solution

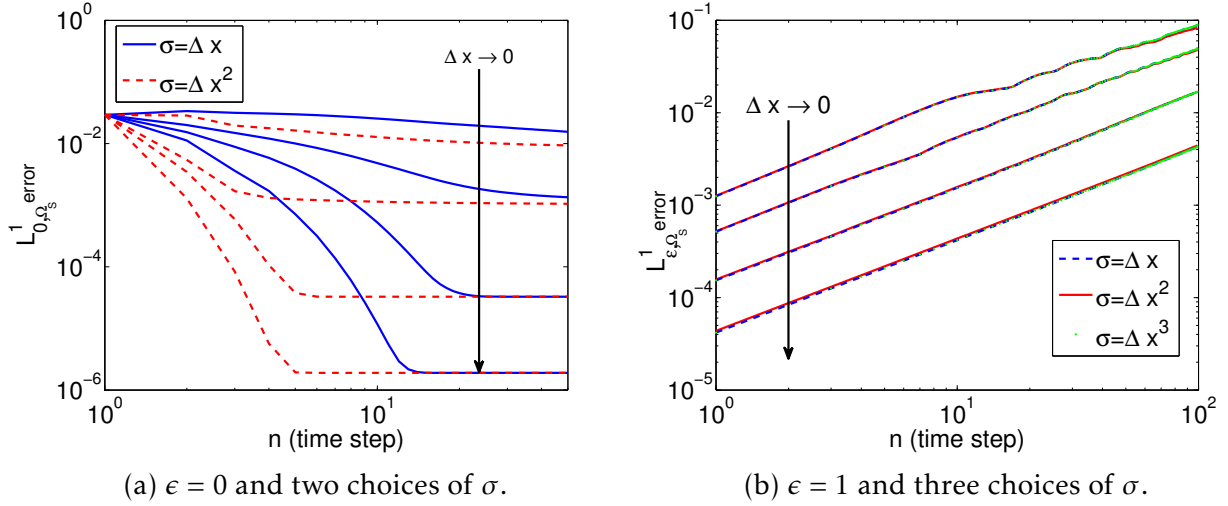


FIGURE 3.6 –  $\mathcal{L}^1_{0,\Omega_S}$  (panel (a)) and  $\mathcal{L}^1_{\epsilon,\Omega_S}$  (panel (b)) errors over time for the (DAMM)-scheme for two different  $\epsilon$ -regimes and several  $\Delta x$ .

TABLE 3.1 – Corresponding to Figure 3.6 (a) with  $\epsilon = 0$ .  $\mathcal{L}^1_{0,\Omega_S}$ -error at final time  $T = 2$  and  $\Delta t = 0.01$ ; number of iterations  $n_{eq}$  and CPU-time  $t_{CPU}$  for reaching the equilibrium plateau; condition nbr. CN of the linear system.

	$\sigma = \Delta x$				$\sigma = \Delta x^2$				
	$\mathcal{L}^1_{0,\Omega_S}$ at time $T$	$n_{eq}$	$t_{CPU}$	CN	$\mathcal{L}^1_{0,\Omega_S}$ at time $T$	$n_{eq}$	$t_{CPU}$	CN	
$\Delta x$	2/20	0.0106	> 200	3.1s	8.4e3	0.0082	175	2.8s	1.3e5
	2/40	0.0011	169	12s	2.6e4	0.0010	66	4.9s	8.5e5
	2/80	3.2737e-5	123	53s	9.1e4	3.2735e-5	10	104s	3.9e6
	2/160	1.8992e-6	19	133s	2.9e5	1.8992e-6	5	1302s	3.5e7

(3.35) whereas the solid lines show the error  $\mathcal{L}^1_{0,\Omega_S}$  w.r.t the limit solution (3.36). In the first phase of the simulation (short times), we recover Figure 3.6 (B), in particular the error  $\mathcal{L}^1_{0,\Omega_S}$  is large and  $\mathcal{L}^1_{\epsilon,\Omega_S}$  increases linearly. At a certain point, these two errors cross, and afterwards (at large times) the solution relaxes towards the weak limit solution (equilibrium plateau), and we recover Figure 3.6 (A), in particular, the error  $\mathcal{L}^1_{\epsilon,\Omega_S}$  is maximal at the end of the simulation, whereas  $\mathcal{L}^1_{0,\Omega_S}$  decreases towards the plateau. To understand Figure 3.7, one has to keep in mind that it corresponds to an intermediate  $\epsilon$ -value ( $\epsilon = 0.1$ ), and that the final simulation time is  $T = 5$ . The physical time is thus  $T_{phys} = 50$ , whereas in Figure 3.6 (B) the physical time was  $T_{phys} = 1$ , so Figure 3.7 combines intermediate  $\epsilon$ -values with long-times asymptotics, which explains why one recovers there the two regimes of Figure 3.6, in particular, the short-time regime (Figure 3.6 (B)) and the long-time regime (Figure 3.6 (A)). In between, the AP-scheme proceeds to the transition between "non-stiff" and "stiff" regimes.

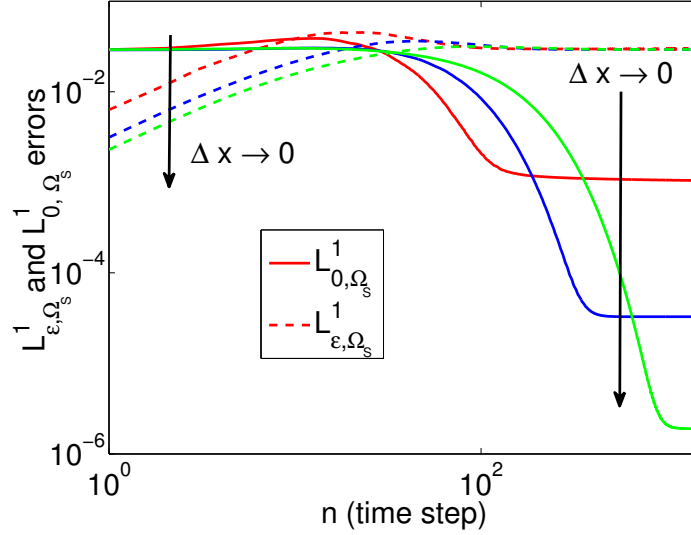


FIGURE 3.7 –  $\mathcal{L}_{\epsilon, \Omega_S}^1$  and  $\mathcal{L}_{0, \Omega_S}^1$  errors over time for the (DAMM)-scheme for several  $\Delta x$  with  $\epsilon = 0.1$ ,  $\sigma = \Delta x$ . Dashed lines refer to the error  $\mathcal{L}_{\epsilon, \Omega_S}^1$  whereas the solid lines show the error  $\mathcal{L}_{0, \Omega_S}^1$ .

### 3.5.5 Choice of the stabilization parameter $\sigma$

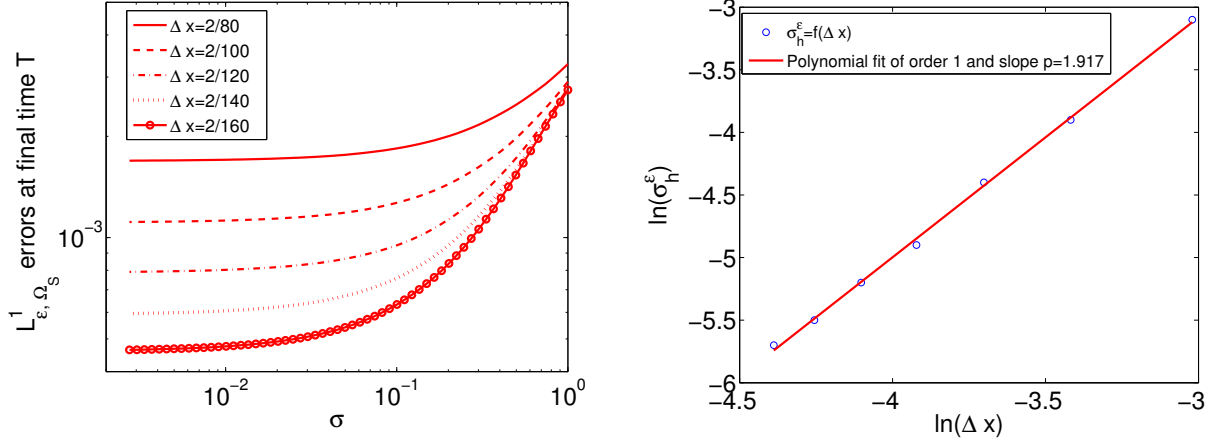
After having given some qualitative observations about the influence of the discretization parameter  $\sigma$  in different  $\epsilon$ -regimes, let us present some ideas to optimize the choice of  $\sigma$ . In Figure 3.8 (a), we plot the  $\mathcal{L}_{\epsilon, \Omega_S}^1$  error at the final time  $T$  for the non-limit regime  $\epsilon = 1$  as a function of  $\sigma$  for several values of  $\Delta x$ . In order to minimize the error  $\mathcal{L}_{\epsilon, \Omega_S}^1$ , the curves suggest to choose a  $\sigma$ -value depending on  $\Delta x$ . To investigate more precisely this dependence, we propose to take as "optimal"  $\sigma$ , for each fixed  $\Delta x$ , the value  $\sigma_h^\epsilon$  computed via

$$\sigma_h^\epsilon := \max \left\{ \sigma \in [\sigma_{min}, 1] \left| \frac{\|\Pi_h(f^\epsilon) - f_h^{\epsilon, \sigma}\|_{L_h^1(\Omega_S)} - \|\Pi_h(f^\epsilon) - f_h^{\epsilon, \sigma_{min}}\|_{L_h^1(\Omega_S)}}{\|\Pi_h(f^\epsilon) - f_h^{\epsilon, \sigma_{min}}\|_{L_h^1(\Omega_S)}} < \eta \right. \right\}, \quad (3.38)$$

(where  $\eta$  is an arbitrary precision) and evaluate how  $\sigma_h^\epsilon$  varies with  $\Delta x$ . In Figure 3.8 (b), we display  $\ln(\sigma_h^\epsilon)$  as a function of  $\ln(\Delta x)$  for  $\eta = 0.01$  and  $\sigma_{min} = 7e - 6$ . The data approach a polynomial line of slope  $p = 1.917$ , suggesting a relation between  $\sigma_h^\epsilon$  and  $\Delta x$  of the form  $\sigma_h^\epsilon = C \Delta x^p$  (with  $C > 0$ ), very close to the relation chosen in the last sections. To end this paragraph, let us briefly analyze the influence of  $\eta$ , the precision criterion, appearing in (3.38). For that, Table 3.2 presents for several values of  $\eta$ , the slope of the polynomial fitting of order 1 of the data  $\ln(\sigma_h^\epsilon) = f(\ln(\Delta x))$ , as well as its determination coefficient  $r^2$ , defined as the square of the correlation coefficient. One notes that the slope is around 2 in each case, validating the choice we have made before, for the stabilization parameter  $\sigma$  in the large  $\epsilon$ -regime, namely  $\sigma_h^\epsilon = (\Delta x)^2$ .

We perform a similar analysis in the limit regime  $\epsilon = 0$ . For that, we display in Figure 3.9 (a) the  $\mathcal{L}_{0, \Omega_S}^1$  error at the final time  $T$  as a function of  $\sigma$ , for several values of  $\Delta x$ . As in the non-limit  $\epsilon$ -regime, the curves suggest us to choose a  $\sigma_h$  depending on  $\Delta x$ . Let us define the application  $\mathcal{Z} : \sigma \mapsto \|\Pi_h(f^0) - f_h^{0, \sigma}\|_{L_h^1(\Omega_S)}$ . To study the dependence between  $\sigma$  and  $\Delta x$ , we choose for each  $\Delta x$  a  $\sigma_h^0$  defined by :





(a)  $\mathcal{L}^1_{\epsilon, \Omega_S}$  error at final time  $T$  as a function of  $\sigma$ . (b) Parameters were  $\eta = 0.01$ ,  $\sigma_{min} = 7e - 6$ .

FIGURE 3.8 – (Non-limit case  $\epsilon = 1$ ). Study of the influence of  $\sigma$  in the  $\mathcal{L}^1_{\epsilon, \Omega_S}$  error (a) and of the relation between  $\sigma_h^\epsilon$  and  $\Delta x$  (b). Time discretization :  $T = 0.1$  and  $\Delta t = 0.01$ .

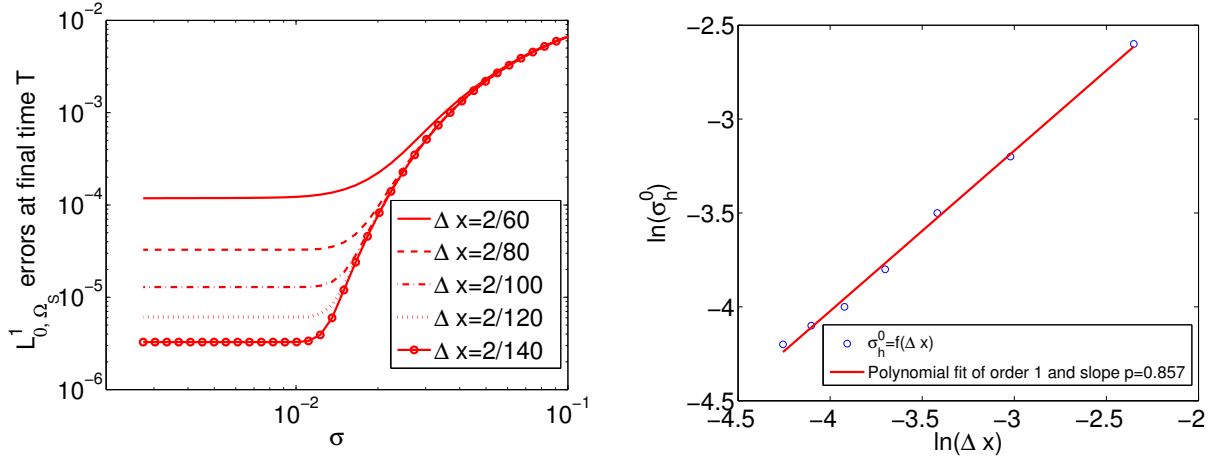
TABLE 3.2 – Polynomial fitting of order 1 for the data  $\ln(\sigma_h^\epsilon) = f(\ln(\Delta x))$  for several values of the precision  $\eta$ . In each case, the slope  $p$  of the line and the determination coefficient  $r^2$  are written.

$\eta$	$p$	$r^2$
$1e - 1$	1.974	0.9975
$5e - 2$	1.974	0.9975
$1e - 2$	1.9175	0.9988
$5e - 3$	1.8675	0.9988
$1e - 3$	1.8597	0.9989

$$\sigma_h^0 := \arg \max_{\sigma \in [\sigma_{min}, 1]} \left| \frac{dZ}{d\sigma} \right| (\sigma), \quad (3.39)$$

and we plot in Figure 3.9 (b) the evolution of  $\ln(\sigma_h^0)$  as a function of  $\ln(\Delta x)$ . The corresponding data follow a linear relation, with a slope of  $p = 0.857$ , meaning that  $\sigma$  can be chosen as  $\sigma = \Delta x$  in the limit regime  $\epsilon = 0$  with the aim to reduce the  $\mathcal{L}^1_{0, \Omega_S}$  error and avoid a bad condition number.

To conclude this first numerical part, one can say that this simple test case permits to make a deep analysis of the (DAMM)-scheme. In particular, the AP behavior of the scheme was confirmed, the orders of convergence in both space and time were confirmed, and the influence of  $\sigma$  as well as its delicate choice have been intensively investigated. This study was enabled by the existence of analytic solutions of the problem, rigorously compared to solutions obtained by the (DAMM)-scheme for several  $\epsilon$ -regimes. Thanks to this verification, the (DAMM)-scheme can be used to resolve more complicated models where no analytic solutions are at hand. This is the topic of the next part.



(a)  $\mathcal{L}^1_{0, \Omega_S}$  error at final time  $T$  as a function of  $\sigma$ .

(b) One chooses  $\sigma_{\min} = 7e - 6$ .

FIGURE 3.9 – (Limit case  $\epsilon = 0$ ). Study of the influence of  $\sigma$  in the  $\mathcal{L}^1_{0, \Omega_S}$  error (a) and of the relation between  $\sigma_h^0$  and  $\Delta x$  (b). Time discretization :  $T = 0.1$  and  $\Delta t = 0.01$ .

## 3.6 Numerical simulations for the Vlasov-Poisson test case

The aim of this section is dual : firstly to solve numerically the Vlasov-Poisson system (3.5) using the (DAMM)-scheme and to simulate some particular physical phenomena (such as the Landau damping or the two-stream instability); and secondly to study the long-time asymptotics  $\epsilon \rightarrow 0$  of the two-stream instability. Note that the literature on the Vlasov-Poisson system is very rich, some theoretical as well as numerical results can be found in the non-exhaustive list [20, 33, 59, 65].

### 3.6.1 The Vlasov-Poisson system and its numerical discretization

In this chapter, we set  $\mathcal{Q} := (0, T) \times (0, L_x) \times (-L_v, L_v)$ . Using the Poisson bracket, the Vlasov-Poisson 1D1V system verified by  $f^\epsilon := f^\epsilon(t, x, v)$  reads

$$(VP)^\epsilon \begin{cases} \partial_t f^\epsilon + \frac{1}{\epsilon} \{f^\epsilon, \Psi^\epsilon\} = 0, \\ -\partial_{xx} \varphi^\epsilon(t, x) = 1 - n^\epsilon(t, x), \end{cases} \quad (3.40)$$

where  $\Psi^\epsilon(t, x, v) := v^2/2 - \varphi^\epsilon(t, x)$  is the stream-function and  $n^\epsilon(t, x) := \int_{\mathbb{R}} f^\epsilon(t, x, v) dv$  denotes the electron density. Due to the fact that this problem is non-linear (unlike the previous one), its study is a more delicate task, and one should think of a linearization procedure.

Following the same reformulation as before, we can construct an Asymptotic-Preserving scheme for the Vlasov-Poisson system by introducing an auxiliary variable  $q^\epsilon(t, x, v)$ . The AP-reformulation of (3.40) is then discretized with the help of the (DAMM)-scheme as before. The determination of the electric field  $E^\epsilon(t, x)$  is guaranteed by the resolution of the discrete Poisson equation. The fully discretized (first order in time) reformulated Vlasov-Poisson system is summarized here for clarity. For each time step  $n$ , one is looking for

$(f_h^{\epsilon,\sigma,n+1}, q_h^{\epsilon,\sigma,n+1})$ , by iterating in  $l \in \mathbb{N}$  like

$$(RVP)_h^{\epsilon,\sigma,n,l} \begin{cases} -\frac{\varphi_{i+1}^{\epsilon,\sigma,n+1,l} - 2\varphi_i^{\epsilon,\sigma,n+1,l} + \varphi_{i-1}^{\epsilon,\sigma,n+1,l}}{\Delta x^2} = 1 - \Delta v \sum_{j=1}^{N_v} f_{i,j}^{\epsilon,\sigma,n+1,l}, \\ \Psi_{i,j}^{\epsilon,\sigma,n+1,l} = \frac{1}{2} v_j^2 - \varphi_i^{\epsilon,\sigma,n+1,l}, \\ f_{i,j}^{\epsilon,\sigma,n+1,l+1} + \Delta t [q_h^{\epsilon,\sigma,n+1,l+1}, \Psi_h^{\epsilon,\sigma,n+1,l}]_{i,j} = f_{i,j}^{\epsilon,\sigma,n}, \\ [f_h^{\epsilon,\sigma,n+1,l+1}, \Psi_h^{\epsilon,\sigma,n+1,l}]_{i,j} = \epsilon [q_h^{\epsilon,\sigma,n+1,l+1}, \Psi_h^{\epsilon,\sigma,n+1,l}]_{i,j} - \sigma q_{i,j}^{\epsilon,\sigma,n+1,l+1}, \end{cases} \quad (3.41)$$

and starting from

$$f_{i,j}^{\epsilon,\sigma,n+1,0} := f_{i,j}^{\epsilon,\sigma,n}.$$

This iterative procedure has been performed to cope with the non-linearity of the Vlasov-Poisson system. In the following simulations, the stopping criterion for these iterations (at  $l = l_f$ ) is

$$\frac{\|f_{i,j}^{\epsilon,\sigma,n+1,l+1} - f_{i,j}^{\epsilon,\sigma,n+1,l}\|_{L_h^1}}{\|f_{i,j}^{\epsilon,\sigma,n+1,l}\|_{L_h^1}} + \frac{\|\varphi_i^{\epsilon,\sigma,n+1,l+1} - \varphi_i^{\epsilon,\sigma,n+1,l}\|_{L_h^1}}{\|\varphi_i^{\epsilon,\sigma,n+1,l}\|_{L_h^1}} < 10^{-2},$$

and we finish by posing

$$f_{i,j}^{\epsilon,\sigma,n+1} := f_{i,j}^{\epsilon,\sigma,n+1,l_f+1}.$$

**Remark 8** As briefly mentioned at the beginning of Section 3.2, in the Vlasov-Poisson test-case, the vector  $\mathbf{b} = {}^\perp \nabla \Psi^\epsilon$  is time-dependent since  $\Psi^\epsilon$  depends on  $f^\epsilon$  via the Poisson equation. Nevertheless, the fixed point iteration procedure that we have employed in (3.41) permits to linearize the problem and during an iteration  $l \rightarrow l+1$  the stream-function  $\Psi^\epsilon$  is considered fixed, given by the previous iteration.

**Remark 9** Note that we wrote the previous system (3.41) without the DIRK time discretization in order to simplify its writing, however the following simulations have been implemented with the (DAMM)-scheme, including a DIRK time discretization.

### 3.6.2 Numerical simulations for weak Landau damping and $\epsilon = 1$

In order to validate our numerical procedure, we are interested in the Landau damping, for which analytic results are at hand. The Landau damping represents the exponential decrease of the electric field energy as a function of time (see [89, 128] for more details). For these simulations, the following initial data (see for example [33, 55]) is considered :

$$f_{in}(x, v) = \frac{1}{\sqrt{2\pi}} (1 + \gamma \cos(kx)) e^{-v^2/2}, \quad (3.42)$$

where  $\gamma$  refers to the amplitude and  $k$  to the mode of the perturbation of the homogeneous equilibrium  $\mathcal{M}(v) = (2\pi)^{-1/2} e^{-v^2/2}$ . In the following simulations, we take  $L_v = 10$ ,  $L_x = 2\pi/k$ ,  $\epsilon = 1$  and  $\sigma = (\Delta x/L_x)^2$ . In this section, we investigate the weak Landau damping, choosing a low amplitude of perturbation  $\gamma$ . According to [117], the weak Landau damping manifests for times  $t < 1/\sqrt{\gamma}$ . Beyond this time, the non-linear effects begin to be significant. Thus, we resolve the Vlasov-Poisson system (3.41) with the above initial

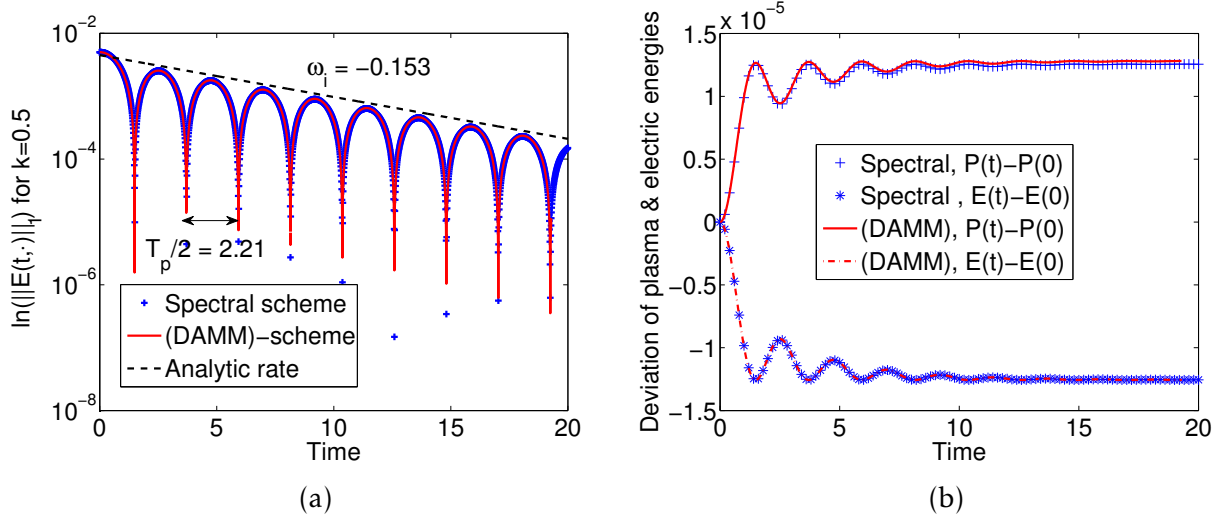


FIGURE 3.10 – (Weak Landau damping for  $\epsilon = 1$ ).  $L^1$ -norm of the electric field (in log-scale) versus time (a) and deviations of plasma and electric energies versus time (b) for both (DAMM) and spectral schemes.  $N_x = N_y = 256$ ,  $\Delta t = 0.01$ ,  $T = 20$ ,  $\sigma = (\Delta x/L_x)^2$ .

condition (3.42) for  $\gamma = 0.001$  and  $k = 0.5$ . To simplify the notation, we shall denote in the following simply by  $f^\epsilon$  our numerical solution obtained by the (DAMM)-scheme.

In Figure 3.10 (a) we represent the evolution in time of the  $L^1$ -norm of the electric field  $\|E^\epsilon(t, \cdot)\|_1$  (in log-scale) obtained from the (DAMM)-scheme. So as to validate efficiently our (DAMM)-scheme, we plot in the same Figure 3.10 (a) the corresponding evolution with a reference spectral scheme which resolves the system (3.40). The curves obtained from the two numerical schemes coincide perfectly. Moreover, we pay attention to the damping rate  $\omega_i$  and the frequency of oscillations  $\omega_p$ , which depend on the perturbation mode  $k$ . Under certain approximations, several formulae of these latter can be found (see for example [106]). One sees that both schemes approach the analytic values (for  $k = 0.5$ ), namely  $\omega_i = -0.153$  and  $\omega_p = 2\pi/T_p = 1.415$ . In Figure 3.10 (b), we plot the deviation (from their initial value) of both electric and plasma energies. These latter are defined as  $\mathcal{E}^\epsilon(t) := 1/2 \int_{-L_x}^{L_x} |E^\epsilon(t, x)|^2 dx$  and  $\mathcal{P}^\epsilon(t) := 1/2 \int_{\mathbb{R}} \int_{-L_x}^{L_x} v^2 f^\epsilon(t, x, v) dx dv$ . The curves indicate clearly that the total energy  $\mathcal{T}^\epsilon := \mathcal{E}^\epsilon + \mathcal{P}^\epsilon$  is conserved in compliance with the theory, for both (DAMM) and spectral schemes. Thus, the weak Landau damping is well simulated by the (DAMM)-scheme.

In Figure 3.11 we displayed the distribution function in phase-space at time  $t = 0$  (panel (a)),  $t = 20$  (panel (c)) and  $t = 40$  (panel (e)). Note that we have plotted the perturbed part of the distribution function  $f^\epsilon$ , meaning  $f^\epsilon - \mathcal{M}(v)$ . In the panels (b), (d), and (f), we represented the cross-sections at  $x = L_x/2$  of the previous plots, at the same times. These figures show us the continuous filamentation of  $f^\epsilon$  over time.

### 3.6.3 Numerical simulations for strong Landau damping and $\epsilon = 1$

We shall perform now the numerical simulations for the non-linear Landau damping by taking a stronger perturbation as in the previous study. Nevertheless, we stay in the non-limit regime  $\epsilon = 1$ . In Figure 3.12, we plot the distribution function  $f^\epsilon(t, x, v)$  at different times, with  $\gamma = k = 0.3$  and the initial data (3.42). Three levels can be pointed out. Up to  $t = 10$  (panel (a)), the linear effects dominate and the behavior of the electric energy is very

close to the linear case. Then, starting from  $t = 20$  (panel (b)), the damping is stopped due to particle trapping, for finally leading to saturation at around  $t = 40$  (panel (d)). The phase-space trapping holes are clearly visible. In Figure 3.13, we plot the space average of the distribution function at several times. The formation of several plateaus is clearly visible at time  $t = 10$  (panel (b)), indicating the trapping of particles in these areas. Over time, this trapping persists, although the numerical diffusion tends to damp these states. Indeed, due to the numerical dissipation, the filamentation is progressively eliminated when the filamentation scale becomes smaller than the velocity grid  $\Delta v$ .

In Figure 3.14 (a), we plot the evolution of the electric energy (in log-scale) as a function of time. Contrary to the weak Landau damping, the growth or decay rates of the oscillations are not known. Nevertheless, we can compare the (DAMM)-scheme to the reference spectral scheme. We observe a good correspondence between these two schemes. In order to carry on the investigations of the strong Landau damping, we look at the evolution of some particular quantities. The Vlasov-Poisson system is well-known to conserve the total particle number (mass), the momentum, the total energy, the  $L^p$ -norms and the entropy. These quantities are given respectively by  $(\Omega_x := (0, L_x))$

$$M^\epsilon(t) := \int_{\mathbb{R} \times \Omega_x} f^\epsilon(t, x, v) dx dv, \quad (3.43)$$

$$\mathcal{M}_0^\epsilon(t) := \int_{\mathbb{R} \times \Omega_x} v f^\epsilon(t, x, v) dx dv, \quad (3.44)$$

$$\mathcal{T}^\epsilon(t) := \mathcal{E}^\epsilon + \mathcal{P}^\epsilon, \quad (3.45)$$

$$C_p^\epsilon(t) := \left( \int_{\mathbb{R} \times \Omega_x} |f^\epsilon(t, x, v)|^p dx dv \right)^{1/p}, \quad (3.46)$$

$$S^\epsilon(t) := \int_{\mathbb{R} \times \Omega_x} -f^\epsilon(t, x, v) \ln(f^\epsilon(t, x, v)) dx dv. \quad (3.47)$$

Due to the presence of the stabilization parameter  $\sigma$ , the conserved quantities introduced previously are no more constant over time when computed via the (DAMM)-scheme. We investigated this in the panels of Figure 3.14. In particular in the panel (b) we see that the total energy is not conserved by the (DAMM)-scheme with 3% of deviation from its initial value. Analogous observations can be done for the mass (panel (c)), the entropy (panel (d)) and the  $L^2$ -norm (panel (f)). Nevertheless, the (DAMM)-scheme conserves the momentum (which is null for the initial condition (3.42)), unlike the spectral scheme. Despite the non-conservation of these quantities, their deviations from their initial value stay weak.

Having carefully considered the Landau damping through several numerical simulations performed by the (DAMM)-scheme, we are interested now in the study of the two-stream instability. Since the Landau damping does not attain an equilibrium (due to the continuous filamentation), it is not suitable for investigating the limit regime  $\epsilon \rightarrow 0$ . As we will see, things are different in the case of the two-stream instability, which permit to investigate the limit  $\epsilon \rightarrow 0$ .

### 3.6.4 Numerical simulations of the two-stream instability, study of the limit case $\epsilon \rightarrow 0$ .

#### Two-stream instability in the non-limit case $\epsilon = 1$

The two stream instability can be thought of as the inverse of the Landau damping. It occurs when the velocity of the particles is slightly greater than the wave velocity  $v_\phi$ . The instability causes a transfer of energy, from the particles to electric field, unlike the Landau damping phenomenon where the exchange of energy occurs from the electric field to the particles. Thus, to simulate this instability, one imposes the following initial data :

$$f_{in}^1(x, v) = \frac{1}{\sqrt{2\pi}} (1 + \gamma \cos(kx)) \frac{e^{-(v-3)^2/2} + e^{-(v+3)^2/2}}{2}. \quad (3.48)$$

As a first step, we keep  $\epsilon = 1$ . We choose  $\gamma = 0.001$ ,  $k = 0.2$ ,  $L_v = 10$ ,  $L_x = 2\pi/k$  and  $\sigma = (\Delta x/L_x)^2$ . In Figure 3.15, we plot the distribution function  $f^\epsilon(t, x, v)$ , solution of (3.40) at different times, with the previous initial condition (3.48). The panels (a) and (c) refer to the (DAMM)-scheme whereas the panels (b) and (d) correspond to the reference spectral scheme. In both cases, the instability grows until the non-linear effects become significant. Over time, the non-linear effects cause a trapping phenomenon. To push ahead with the investigations, we plot in Figure 3.16 the evolution over time of  $\ln(\|E^\epsilon(t, \cdot)\|_1)$  (a). The analytic value of the growth rate for the electric field, *i.e.*  $\omega_i(k = 0.2) = 0.2548$ , is very close to the numerical value observed and the curves obtained via the two numerical schemes (DAMM and spectral) coincide. As we made for the non-linear Landau damping, we examine in Figure 3.16 the conservation of several quantities over time, for both (DAMM) and spectral schemes. As for the non-linear Landau damping, only the momentum (panel (e)) is well conserved by the (DAMM)-scheme, the total energy (panel (b)), the mass (panel (c)), the entropy (panel (d)) and the  $L^2$ -norm (panel (f)) indicate weak deviations from their initial value. Nevertheless, we will see in the next section the main advantages of the (DAMM)-scheme, when compared to standard schemes.

#### Two-stream instability in the limit regime $\epsilon \rightarrow 0$

In order to conclude this section, we show that the AP-property of the (DAMM)-scheme can be useful when an equilibrium is reached. Recall that in the case of the Vlasov-Poisson system (3.40), passing to  $\epsilon \rightarrow 0$  is equivalent to passing to  $t \rightarrow \infty$ . Thus, the (DAMM)-scheme seems suitable to study the long-time behavior of the non-linear two-stream instability. In this part, we modify the initial condition, taking

$$f_{in}^2(x, v) = \frac{1}{\sqrt{2\pi}} v^2 e^{-v^2/2} (1 + \gamma \cos(kx)). \quad (3.49)$$

Although no rigorous proofs exist, the two-stream instability leads (in a certain weak sense) to a BGK (Bernstein-Greene-Kruskal) equilibrium after the growth phase. In Figure 3.17, the qualitative behavior of such equilibrium is visible. We have plotted the initial condition (3.49) in the panel (a), then we resolve the Vlasov-Poisson system with the (DAMM)-scheme for  $\epsilon = 0$ ,  $\sigma = (\Delta x/L_x)^2$ ,  $\Delta t = 0.01$ ,  $N_x = N_y = 256$ ,  $L_x = 2\pi/k$ ,  $L_v = 5$ , and the initial condition (3.49), with  $k = 0.5$  and  $\gamma = 0.05$ . From the first time iteration  $n = 3$  (panel (b)), the equilibrium seems to be attained and the filamentations are smoothed out. Note the formation of the separatrix which connects the saddle points at  $v = 0$  and  $x = 0 = 4\pi$ . Due to the topological conservation of the Vlasov-Poisson equation (see [74]),

the distribution function keeps over time the nature and the number of its extrema. Panels (c) and (d) represent the distribution function  $f^\epsilon$  at the time iteration  $n = 15$  and  $n = 50$ , respectively. Note that the separatrix, clearly visible in the panel (b), is progressively smoothed out due to the numerical dissipation of the scheme. Besides, the value of the central extremum in  $(2\pi, 0)$  remains essentially constant in time. Thus, the (DAMM)-scheme conserves the nature and the position of this latter, meaning that the particle trapping is well-reproduced by our scheme. In Figure 3.18, we have plotted the contours of the distribution function  $f^\epsilon$  at the same times. We see clearly in the center the particle trapping on the panels (b), (c) and (d).

In order to confirm this BGK saturation, we shall check if the contours of the distribution function  $f^0$  are aligned with the contours of the stream-function  $\Psi^0 = v^2/2 - \varphi^0$ , as one expects that, in the limit  $\epsilon \rightarrow 0$ ,  $f^0$  depends only on  $\Psi^0$ . Thanks to the AP-property of our (DAMM)-scheme, we can obtain this equilibrium with a very low numerical cost, without too much numerical pollution. Few iterations are effectively needed to reach this equilibrium. To put into evidence the dependence  $f^0(\Psi^0)$ , we use the fitting proposed by Heath and al [74], namely

$$f_{fit}^0 = a(\Psi^0 + \varphi_M)(\Psi^0 + \Psi^\star) e^{-\beta\Psi^0}, \quad (3.50)$$

where  $a$  and  $\beta$  are fitting parameters to be found numerically,  $\varphi_M$  is the maximum of  $\varphi^0$  and  $\Psi^\star$  is defined by

$$\Psi^\star = \frac{\varphi_M - \beta\Psi_M\varphi_M + 2\Psi_M - \beta\Psi_M^2}{\beta\varphi_M + \beta\Psi_M - 1},$$

where  $\Psi_M$  is the value at which  $f^0$  attains its maximum. We choose  $a = 0.2948$  and  $\beta = 1.20$ . From the numerical simulation, we extract  $\varphi_M = 0.60$ ,  $\Psi_M = 0.93$  and thus  $\Psi^\star = 0.90$ . In Figure 3.19, we plot the evolution of  $f^0(\Psi^0)$  as compared to the fitting distribution  $f_{fit}^0$ . Panel (a) represents  $f^0(\Psi^0)$  at time  $t = 0$ , clearly, there is no alignment between  $f^0$  and  $\Psi^0$ , as expected. The panels (c) and (e) which zoom the panel (a) in two regimes confirm this affirmation, we see clearly the non-functional structure of the plot (multi-valued function). However, in panel (b), we plot the same evolution but after fifty time iterations. One notes a very good correspondence between the numerical curve and the fitting one. The panels (d) and (f) show a good alignment of the points, showing that a BGK equilibrium is attained. Nevertheless, we observe an anormal inflexion of the curve near to the point  $f^0(\Psi^0 = 0)$ . A similar phenomenon was observed in [28], with a different numerical scheme. In the panel (f), we examine  $f^0(\Psi^0)$  near to its minimum. We pay attention here that there is no splitting phenomenon, confirming that the saturation is totally achieved.

To summarize, the (DAMM)-scheme permits, by passing to the limit  $\epsilon \rightarrow 0$ , to obtain a BGK equilibrium with a low number of iterations, permitting to control the accumulation of the errors. This is an essential advantage, as compared to standard schemes.

### 3.7 Concluding remarks and perspectives

The long-time behavior of the Vlasov-Poisson system is a challenging problem, requiring some investigations. Numerically, it is arduous to obtain a solution avoiding numerical pollution in such time asymptotics. We have developed an asymptotic-preserving scheme, based on a micro-macro decomposition coupled with a stabilization procedure

in order to limit this difficulty. The analysis of the two-stream instability has shown the remarkable properties of the (DAMM)-scheme, permitting to attain a BGK-like equilibrium in few time-step iterations with low numerical costs and small errors. Nevertheless, the (DAMM)-scheme could be improved, notably through its stabilization part. The circle test case helped us a lot to better understand the choice of the stabilization parameter. But the Vlasov-Poisson test case shows that this parameter alters the conservation properties of the system. One may imagine for a future work to replace the stabilization parameter by a more general operator which could improve the conservation properties of the Vlasov-Poisson equation.



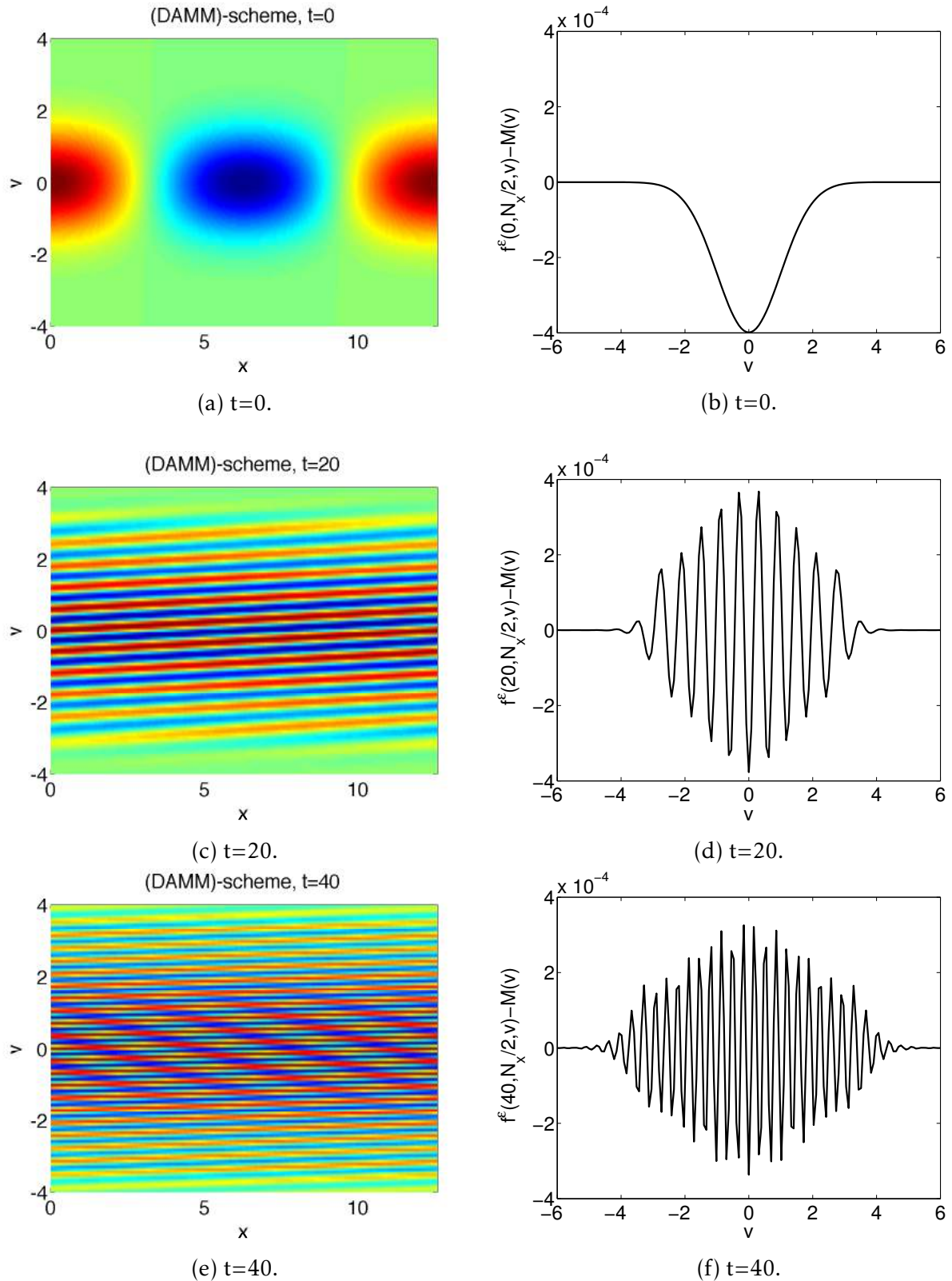


FIGURE 3.11 – (Weak Landau damping for  $\epsilon = 1$ ) Zoom of the distribution function  $f^\epsilon(t, x, v) - M(v)$  at different times with  $k = 0.5$  and  $\gamma = 0.001$ , obtained with the (DAMM)-scheme (a), (c) and (e) and cross-sections at  $x = L_x/2$  of this latter (b), (d) and (f). Mesh size :  $N_x = N_y = 256$ . Other parameters were :  $\Delta t = 0.01$ ,  $T = 40$ ,  $\epsilon = 1$  and  $\sigma = (\Delta x/L_x)^2$ .

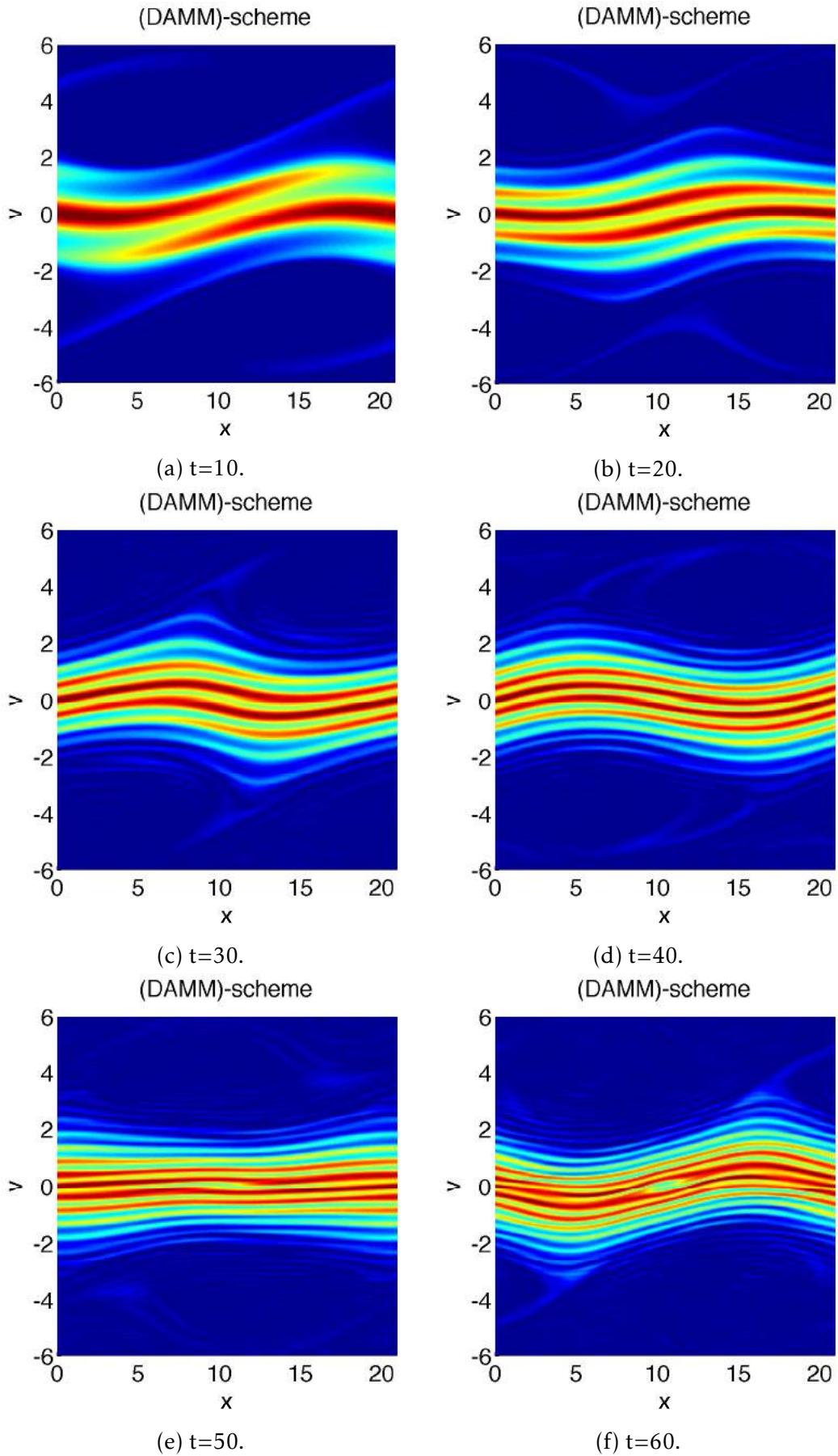


FIGURE 3.12 – (Strong Landau damping for  $\epsilon = 1$ ) Zoom of the distribution function  $f^\epsilon(t, x, v)$  at different times with  $k = \gamma = 0.3$ , obtained with the (DAMM)-scheme. Mesh size :  $N_x = N_v = 256$ . Parameters were :  $\Delta t = 0.01$ ,  $T = 60$ ,  $\epsilon = 1$  and  $\sigma = (\Delta x/L_x)^2$ .

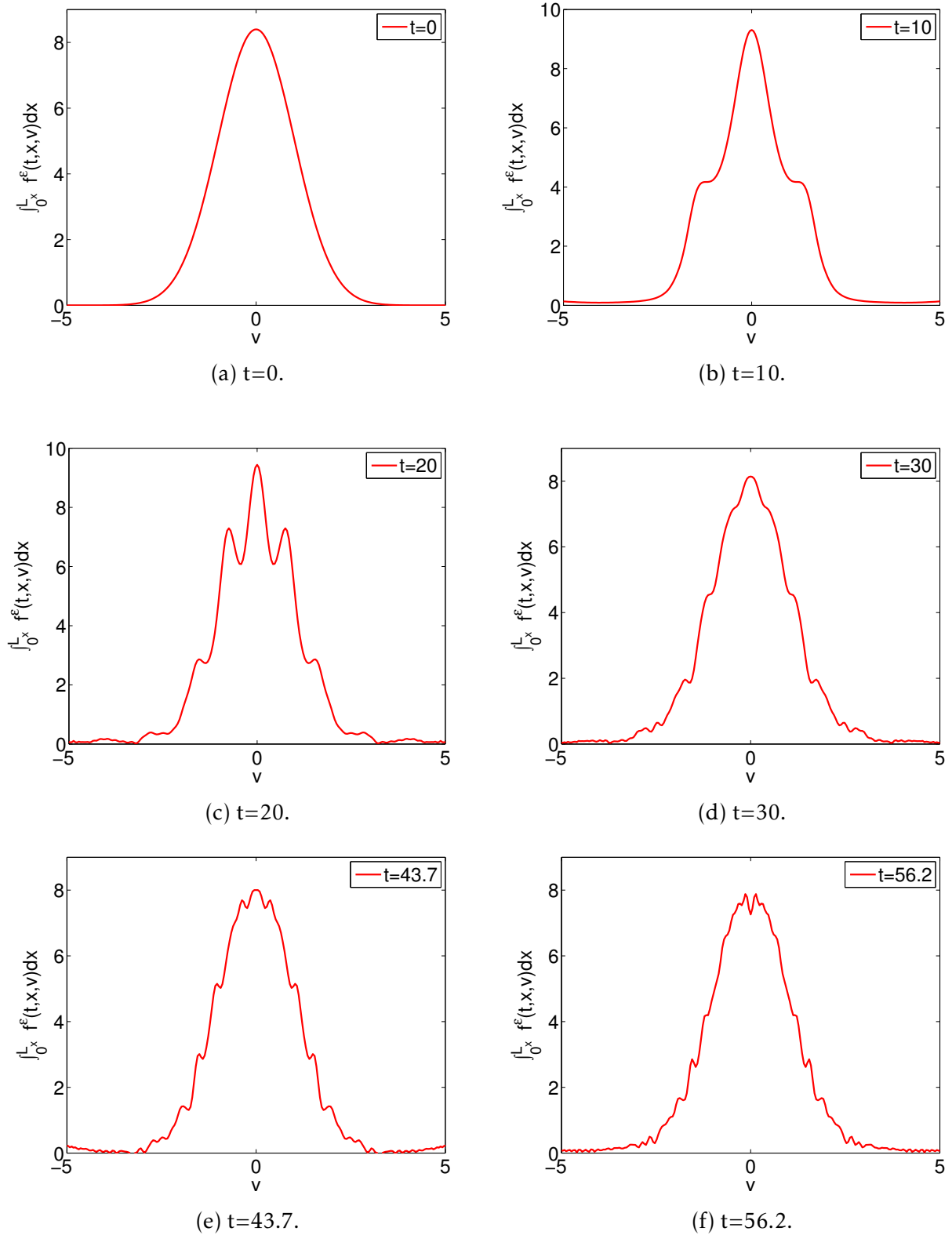
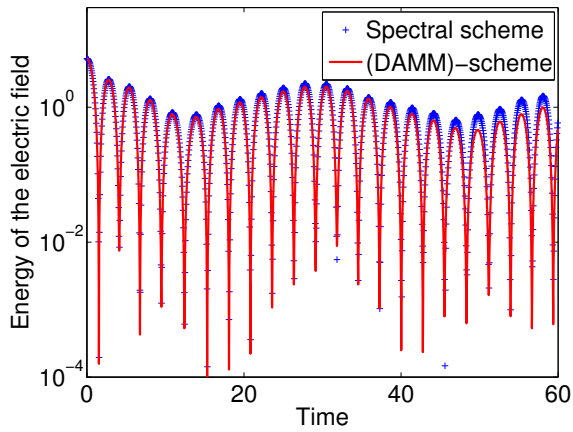
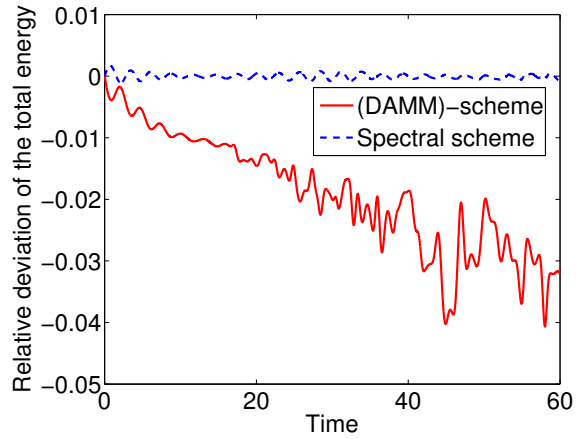


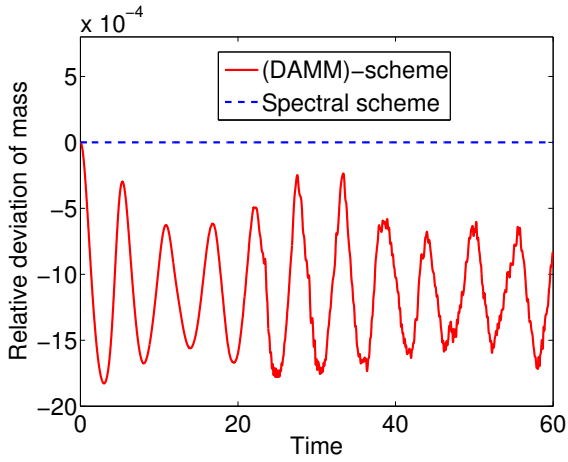
FIGURE 3.13 – (Strong Landau damping for  $\epsilon = 1$ ) Spatial average of the distribution function  $f^\epsilon(t, x, v)$  at different times with  $k = \gamma = 0.3$ , obtained with the (DAMM)-scheme. Mesh size :  $N_x = N_y = 256$ . Parameters were :  $\Delta t = 0.01$ ,  $T = 60$ ,  $\epsilon = 1$  and  $\sigma = (\Delta x/L_x)^2$ .



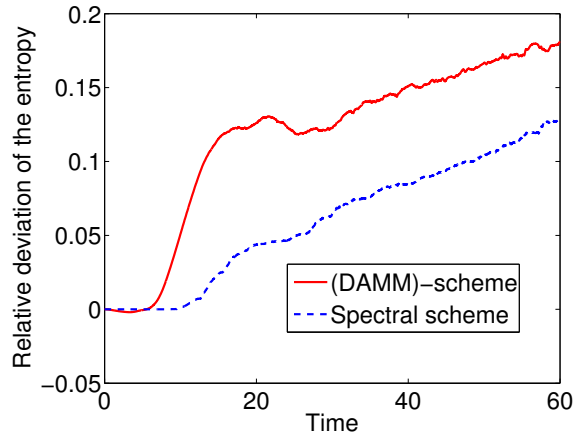
(a) Strong Landau damping.



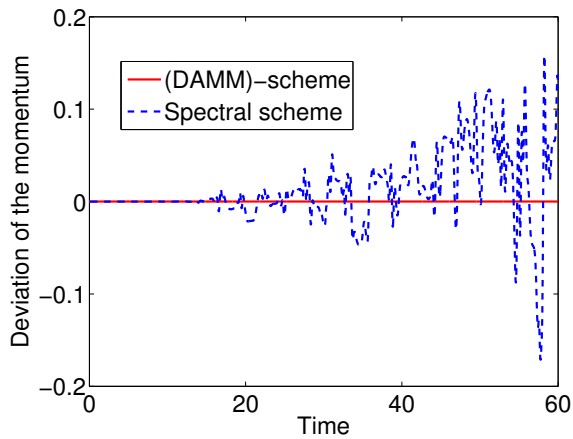
(b) Total energy.



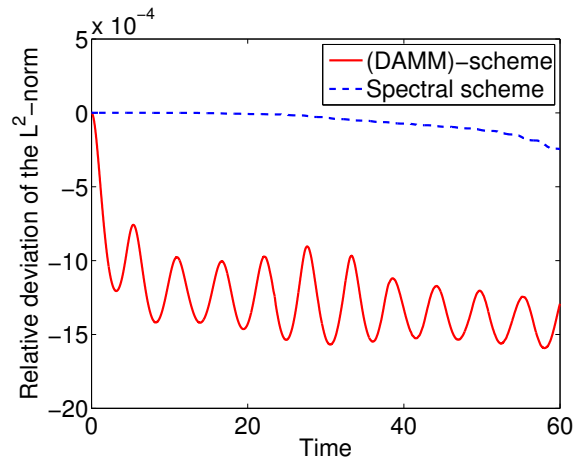
(c) Mass.



(d) Entropy.



(e) Momentum.



(f)  $L^2$ -norm.

FIGURE 3.14 – (Strong Landau damping for  $\epsilon = 1$ ) Energy of the electric field versus time (a) and deviation over time for both (DAMM) and spectral schemes of several quantities (b), (c), (d), (e) and (f). Mesh size :  $N_x = N_y = 256$ .  $T = 60$ ,  $\Delta t = 0.01$ ,  $\epsilon = 1$ , and  $\sigma = (\Delta x/L_x)^2$ .

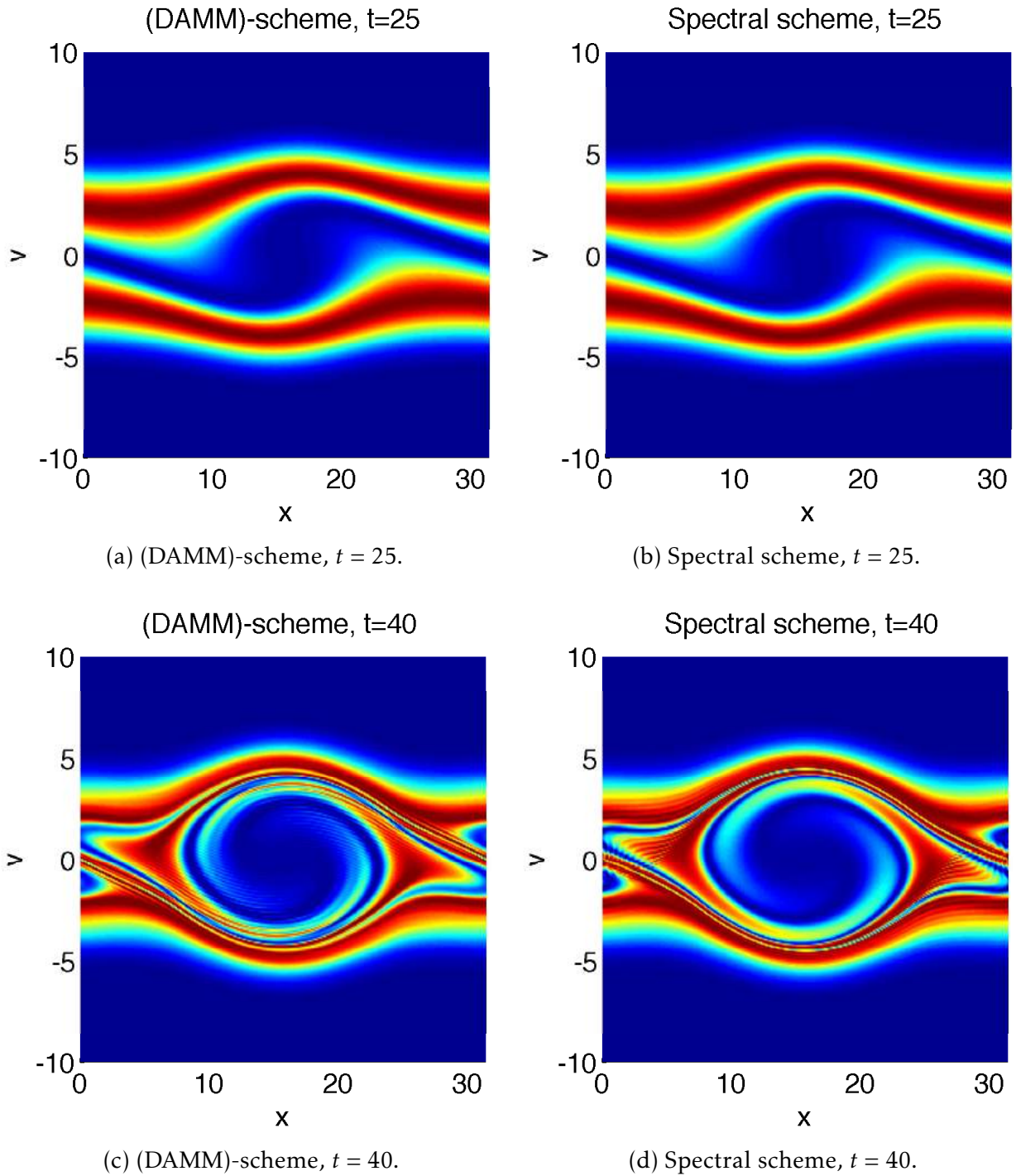
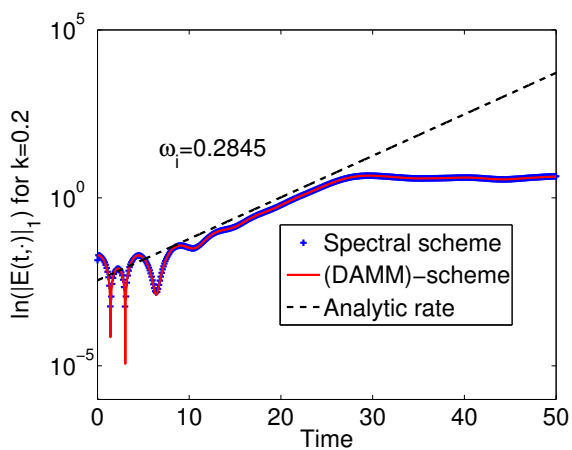
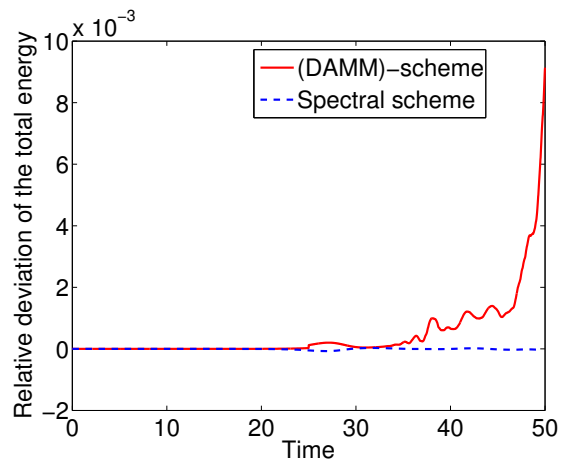


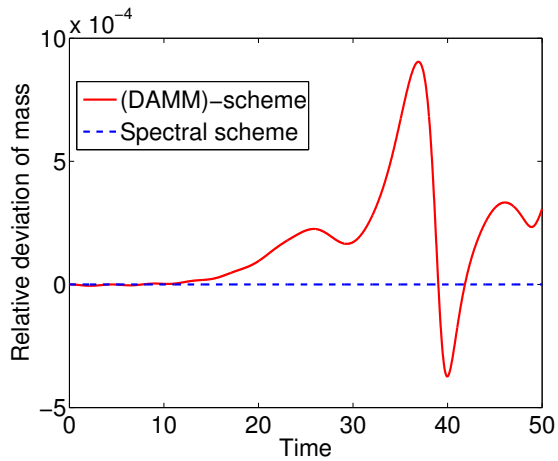
FIGURE 3.15 – (Two-stream instability for  $\epsilon = 1$  and  $f_{in}^1$ ) Distribution function  $f^\epsilon(t, x, v)$  at different times with  $k = 0.2$  and  $\gamma = 0.001$  for the (DAMM)-scheme ((a) and (c)) and the spectral scheme ((b) and (d)).  $T = 50$ ,  $N_x = 256$ ,  $N_y = 256$ ,  $\Delta t = 0.1$ ,  $\epsilon = 1$  and  $\sigma = (\Delta x/L_x)^2$ .



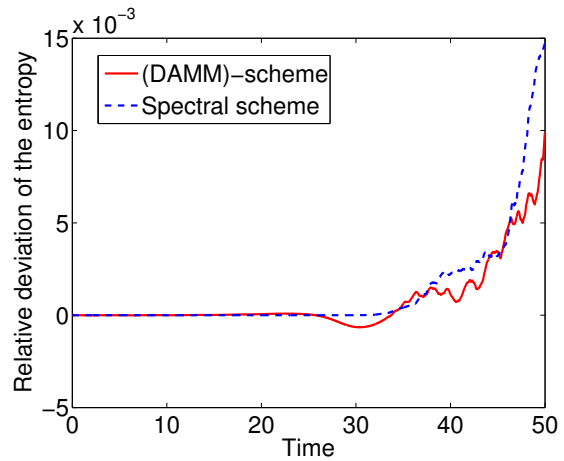
(a) Electric Field (in log value).



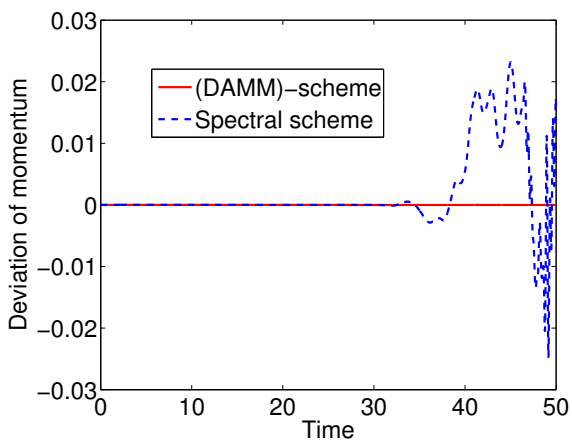
(b) Total energy.



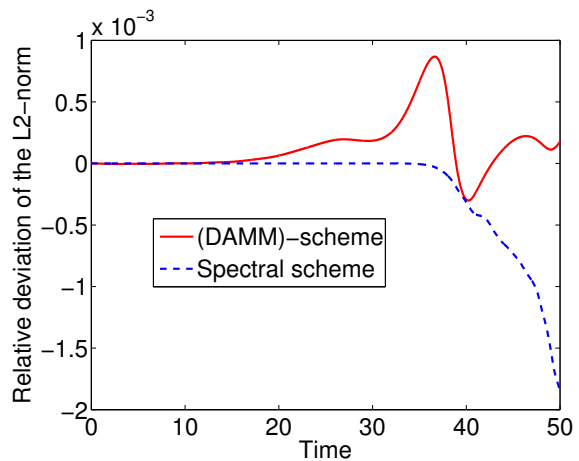
(c) Mass.



(d) Entropy.



(e) Momentum.



(f)  $L^2$ -norm.

FIGURE 3.16 – (Two-stream instability for  $\epsilon = 1$  and  $f_{in}^1$ ) Electric field versus time (a) and deviation over time for both (DAMM) and spectral schemes of several quantities (b), (c), (d), (e) and (f). Mesh size :  $N_x = N_y = 256$ .  $T = 50$ ,  $\Delta t = 0.01$ ,  $\epsilon = 1$ , and  $\sigma = (\Delta x/L_x)^2$ .

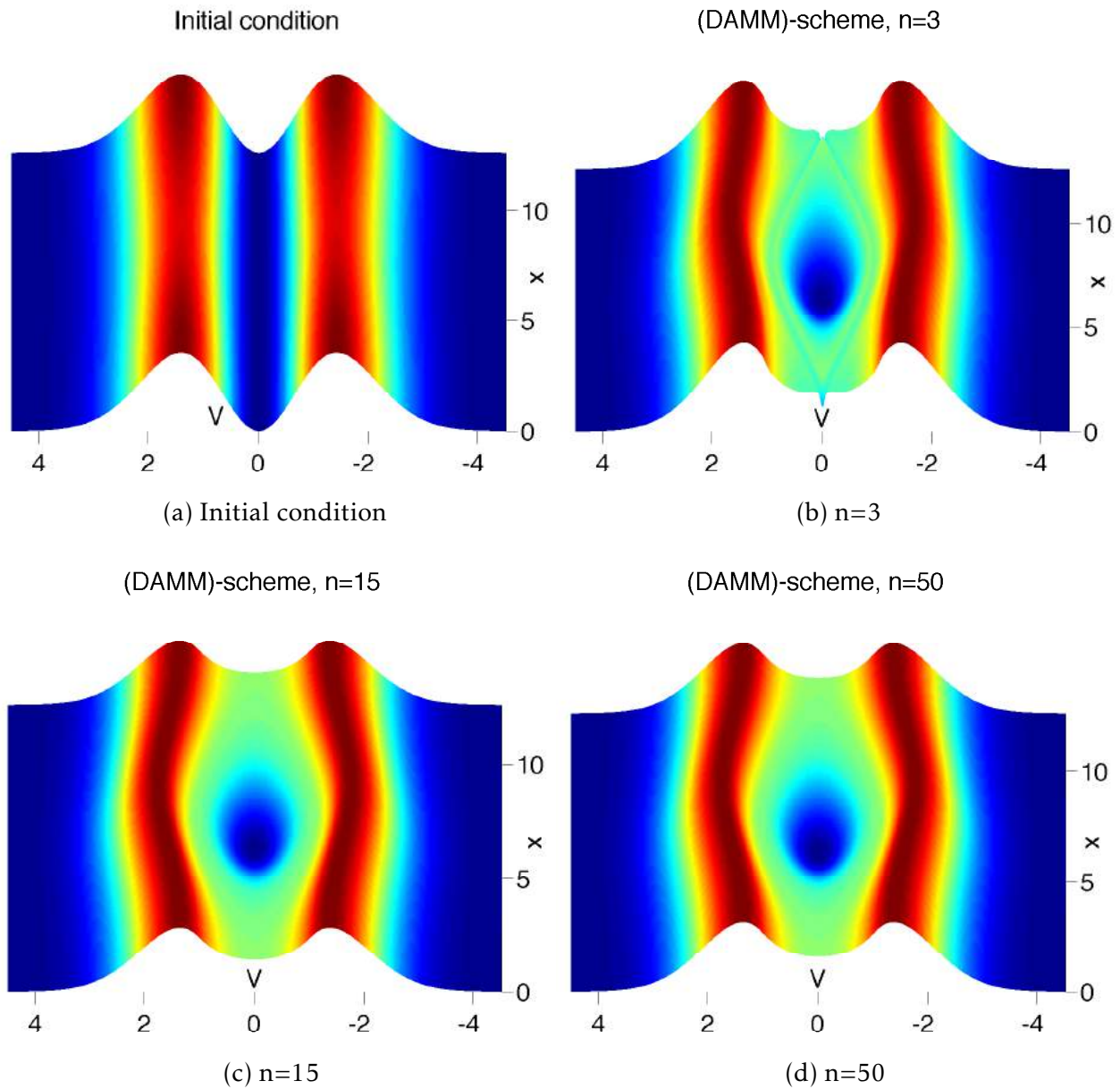


FIGURE 3.17 – (Two-stream instability for  $\epsilon = 0$  and  $f_{in}^2$ ) Distribution function  $f^0(t, x, v)$  at different times for the two stream instability with  $k = 0.5$  and  $\gamma = 0.05$  via the (DAMM)-scheme. Parameters were  $N_x = N_y = 256$ ,  $\Delta t = 0.1$ , and  $\sigma = (\Delta x/L_x)^2$ .

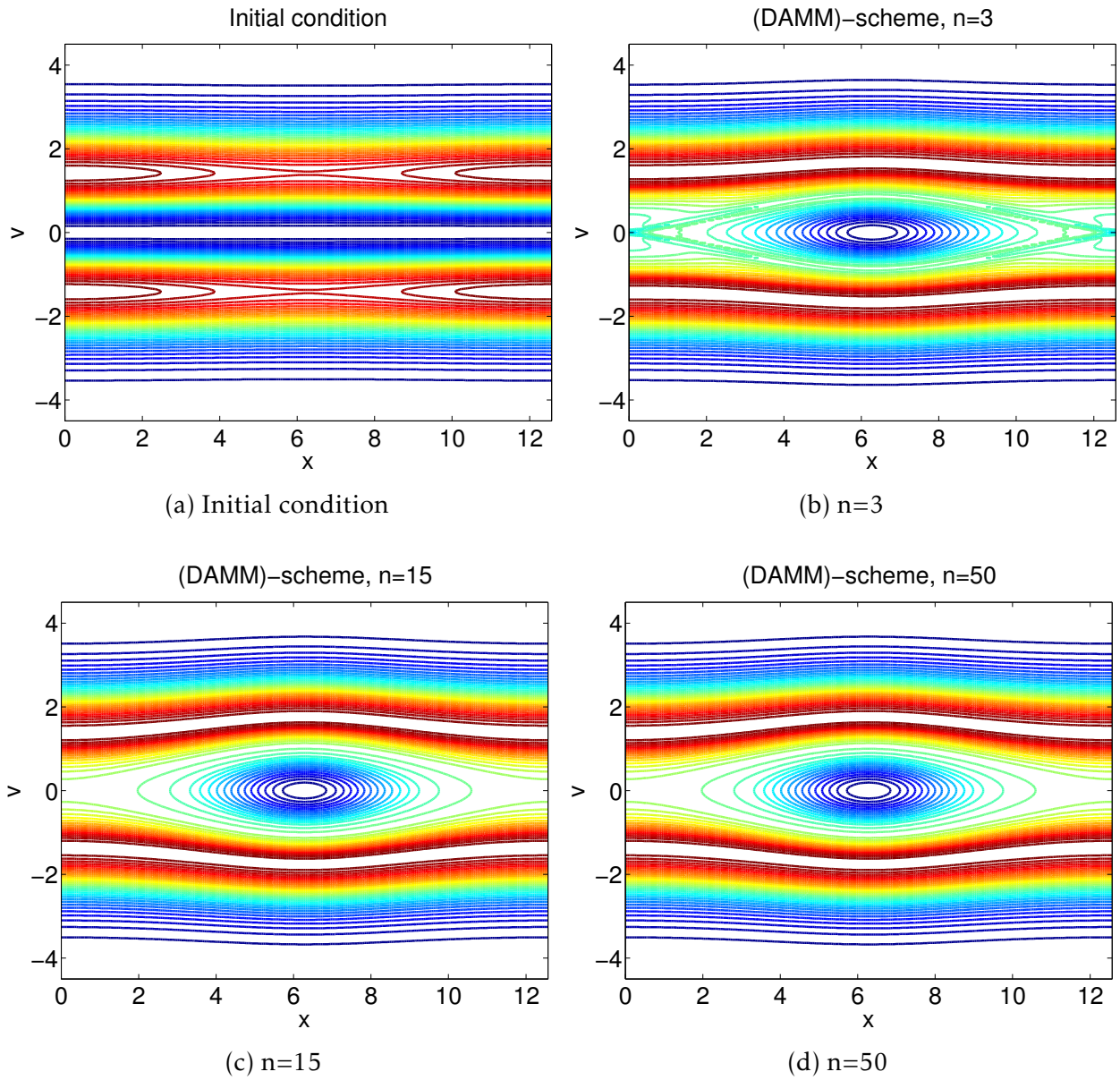


FIGURE 3.18 – (Two-stream instability for  $\epsilon = 0$  and  $f_{in}^2$ ) Contour plots of the distribution function  $f^0(t, x, v)$  at different times for the two stream instability with  $k = 0.5$  and  $\gamma = 0.05$  via the (DAMM)-scheme. Parameters were  $N_x = N_y = 256$ ,  $\Delta t = 0.1$ , and  $\sigma = (\Delta x/L_x)^2$ .



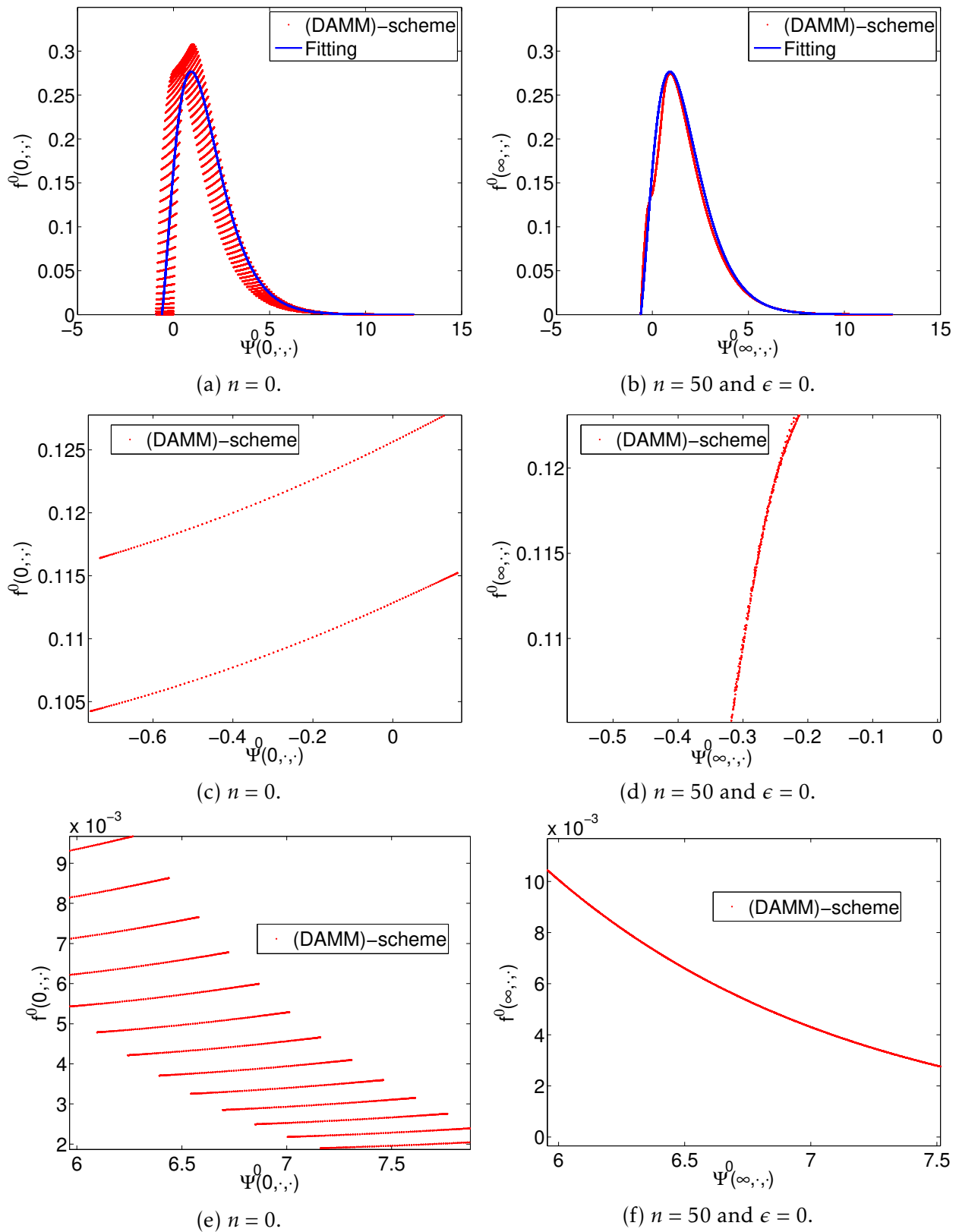


FIGURE 3.19 – (Two-stream instability for  $\epsilon = 0$  and  $f_{in}^2$ ) Plot of the distribution function at times  $t = 0$  (panels (a), (c), (e)) and  $t = 50\Delta t$  (panels (b), (d), (f)), with  $\Delta t = 0.01$  as a function of  $\Psi^0$ . Mesh size :  $N_x = N_y = 256$ . Stabilization parameter :  $\sigma = (\Delta x/L_x)^2$ .



---

---

## CHAPITRE 4

---

# NUMERICAL AND ANALYTICAL STUDY OF A FORCED VORTICITY EQUATION AND ITS ASYMPTOTIC-PRESERVING RESOLUTION

Le contenu de ce chapitre correspond à un article écrit avec C. Negulescu et M. Ottaviani :

*Numerical and analytical study of a forced vorticity equation and its asymptotic-preserving resolution.*

### **Abstract.**

Dans ce chapitre, un schéma préservant l'asymptotique est développé afin de résoudre un système de Vorticité-Poisson avec terme de forçage, qui apparaît souvent afin de décrire des écoulements turbulents. Deux cas sont étudiés. Tout d'abord, nous étudions numériquement un cas non-forcé à travers l'écoulement de Taylor-Green, afin principalement de valider notre procédure numérique. Puis, l'écoulement de Kolmogorov avec un terme de forçage sinusoïdal est considéré, cette fois numériquement et analytiquement. Un résultat analytique à propos de l'instabilité linéaire de l'écoulement est donné. Les propriétés préservant l'asymptotique de notre schéma permettent d'étudier l'état final de l'écoulement (après l'apparition de l'instabilité).

## 4.1 Motivation, background and objectives

The aim of the present paper is to investigate the following nonlinear singularly perturbed problem coming from incompressible fluid mechanics :

$$(\mathcal{V})^\epsilon \begin{cases} \partial_t \omega^\epsilon + \frac{1}{\epsilon} \mathbf{u}^\epsilon \cdot \nabla \omega^\epsilon = (\Delta \omega^\epsilon - \Delta \omega_{eq}), & (t, \mathbf{x}) \in \mathbb{R}^+ \times \Omega, \\ -\Delta \Psi^\epsilon = \omega^\epsilon, & \mathbf{u}^\epsilon = {}^\perp \nabla \Psi^\epsilon, \quad (t, \mathbf{x}) \in \mathbb{R}^+ \times \Omega, \end{cases} \quad (4.1)$$

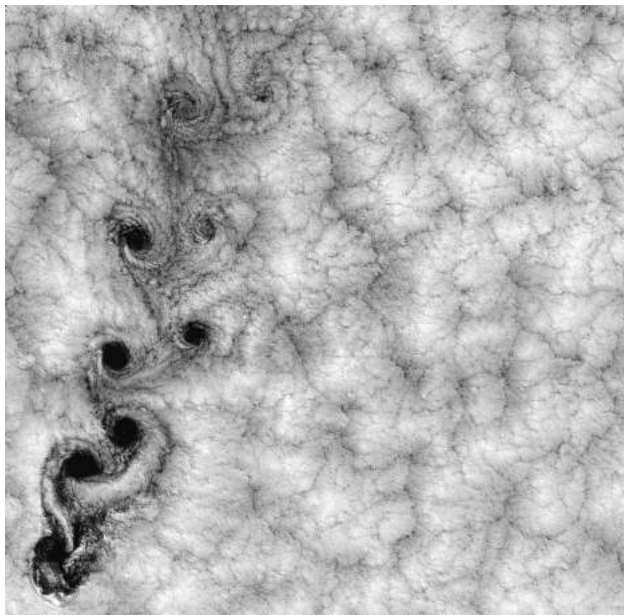
where  $\mathbf{x} := (x, y) \in \Omega$  with  $\Omega$  an open, bounded domain of  $\mathbb{R}^2$ , in our case a rectangle. The problem is associated with periodic boundary conditions and consists of an evolution equation for the vorticity  $\omega^\epsilon$ , coupled to the Poisson equation for the computation of the stream-function  $\Psi^\epsilon$ . This latter function is fixed by imposing the constraint  $\langle \Psi^\epsilon \rangle = 0$  where  $\langle \cdot \rangle$  denotes the average over the domain  $\Omega$ . The perpendicular gradient operator is defined as  ${}^\perp \nabla := (\partial_y, -\partial_x)$ . The parameter  $\epsilon \in (0, 1)$  models the stiffness of the problem, and its physical significance will be explained in Section 4.1.1. The term  $\omega_{eq}$  appearing on the right hand side of system (4.1) denotes a forcing term. In this paper, we examine two well-known vortex-problems : the Taylor-Green flow (unforced case  $\omega_{eq}(\mathbf{x}) = 0$ ) and the Kolmogorov flow (forced case  $\omega_{eq}(\mathbf{x}) = \cos(x)$ ), with special emphasis on the stability properties of these flows. The study of such simple flows is precious, as they help to acquire intuition and insight by displaying the particular behavior of the flow under the different physical mechanics and also by providing ideal test cases for checking the numerical methods. Our main focus will be on the numerical investigation of the instability of equilibria-flows described by (4.1), and this via an *Asymptotic-Preserving-scheme* [80, 112]. In particular, we are interested in the validation of the proposed AP-scheme, which was especially designed to capture efficiently and accurately all  $\epsilon$ -regimes, notably the  $\epsilon \rightarrow 0$  limit regime.

The vorticity is an essential quantity, related to the angular momentum of a fluid and permitting to give useful information about the transition to turbulence. One can think that the restriction to 2D flows is somehow artificial, and underline moreover that the 2D vorticity (and thus the 2D turbulence) behaves completely differently as compared to 3D vorticity. However a lot of flows in nature exhibit 2D dynamics, as for example flows in geophysical, astrophysical and oceanographic applications which are submitted to a strong geometrical confinement, such that the vertical motion can be neglected (hurricanes, the Great Red Spot of Jupiter, the vortex Gulf Stream, etc.). We refer the interested reader to [101] for a review of vortex dynamics in nature. In the context of plasma physics, such quasi-2D flows can occur as a consequence of the strong magnetic field which induces a strongly anisotropic dynamics. The earliest model equation was derived by Hasegawa and Mima [17]. It describes the evolution of the electric potential in a plane perpendicular to the magnetic field and it is only slightly more general than the model studied in the present work.

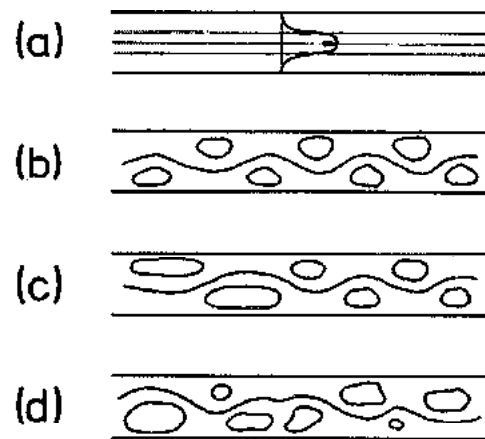
The vorticity equation (4.1) permits to treat several flow instabilities, depending on the choice of the initial condition and the boundary conditions. In the simplest geometry and without boundaries (periodic boundary conditions), it permits to study the intrinsic properties of turbulence and of the instability at its origin, meaning the turbulence not generated by some solid objet (as a wall for ex.), but by inherent properties of the flow. The two cases we shall present here, the Taylor-Green instability and the Kolmogorov instability, are very simple test cases, the relevance of whom lie especially in the theoretical analysis and in the validation of a numerical scheme, which is our main aim here.

The Taylor-Green flow is a particular 2D vortex flow, which evolves from a single-Fourier mode initial condition (for ex.  $\omega_{in} = 2 \sin x \sin y$ ), without forcing ( $\omega_{eq}(\mathbf{x}) = 0$ ). The analytic solution is known in this case and will be considered as the time-dependent equilibrium-solution to be perturbed. Without perturbation the initial vortices are simply damped in time by dissipation, without changing or mixing or evolving somehow. It is however well-known, that this basis-solution is unstable, giving rise, when perturbed, to mixing processes or filamentation, which end up with the appearance of a final state, which is completely different from the initial one. The Taylor-Green flow is closely related to the Von Karman flow (see Figure 4.1, panel (a)).

The Kolmogorov flow corresponds to a 2D, one-directional shear flow, with sinusoidal velocity profile, and is maintained in the viscous version by a forcing term ( $\omega_{eq}(\mathbf{x}) = \cos(x)$ ). It admits a stationary solution, the laminar Kolmogorov flow, which becomes unstable through perturbation. One observes then the emergence of complex time-dependent structures, leading first to a characteristic island pattern (also sometimes called cat's eye), then, as the systems evolves, to a complete rearrangement of the flow which leads to a stationary end state. The Kolmogorov flow is closely related to a channel flow without boundaries (see Figure 4.1, panel (b)).



(a) Von Karman flow.



(b) Channel flow.

FIGURE 4.1

Studying the instability of shear flows is a numerically difficult problem. Indeed, the occurrence of several time and space scales (eddies with a wide range of sizes), as well as the sensibility of the solution (unstable equilibria) to numerical errors, induces several numerical problems. Furthermore, if one is interested to capture long-time asymptotics, efficient numerical schemes are needed, permitting to control the accumulation of the errors over the time. Using traditional simulation strategies, based only on brute force are likely to fail. Numerical strategies must take full advantage of the asymptotic properties of the underlying problem in order to be efficient. In view of all these difficulties, our present aim is to introduce and validate an *Asymptotic-Preserving method*, which allows to overcome the above mentioned complications, in so far that this scheme shall permit to reach directly an equilibrium state (when it exists), jumping over the intermediate, mixing states, and as

a consequence, avoiding the error accumulation. In particular, being based on asymptotic arguments, the scheme permits to largely reduce the computation time and thus to study numerically the long-time scale dynamics of our problems. The aim of this work is hence to show that an AP-method can bring important gains w.r.t standard methods in terms of numerical efficiency.

For numerical reasons, let us reformulate the system  $(\mathcal{V})^\epsilon$  as :

$$(\mathcal{V})^\epsilon \begin{cases} \partial_t \omega^\epsilon + \frac{1}{\epsilon} \{\omega^\epsilon, \Psi^\epsilon\} = (\Delta \omega^\epsilon - \Delta \omega_{eq}), & (t, \mathbf{x}) \in \mathbb{R}^+ \times \Omega, \\ -\Delta \Psi^\epsilon = \omega^\epsilon, & (t, \mathbf{x}) \in \mathbb{R}^+ \times \Omega, \end{cases} \quad (4.2)$$

where  $\{\omega^\epsilon, \Psi^\epsilon\} = \partial_x \omega^\epsilon \partial_y \Psi^\epsilon - \partial_y \omega^\epsilon \partial_x \Psi^\epsilon$  denotes the Poisson bracket. This rewriting is more adequate for simulations, as we will discretize the Poisson bracket via the *Arakawa scheme*, designed for such an advection term.

The outline of this paper is the following. In the remaining of Section 4.1, the physical origin of system (4.1) is explained as well as its multiscale behavior and useful quantities are introduced. Section 4.2 deals with the Asymptotic-Preserving reformulation of our original problem (4.1) in order to obtain a more regular problem. That reformulation is based on a Micro-Macro decomposition of the unknown and is completed by a regularization procedure. The second part of Section 4.2 presents the numerical scheme that we develop in this paper. Section 4.3 focuses on the numerical study of the Taylor-Green test case. The main purpose of this Section is to validate our numerical procedure. Section 4.4 is dedicated to the study of the Kolmogorov flow. A result based on the relation between the growth rate of the linear instability of this flow and the aspect ratio of the domain is given. Furthermore, the nonlinear stage as well as the final state of the Kolmogorov flow are investigated. In particular, a fitting solution as well as the  $\Psi - \omega$  relation of the final state are proposed and compared to the numerical solution obtained by our numerical procedure. This article ends with Section 4.5 where some conclusions and perspectives are given.

### 4.1.1 Non-dimensional Vorticity-Poisson system

At the basis of our analysis are the incompressible Navier-Stokes equations, describing a bi-dimensional flow with velocity  $\mathbf{u} := (u_x, u_y, 0)^T$ , pressure  $p$  and forcing term  $\mathbf{u}_{eq}$  :

$$(\mathcal{NS}) \begin{cases} \partial_t \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\nabla p + \nu (\Delta \mathbf{u} - \Delta \mathbf{u}_{eq}), & (t, \mathbf{x}) \in \mathbb{R}^+ \times \Omega, \\ \nabla \cdot \mathbf{u} = 0, & (t, \mathbf{x}) \in \mathbb{R}^+ \times \Omega, \end{cases} \quad (4.3)$$

where  $\nu$  denotes the kinematic viscosity of the fluid. The divergence-free constraint of  $\mathbf{u}$  leads to the existence of a scalar stream-function  $\Psi$  such that  $\mathbf{u} = {}^\perp \nabla \Psi = (\partial_y \Psi, -\partial_x \Psi, 0)^T$ . Recalling that the vorticity is defined as  $\omega = (\nabla \times \mathbf{u}) \cdot \mathbf{e}_z$ , let us take the rotational of (4.3) and then the scalar product with  $\mathbf{e}_z := (0, 0, 1)^T$ , to rewrite this system in the equivalent vorticity/stream-function form

$$(\mathcal{V}) \begin{cases} \partial_t \omega + \{\omega, \Psi\} = \nu (\Delta \omega - \Delta \omega_{eq}), & (t, \mathbf{x}) \in \mathbb{R}^+ \times \Omega, \\ -\Delta \Psi = \omega, & (t, \mathbf{x}) \in \mathbb{R}^+ \times \Omega, \end{cases} \quad (4.4)$$

which is the "physical" version of the Vorticity-Poisson equation (4.2). The vorticity evolution problem contains some important physics which play an essential role in the understanding and visualization of turbulence, such that it acquired a big attractivity [94].

In this work, we will consider different time scales. In order to identify these latter, we proceed now to an adimensionalization to highlight the regimes of interest. For each function generically denoted by  $n(\cdot)$ , we set  $n(\cdot) := \bar{n} n'(\cdot)$ , where  $\bar{n}$  refers to the characteristic scale of  $n(\cdot)$  and  $n'$  the non-dimensional associated function. With these notation, let us fix now the spatial and time characteristic lengths  $\bar{x}$  and  $\bar{t}$  of the flow. Starting from this couple  $(\bar{x}, \bar{t})$ , we deduce the characteristic scales of the other quantities :

$$\bar{u} = \frac{\bar{x}}{\bar{t}}, \quad \bar{\omega} = \frac{\bar{u}}{\bar{x}}, \quad \bar{\Psi} = \bar{x} \bar{u}.$$

Replacing in the vorticity problem (4.4) each quantity  $n(\cdot)$  by  $\bar{n} n'$ , we obtain

$$(\mathcal{V}) \begin{cases} \partial_{t'} \omega' + \{\omega', \Psi'\} = \frac{\nu \bar{t}}{\bar{x}^2} (\Delta \omega' - \Delta \omega'_{eq}), & (t', \mathbf{x}) \in \mathbb{R}^+ \times \Omega, \\ -\Delta \Psi' = \omega', & (t', \mathbf{x}) \in \mathbb{R}^+ \times \Omega. \end{cases} \quad (4.5)$$

Introducing now a viscous time scale  $\tau_\nu := \frac{\bar{x}^2}{\nu}$ , the previous system writes

$$(\mathcal{V}) \begin{cases} \partial_{t'} \omega' + \{\omega', \Psi'\} = \frac{\bar{t}}{\tau_\nu} (\Delta \omega' - \Delta \omega'_{eq}), & (t', \mathbf{x}) \in \mathbb{R}^+ \times \Omega, \\ -\Delta \Psi' = \omega', & (t', \mathbf{x}) \in \mathbb{R}^+ \times \Omega. \end{cases} \quad (4.6)$$

Supposing that the viscous time scale  $\tau_\nu$  is very long compared to the observation time scale  $\bar{t}$ , we introduce the stiffness parameter  $\epsilon$  standing for this ratio :

$$\epsilon := \frac{\bar{t}}{\tau_\nu} \in (0, 1).$$

If we are now interested in describing phenomena arising on non-viscous time scales, the vorticity model simply reads :

$$(\mathcal{V}) \begin{cases} \partial_{t'} \omega + \{\omega, \Psi\} = \epsilon (\Delta \omega - \Delta \omega_{eq}), & t' \in (0, T), \\ -\Delta \Psi = \omega, \end{cases} \quad (4.7)$$

with  $T \in \mathbb{R}^+$  a fixed final time. Yet, if we are interested in characterizing phenomena arising on the viscous time scale, we set  $t = \epsilon t'$ , and the Vorticity-Poisson system writes

$$(\mathcal{V})^\epsilon \begin{cases} \partial_t \omega^\epsilon + \frac{1}{\epsilon} \{\omega^\epsilon, \Psi^\epsilon\} = (\Delta \omega^\epsilon - \Delta \omega_{eq}), & t \in (0, T), \\ -\Delta \Psi^\epsilon = \omega^\epsilon. \end{cases} \quad (4.8)$$

with  $T$  linked to  $\mathcal{T}$  by the relation  $T = \epsilon \mathcal{T}$ . It is this last part which is the starting point of our investigations.

### 4.1.2 Quantities of interest

Important aspects of 2D turbulence can be recovered by studying some important integral (average) quantities, permitting to evaluate somehow the vorticity of the flow and the energy/enstrophy cascades. Let us introduce here these physical quantities (see [34]). In the following,  $|\Omega|$  denotes the measure of the domain  $\Omega$ .

## The kinetic energy

The kinetic energy can be linked to the vorticity field, and is given mathematically by :

$$\mathcal{K}^\epsilon(t, \omega^\epsilon) = \frac{1}{2|\Omega|} \int_{\Omega} \Psi^\epsilon(t, x, y) \omega^\epsilon(t, x, y) dx dy = \frac{1}{2|\Omega|} \int_{\Omega} |\mathbf{u}^\epsilon(t, x, y)|^2 dx dy.$$

In the case where the forcing term vanishes, we expect a monotonic decreasing in time of this quantity, due to the viscosity which dissipates the energy of the system.

## Enstrophy and Palinstrophy

We consider now two essential quantities concerning directly the dynamics of the vorticity field and which are particularly useful for the study of turbulence. Mathematically, the enstrophy  $\mathcal{E}^\epsilon(t, \omega^\epsilon)$  and the palinstrophy  $\mathcal{P}^\epsilon(t, \omega^\epsilon)$  are defined respectively by

$$\begin{aligned} \mathcal{E}^\epsilon(t, \omega^\epsilon) &= \frac{1}{2|\Omega|} \int_{\Omega} [\omega^\epsilon(t, x, y)]^2 dx dy, \\ \mathcal{P}^\epsilon(t, \omega^\epsilon) &= \frac{1}{2|\Omega|} \int_{\Omega} [\nabla \omega^\epsilon(t, x, y)]^2 dx dy. \end{aligned}$$

The enstrophy measures the intensity of the vorticity field or in other words the strength of the fluid rotation and represents some form of energy. The palinstrophy quantifies how the vorticity is transferred within the domain and permits to recognize the merging of vortices. Some relations connect the kinetic energy, the enstrophy and the palinstrophy, summarized in the following Proposition.

**Proposition 1** 1. *The time evolution of the kinetic energy is given by the relation*

$$\frac{d}{dt} \mathcal{K}^\epsilon(t, \omega) = -2(\mathcal{E}^\epsilon(t, \omega) + I^\epsilon(t, \mathbf{u})),$$

where  $I^\epsilon(t, \mathbf{u}^\epsilon) = \frac{1}{2|\Omega|} \int_{\Omega} \Delta \mathbf{u}_{eq} \cdot \mathbf{u}^\epsilon dx dy$  denotes the input energy through the external forcing velocity  $\mathbf{u}_{eq}$ .

2. *The time evolution of the enstrophy satisfies*

$$\frac{d}{dt} \mathcal{E}^\epsilon(t, \omega^\epsilon) = -2(\mathcal{P}^\epsilon(t, \omega^\epsilon) + J^\epsilon(t, \omega^\epsilon)),$$

where  $J^\epsilon(t, \omega^\epsilon) = \frac{1}{2|\Omega|} \int_{\Omega} \Delta \omega_{eq} \omega^\epsilon dx dy$  describes the enstrophy production through the external forcing  $\omega_{eq}$ .

In particular, we expect that the kinetic energy and the enstrophy decrease in time when the forcing term is zero. In order to end this part, we introduce two quantities, the Rayleigh quotient  $\mathcal{R}^\epsilon(t, \omega^\epsilon)$ , and the dissipation quotient  $\Lambda^\epsilon(t, \omega^\epsilon)$  :

$$\mathcal{R}^\epsilon(t, \omega^\epsilon) = \frac{\mathcal{E}^\epsilon(t, \omega^\epsilon)}{\mathcal{K}^\epsilon(t, \omega^\epsilon)}, \quad \Lambda^\epsilon(t, \omega^\epsilon) = \frac{\mathcal{P}^\epsilon(t, \omega^\epsilon)}{\mathcal{E}^\epsilon(t, \omega^\epsilon)}.$$

The time derivative of the Rayleigh quotient verifies the following relation, obtained after simple computations

$$\frac{d}{dt} \mathcal{R}^\epsilon(t, \omega^\epsilon) = -2\mathcal{R}^\epsilon(t, \omega^\epsilon) [\Lambda^\epsilon(t, \omega^\epsilon) - \mathcal{R}^\epsilon(t, \omega^\epsilon) + \mathcal{K}^\epsilon(t, \omega^\epsilon) (J^\epsilon(t, \omega^\epsilon) / \mathcal{R}^\epsilon(t, \omega^\epsilon) - I^\epsilon(t, \mathbf{u}^\epsilon))].$$



## Vorticity thickness

The last quantity of interest is the vorticity thickness, which characterizes the length's evolution of the vortices. Its mathematical expression is given by :

$$\delta^\epsilon(t, \omega^\epsilon) = \frac{2}{\sup_{y \in \Omega_y} |\langle \omega^\epsilon \rangle(t, y)|}, \quad \text{with } \langle \omega^\epsilon \rangle(t, y) = \frac{1}{L_x} \int_0^{L_x} \omega^\epsilon(t, x, y) dx,$$

and  $\Omega_y = [0, L_y]$ .

Before going on, let us say here some words about the characteristic features of 2D turbulence. Turbulent flows are far from being in equilibrium and seem to be very chaotic. One of the characteristic features of turbulence is the occurrence of eddies which span over a wide range of space scales. Energy is injected in the system at one scale, drives the turbulence -in particular the production of vortices- and is finally removed from the system at another scale. In 2D turbulence, the direction of this energy cascade is inverse to the one of 3D turbulence, namely energy moves in 2D from smaller to larger scales. The consequence of this is that, contrary to 3D turbulence which is largely incoherent, 2D turbulent flows are rather self-organized and form large-scale, long-lived coherent dynamical structures. It is exactly what we shall observe our future simulations (see Figure 4.10).

## 4.2 Numerical scheme

Let us present in this section the numerical scheme we used for an efficient resolution of (4.1). It will be firstly based on an Asymptotic-Preserving reformulation and secondly on the use of the Arakawa scheme for the advection operators as well as the DIRK (Diagonal Implicit Runge-Kutta) approach for the time derivative.

### 4.2.1 Asymptotic-Preserving reformulation

An Asymptotic-Preserving scheme is a numerical approach designed to solve efficiently singularly-perturbed problems, denoted generically  $P^\epsilon$ , which contain some small parameter  $\epsilon \in [0, 1]$ . A precise definition of an AP-scheme is given in the next definition (see also commutative diagram in Figure 4.2).

**Definition 2** Consider a singularly-perturbed problem  $P^\epsilon$ , whose solution is assumed to converge (in a certain sense) toward the solution of a limit problem  $P^0$ . An AP-scheme for  $P^\epsilon$ , denoted  $P^{\epsilon, \mathbf{h}}$ , is a numerical scheme which enjoys the following properties :

- The AP-scheme  $P^{\epsilon, \mathbf{h}}$  is stable in a suitable sense, uniformly in  $\epsilon$ .
- For fixed stiffness parameter  $\epsilon > 0$ , the AP-scheme  $P^{\epsilon, \mathbf{h}}$  provides a consistent discretization of the problem  $P^\epsilon$ , as  $\mathbf{h} \rightarrow 0$ .
- For fixed discretization parameters  $\mathbf{h} > 0$ , the AP-scheme  $P^{\epsilon, \mathbf{h}}$  provides in the limit  $\epsilon \rightarrow 0$  a consistent discretization of the limit problem  $P^0$ .

Physical problems arising in nature are complex and contain several time and space scales, which are difficult to capture in their globality with standard schemes. Classical numerical procedures such as explicit methods are submitted to a CFL-condition, which forces the user to solve also the small scales of the problem, which can be sometimes undesirable. This leads necessary to huge computational costs. Fully implicit methods can

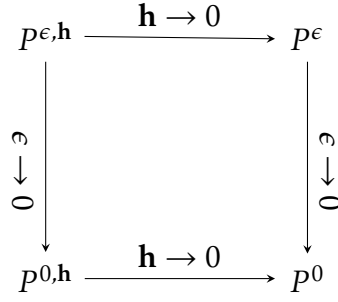


FIGURE 4.2 – Commutative properties of AP-schemes.

sometimes suffer from accuracy problems when the perturbation parameter is too small, as they do not capture the asymptotic limit. In such multi-scale frameworks, AP-schemes become interesting, as they permit to eliminate those scales which are not relevant in the considered study without loosing accuracy. This aspect will be illustrated in the last part of this work.

The AP-reformulation we shall introduce for the resolution of (4.8) is based on the following decomposition of the vorticity field  $\omega^\epsilon$  into a macroscopic and a microscopic part, namely

$$\omega^\epsilon = \chi^\epsilon + \epsilon \xi^\epsilon, \quad (4.9)$$

where  $\chi^\epsilon$  denotes the macroscopic part, which belongs to the kernel of the dominant transport operator  $\mathcal{T}_{\Psi^\epsilon} := \{\cdot, \Psi^\epsilon\}$  and  $\xi^\epsilon$  denotes the microscopic part.

Inserting now (4.9) in (4.8), permits to obtain the following system for the unknowns  $(\omega^\epsilon, \xi^\epsilon, \Psi^\epsilon)$ :

$$(\mathcal{M}\mathcal{M})^\epsilon \begin{cases} \partial_t \omega^\epsilon + \{\xi^\epsilon, \Psi^\epsilon\} = (\Delta \omega^\epsilon - \Delta \omega_{eq}), & (t, \mathbf{x}) \in (0, T) \times \Omega, \\ \{\omega^\epsilon, \Psi^\epsilon\} - \epsilon \{\xi^\epsilon, \Psi^\epsilon\} = 0, & (t, \mathbf{x}) \in (0, T) \times \Omega, \\ -\Delta \Psi^\epsilon = \omega^\epsilon, & (t, \mathbf{x}) \in (0, T) \times \Omega. \end{cases} \quad (4.10)$$

This reformulation of (4.8) is now a regular problem in  $\epsilon$ , permitting to capture for  $\epsilon \rightarrow 0$  the limit problem  $(\mathcal{M}\mathcal{M})^0$ . Nevertheless, it is ill-posed since  $\xi^\epsilon$  is not unique. To overcome this new problem, we shall employ a regularization technique already proposed in [53], by inserting the term  $\sigma \xi^\epsilon$  in the second equation of (4.10). We obtain then the final reformulated system :

$$(\mathcal{M}\mathcal{M})_\sigma^\epsilon \begin{cases} \partial_t \omega^{\epsilon, \sigma} + \{\xi^{\epsilon, \sigma}, \Psi^{\epsilon, \sigma}\} = (\Delta \omega^{\epsilon, \sigma} - \Delta \omega_{eq}), & (t, \mathbf{x}) \in (0, T) \times \Omega, \\ \{\omega^{\epsilon, \sigma}, \Psi^{\epsilon, \sigma}\} - \epsilon \{\xi^{\epsilon, \sigma}, \Psi^{\epsilon, \sigma}\} + \sigma \xi^{\epsilon, \sigma} = 0, & (t, \mathbf{x}) \in (0, T) \times \Omega, \\ -\Delta \Psi^{\epsilon, \sigma} = \omega^{\epsilon, \sigma}, & (t, \mathbf{x}) \in (0, T) \times \Omega. \end{cases} \quad (4.11)$$

Remark that due to the regularization,  $(\mathcal{M}\mathcal{M})_\sigma^\epsilon$  is no more equivalent to  $(\mathcal{M}\mathcal{M})^\epsilon$ , however  $\sigma$  shall be chosen small enough, of the order of the truncation error, in order not to modify too much the initial problem.

## 4.2.2 Numerical discretization

Let us detail now the numerical discretization of the reformulated system (4.11).

### Discretization parameters

In what follows, we consider a bounded simulation domain  $\Omega_S := (0, L_x) \times (0, L_y)$  and all considered functions are supposed to be doubly periodic in  $x$  and  $y$ . The time interval  $[0, T]$ ,  $T > 0$ , is discretized as follows :

$$t^n := n\Delta t, \quad \Delta t := T/N_t, \quad n \in [[0, N_t]], \quad N_t \in \mathbb{N}.$$

In order to provide a uniform mesh on the domain  $\Omega_S$ , we define the grid spacings as follow :

$$x_i := (i-1)\Delta x, \quad y_j := (j-1)\Delta y, \quad \Delta x := L_x/N_x, \quad \Delta y := L_y/N_y,$$

where  $i \in [[1, N_x + 1]]$ , and  $j \in [[1, N_y + 1]]$ . For any function  $f : [0, T] \times \Omega_S \rightarrow \mathbb{R}$ ,  $f_{i,j}^n$  refers to the numerical approximation of  $f(t^n, x_i, y_j)$ , and  $f_h^n$  shall denote the discrete grid-function  $(f_{i,j}^n)_{i,j}$ . Therefore, thanks to the periodicity boundary conditions, we put :

$$f_{N_x+1,j}^n = f_{1,j}^n, \quad f_{i,N_y+1}^n = f_{i,1}^n, \quad \forall (n, i, j) \in [[0, N_t]] \times [[1, N_x + 1]] \times [[1, N_y + 1]].$$

### Space semi-discretization

Let us proceed with the discretization of the operators appearing in (4.11). For the discrete Laplace operator evaluated at the point  $(x_i, y_j)$ , one has

$$[\Delta_h f_h]_{i,j} := \frac{1}{\Delta x^2} (f_{i+1,j} - 2f_{i,j} + f_{i-1,j}) + \frac{1}{\Delta y^2} (f_{i,j+1} - 2f_{i,j} + f_{i,j-1}).$$

For the Poisson brackets operator, let us use the second order Arakawa discretization [6]. For two functions  $f, g : \Omega_S \rightarrow \mathbb{R}$ , the discrete version of the Poisson bracket  $\{f, g\}$  evaluated at the point  $(x_i, y_j)$  is expressed by :

$$\begin{aligned} [f_h, g_h]_{i,j} := & \frac{1}{12\Delta x\Delta y} \left( f_{i+1,j} \mathcal{A}_{i,j} + f_{i-1,j} \mathcal{B}_{i,j} + f_{i,j+1} \mathcal{C}_{i,j} + f_{i,j-1} \mathcal{D}_{i,j} \right. \\ & \left. + f_{i+1,j+1} \mathcal{E}_{i,j} + f_{i-1,j-1} \mathcal{F}_{i,j} + f_{i-1,j+1} \mathcal{G}_{i,j} + f_{i+1,j-1} \mathcal{H}_{i,j} \right). \end{aligned}$$

where the coefficients write

$$\begin{aligned} \mathcal{A}_{i,j} &:= g_{i,j+1} - g_{i,j-1} + g_{i+1,j+1} - g_{i+1,j-1}, & \mathcal{E}_{i,j} &:= g_{i,j+1} - g_{i+1,j}, \\ \mathcal{B}_{i,j} &:= g_{i,j-1} - g_{i,j+1} - g_{i-1,j+1} + g_{i-1,j-1}, & \mathcal{F}_{i,j} &:= g_{i,j-1} - g_{i-1,j}, \\ \mathcal{C}_{i,j} &:= g_{i-1,j} - g_{i+1,j} - g_{i+1,j+1} + g_{i-1,j+1}, & \mathcal{G}_{i,j} &:= g_{i-1,j} - g_{i,j+1}, \\ \mathcal{D}_{i,j} &:= g_{i+1,j} - g_{i-1,j} + g_{i+1,j-1} - g_{i-1,j-1}, & \mathcal{H}_{i,j} &:= g_{i+1,j} - g_{i,j-1}. \end{aligned}$$

We finish this paragraph with the discretization of the Poisson equation  $-\Delta f = g$ . In order to limit the computational cost of the scheme, we resolve this latter by the Fourier method. For a discrete grid-function  $(f_{i,j})_{i,j}$ , we denote by  $(\widehat{f_{p,q}})_{p,q}$  its discrete Fourier transform, where  $p \in [[-N_x/2, N_x/2]]$  and  $q \in [[-N_y/2, N_y/2]]$ . Introducing the quantities

$$k_p = \frac{2\pi p}{L_x}, \quad k_q = \frac{2\pi q}{L_y},$$

Thus, the Poisson equation  $-\Delta f = g$  becomes, after the application of the discrete Fourier transform,

$$\widehat{f_{p,q}} = \frac{\widehat{g_{p,q}}}{k_p^2 + k_q^2}.$$

And denoting by  $\widetilde{f}$  the inverse discrete Fourier transform of  $f$ , we obtain

$$(f_{i,j})_{i,j} = \widetilde{(\widehat{f_{p,q}})_{p,q}}.$$

### Time discretization

In order to achieve second-order accuracy in time for the problem (4.11), we use a DIRK (Diagonally-Implicit Runge-Kutta) approach. The RK-method is recalled here for clarity reasons [2].

We consider a problem of the form  $\partial_t f = L(f) + g(t)$ , where  $L$  denotes some differential operator and  $g$  a source term. An  $r$ -stage Runge-Kutta approach is characterized by its Butcher table

$$\begin{array}{c|ccc} c_1 & a_{11} & \dots & a_{1r} \\ \vdots & \vdots & & \vdots \\ c_r & a_{r1} & \dots & a_{rr} \\ \hline & b_1 & \dots & b_r \end{array}.$$

For a given  $u^n$ , the subsequent  $u^{n+1}$  is defined by the formula

$$u^{n+1} = u^n + \Delta t \sum_{j=1}^r b_j (L(u_j) + g(t + c_j \Delta t)),$$

where each  $u_i$  is defined by

$$u_i = u^n + \Delta t \sum_{j=1}^r a_{i,j} (L(u_j) + g(t + c_j \Delta t)).$$

Note that in the case where  $b_j = a_{rj}$  for  $j = 1, \dots, r$ , then  $u^{n+1}$  is equal to the last stage of the method, namely  $u_r$ . For our problem (4.11), we consider the following 2-stage Butcher table

$$\begin{array}{c|cc} \mu & \mu & 0 \\ 1 & 1 - \mu & \mu \\ \hline & 1 - \mu & \mu \end{array}.$$

In all our simulations, we choose  $\mu := 1 - 1/\sqrt{2}$ . With this choice, the method is  $L$ -stable (see [72], Sec IV-3).

## Fixed point procedure and final scheme

Due to the nonlinearity of our problem (4.11), in particular the function  $\Psi^{\epsilon,\sigma}$  is linked to the function  $\omega^{\epsilon,\sigma}$  via the Poisson equation, we have to develop a fixed point procedure in order to resolve numerically the problem (4.11) with an implicit numerical method. If we choose  $l \in \mathbb{N}$  as an iteration index, the stream-function  $\Psi^{\epsilon,\sigma}$  can be considered as fixed during the iteration step  $l \rightarrow l+1$ , given by the previous iteration. We denote by  $\Delta t^\star$  the product  $\mu \Delta t$ .

For each time step  $n \in \mathbb{N}$ , we are looking thus for  $(\omega_h^{\epsilon,\sigma,n+1}, \xi_h^{\epsilon,\sigma,n+1})$ , by iterating in  $l \in \mathbb{N}$  as follows :  
starting from

$$\omega_{i,j}^{\epsilon,\sigma,n+1,0} := \omega_{i,j}^{\epsilon,\sigma,n}, \quad \text{we have}$$

$$\left\{ \begin{array}{l} \textbf{Poisson solver :} \\ \widehat{\Psi_{p,q}^{\epsilon,\sigma,n+1,l}} = \frac{\widehat{\omega_{p,q}^{\epsilon,\sigma,n+1,l}}}{k_p^2 + k_q^2}. \\ \textbf{Stage 1 :} \\ \omega_{1,i,j}^{\epsilon,\sigma,n+1,l+1} + \Delta t^\star [\xi_{1,h}^{\epsilon,\sigma,n+1,l+1}, \Psi_h^{\epsilon,\sigma,n+1,l}]_{i,j} = \omega_{i,j}^{\epsilon,\sigma,n} + \Delta t^\star \left( \left[ \Delta \omega_{1,h}^{\epsilon,\sigma,n+1,l+1} \right]_{i,j} - \left[ \Delta \omega_{eq,h} \right]_{i,j} \right), \\ [\omega_{1,h}^{\epsilon,\sigma,n+1,l+1}, \Psi_h^{\epsilon,\sigma,n+1,l}]_{i,j} - \epsilon [\xi_{1,h}^{\epsilon,\sigma,n+1,l+1}, \Psi_h^{\epsilon,\sigma,n+1,l}]_{i,j} + \sigma \xi_{1,i,j}^{\epsilon,\sigma,n+1,l+1} = 0. \\ \textbf{Stage 2 :} \\ \omega_{2,i,j}^{\epsilon,\sigma,n+1,l+1} + \Delta t^\star [\xi_{2,h}^{\epsilon,\sigma,n+1,l+1}, \Psi_h^{\epsilon,\sigma,n+1,l}]_{i,j} = \omega_{i,j}^{\epsilon,\sigma,n} + \frac{1-\mu}{\mu} (\omega_{1,i,j}^{\epsilon,\sigma,n+1,l+1} - \omega_{i,j}^{\epsilon,\sigma,n}) \\ + \Delta t^\star \left( \left[ \Delta \omega_{2,h}^{\epsilon,\sigma,n+1,l+1} \right]_{i,j} - \left[ \Delta \omega_{eq,h} \right]_{i,j} \right), \\ [\omega_{2,h}^{\epsilon,\sigma,n+1,l+1}, \Psi_h^{\epsilon,\sigma,n+1,l}]_{i,j} - \epsilon [\xi_{2,h}^{\epsilon,\sigma,n+1,l+1}, \Psi_h^{\epsilon,\sigma,n+1,l}]_{i,j} + \sigma \xi_{2,i,j}^{\epsilon,\sigma,n+1,l+1} = 0. \\ \textbf{Final stage :} \\ (\omega_{i,j}^{\epsilon,\sigma,n+1,l+1}, \xi_{i,j}^{\epsilon,\sigma,n+1,l+1}) = (\omega_{2,i,j}^{\epsilon,\sigma,n+1,l+1}, \xi_{2,i,j}^{\epsilon,\sigma,n+1,l+1}). \end{array} \right. \quad (4.12)$$

In all of the subsequent simulations, the stopping criterion chosen for these iterations (at  $l = l_f$ ) is

$$\frac{\|\omega_h^{\epsilon,\sigma,n+1,l+1} - \omega_h^{\epsilon,\sigma,n+1,l}\|_{L_h^1}}{\|\omega_h^{\epsilon,\sigma,n+1,l}\|_{L_h^1}} + \frac{\|\Psi_h^{\epsilon,\sigma,n+1,l+1} - \Psi_h^{\epsilon,\sigma,n+1,l}\|_{L_h^1}}{\|\Psi_h^{\epsilon,\sigma,n+1,l}\|_{L_h^1}} < 10^{-2},$$

where  $\|f_h\|_{L_h^1} := \Delta x \Delta y \sum_{i=1}^{N_x+1} \sum_{j=1}^{N_y+1} |f_{i,j}|$  denotes the discrete  $L_1$ - norm of a grid function  $f_h$ . We end the procedure by setting

$$\begin{aligned} \omega_{i,j}^{\epsilon,\sigma,n+1} &:= \omega_{i,j}^{\epsilon,\sigma,n+1,l_f+1}, \\ \xi_{i,j}^{\epsilon,\sigma,n+1} &:= \xi_{i,j}^{\epsilon,\sigma,n+1,l_f+1}. \end{aligned}$$

Finally, the new stream-function at iteration  $n + 1$  is computed then via

$$(\Psi_{i,j}^{\epsilon,\sigma,n+1})_{i,j} = (\widehat{\Psi_{p,q}^{\epsilon,\sigma,n+1}})_{p,q}, \quad \text{with} \quad \widehat{\Psi_{p,q}^{\epsilon,\sigma,n+1}} := \frac{\widehat{\omega_{p,q}^{\epsilon,\sigma,n+1}}}{k_p^2 + k_q^2}.$$

### 4.3 Numerical study of the Taylor-Green vortex case

The aim of this section is to validate the previously presented (DAMM)-scheme with a well-known benchmark, namely the Taylor-Green flow, without forcing term :  $\omega_{eq} \equiv 0$ .

The Taylor-Green flow is a well-known analytic solution of the system (4.8). This flow is characterized by a periodic array of counter-rotating vortices decaying in time. In the following, we choose the box domain  $\Omega_S = (0, L_x) \times (0, L_y)$ , with  $L_x = L_y = 2\pi$ , associated with periodic boundary conditions. At the initial time,  $t = 0$ , we consider the vorticity field  $\omega_{in}^A(x, y) = 2 \sin(x) \sin(y)$ . This leads to the following exact solution of the vorticity equation (4.8)  $\omega^{ex}(t, x, y) = 2 \exp(-2t) \sin(x) \sin(y)$ , which corresponds to a velocity field of  $\mathbf{u}^{ex}(t, x, y) = \exp(-2t) (\cos(y) \sin(x), -\cos(x) \sin(y))^T$ , and a stream-function  $\Psi_{ex}(t, x, y) = \exp(-2t) \sin(x) \sin(y)$ . Thus, over time, the solution corresponds to the initial condition dissipated due to viscosity. In particular, since the stream-function is proportional to the vorticity field in that case, the Poisson bracket  $\frac{1}{\epsilon} \{\omega^{ex}, \Psi^{ex}\}$  vanishes.

#### 4.3.1 The Taylor-Green vortex for large values of $\epsilon$ .

Before starting the simulations, we recall that we denote by  $\mathcal{T}$  the physical time and by  $T$  the simulation one. These two times are linked by the relation :  $T = \epsilon \mathcal{T}$ . Let us first validate the damping rate of the Taylor-Green vortices. The initial condition  $\omega_{in}^A$  was introduced just before. In Figure 4.3 (a), we plot a cut of the solution  $\omega^{\epsilon,\sigma}$  at  $y = \pi/4$  obtained with the (DAMM)-scheme (4.12) at three physical times :  $t = 0$ ,  $t = T/2$  and  $t = T$ , with  $\mathcal{T} = 1$  and  $\epsilon = 1$  (remark that in this case,  $t = t'$  and  $T = \mathcal{T}$ ). Moreover, we superpose the analytical solution  $\omega^{ex}$  at the same times. We observe a perfect correspondance between the two solutions, for the considered times. The dissipation of the vortices caused by the viscosity is moreover clearly visible. In Figure 4.3 (b), we plot the evolution in time of the vorticity maximum, obtained with the (DAMM)-scheme for two values of  $\epsilon$ , namely  $\epsilon = 1$  and  $\epsilon = 0.1$  and  $\mathcal{T} = 1$ . As expected, the dissipation is stronger for higher  $\epsilon$ . As before, we add in Figure 4.3 (b) the analytic solution. The correspondance between the two solutions is clearly visible, permitting to confirm that the (DAMM)-scheme mimics the expected behavior of the flow in this case.

Now, let us show that the (DAMM)-scheme verifies the relations established earlier about the kinetic energy, the enstrophy, and the palinstrophy. For that, in Figure 4.4 (a), we have plotted the evolution in time of these quantities, for  $\epsilon = 1$ . We observe the decrease towards zero of all these quantities, as expected, since the flow is just damped by viscosity. Furthermore, the link between the derivative in time of the kinetic energy and the enstrophy, as well as the link between the time derivative of the enstrophy and the palinstrophy is also well-reproduced by the (DAMM)-scheme. In Figure 4.4 (b), we plot the enstrophy as a function of the kinetic energy. There is clearly a linear relation between the two quantities, which indicates that the Rayleigh quotient is constant over time. Moreover, if we plot the evolution of  $\Lambda(\omega^{\epsilon,\sigma})$  as a function of the kinetic energy, we find the same straight line,

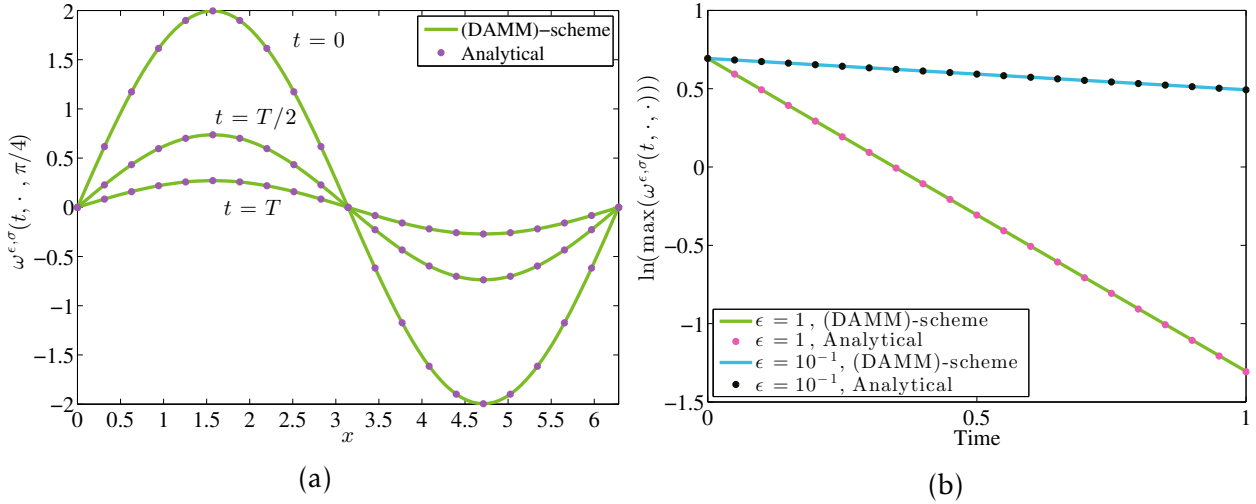


FIGURE 4.3 – (Taylor-Green vortex solution for large  $\epsilon$ -values and initial condition  $\omega_{in}^A$ ). Cut at  $y = \pi/4$  of the vorticity field versus  $x$  at three different times (a). Evolution in time of the vorticity maximum for two different values of  $\epsilon$  :  $\epsilon = 1$  and  $\epsilon = 0.1$  (b). Parameters were  $N_x = N_y = 100$ ,  $\Delta t = 0.001$ ,  $T = 1$  and  $\sigma = (\Delta x/L_x)^2$ .

thus  $\Lambda^{\epsilon, \sigma}(\omega^{\epsilon, \sigma}) = \mathcal{R}^{\epsilon, \sigma}(\omega^{\epsilon, \sigma}) = cte$ , meaning the relation connecting  $\mathcal{R}^{\epsilon, \sigma}$ ,  $\frac{d}{dt}\mathcal{R}^{\epsilon, \sigma}$  and  $\Lambda^{\epsilon, \sigma}$  is also verified by the (DAMM)-scheme.

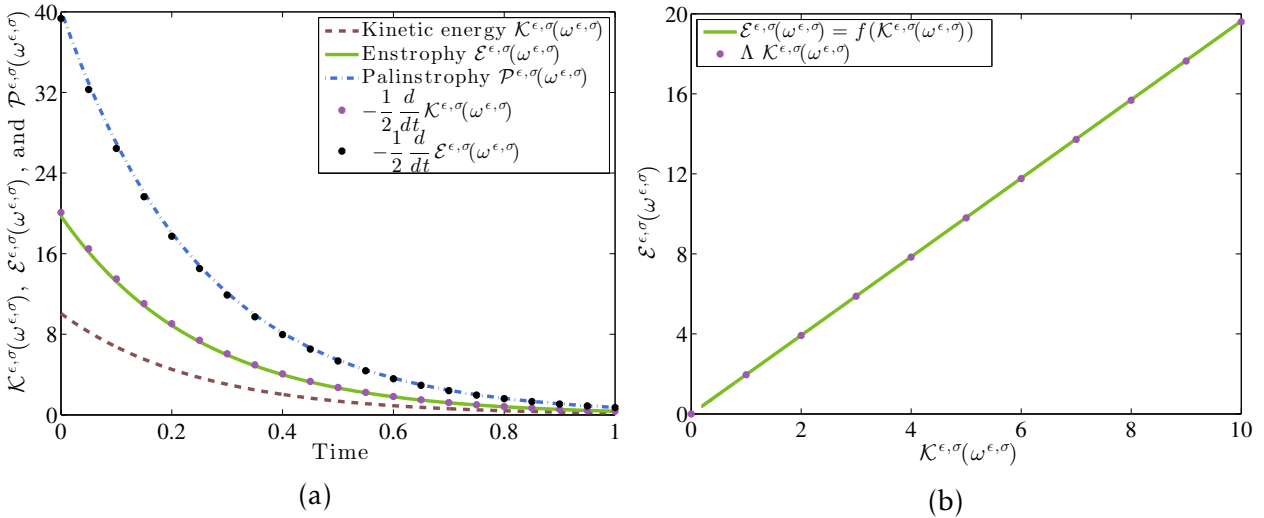


FIGURE 4.4 – (Taylor-Green vortex for  $\epsilon = 1$ , with initial condition  $\omega_{in}^A$ ). Evolution in time of the kinetic energy, enstrophy, palinstrophy, and time derivatives of the kinetic energy and enstrophy (a); Kinetic energy-enstrophy diagram showing that the Rayleigh quotient is constant over time (b). Parameters were  $N_x = N_y = 100$ ,  $\Delta t = 0.001$ ,  $T = 1$ ,  $\sigma = (\Delta x/L_x)^2$ .

As a second test, let us now perturb the initial condition as follows  $\omega_{in}^B = \omega_{in}^A + \beta \mathcal{N}(x, y)$ , where  $\beta = 0.1$  and  $\mathcal{N}(x, y)$  represents a noise, *i.e* a random value between 0 and 1, at each point  $(x, y)$ . Due to this noise, the Poisson bracket  $\{\omega^{\epsilon, \sigma}, \Psi^{\epsilon, \sigma}\}$  is no longer zero and the symmetry of the initial vorticity field is broken. In Figure 4.5, we plot the vorticity field  $\omega^{\epsilon, \sigma}$  at different times, obtained with the (DAMM)-scheme from the noisy initial condition  $\omega_{in}^B$ . Starting from four perturbed counter-rotating vortices (a), we observe this time the beginning of an instability and of the movement of the vortices (b). Then, a filamentation

of the vortices (c), and a wavy pattern of the two occurring stripes appears (d). Vorticity is tightened and elongated by the flow like a sort of cream. This continuous filamentation of the vorticity is related to the enstrophy cascade. Finally, the flow begins to stabilize (e), and to approach a two-stripe pattern (f), although no equilibrium seems to be attained at this time. This sequence of figures shows a classical aspect of turbulence in the two dimensional case, *i.e.* the appearance of coherent structures embedded in the filamentation picture. Here the initial characteristic length of the vorticity patch is  $L = \pi$  and attains  $2\pi$  at the final time. This instability was recently studied in [122] where similar figures were obtained, with a different numerical scheme, namely a high accuracy compact scheme on a non-uniform grid coupled with a fourth-order Runge-Kutta method.

As we made previously for the non-perturbed Taylor-Green flow, we investigate now the evolution over time of the quantities of interest. In Figure 4.6 (a), we plot the evolution over time of the kinetic energy and the enstrophy. One sees that due to the viscosity, the enstrophy and the kinetic energy decay in time, as expected. Yet, in Figure 4.6 (b) we observe several large picks of the palinstrophy around the physical times of  $t' = 75$  and  $t' = 85$ . These spontaneous outbursts correspond to the merging of the likely-colored vortices, merging during which the gradient of the vorticity is locally maximal. As expected, the sudden increase of palinstrophy from  $t' \approx 70$  to  $t' \approx 115$ , corresponding to the filamentation and the merging phases of vortices, coincides with the maximal decrease phases of both enstrophy and kinetic energy.

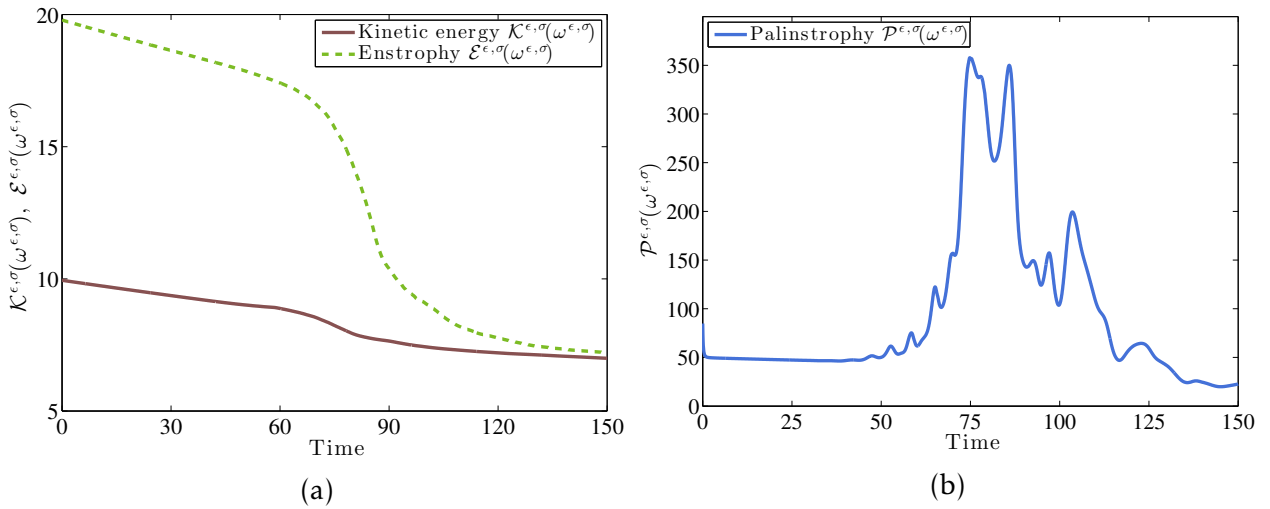


FIGURE 4.6 – (Taylor-Green vortex with initial condition  $\omega_{in}^B$ ). Evolution over time of kinetic energy and enstrophy (a), and of the palinstrophy (b). Parameters were  $\epsilon = 5e-4$ ,  $N_x = N_y = 256$ ,  $N_t = 2500$ ,  $T = 150$ , and  $\sigma = (\Delta x/L_x)^2$ .

This first test case permitted us to validate our numerical procedure in the non-stiff case  $\epsilon = 1$ . Indeed, the (DAMM)-scheme is able to recover the multiple relations linking several physical quantities of interest like the kinetic energy, the enstrophy, the palinstrophy and so one. Moreover, it permits also to observe an unstable flow when the initial condition of the Taylor-Green flow is perturbed. The next step is to focus on a different test case, namely the Kolmogorov flow, and to investigate the viscous-time asymptotics  $\epsilon \rightarrow 0$ .



## 4.4 Study of a forced case - Kolmogorov flow

Very often in applications, like atmosphere and ocean motions, the fluid develops robust large-scale coherent, essentially 2D structures, which are a result of many competing forces, like dissipation, nonlinearity and external forcing. We shall therefore study in this section such flows and their stability properties.

Let us thus consider (4.1) with forcing term  $\omega_{eq}(x, y) = \cos(x)$  and a domain  $\Omega = (-\pi, \pi) \times \left(-\frac{\pi}{K}, \frac{\pi}{K}\right)$  corresponding to a flat torus, where  $1/K \in \mathbb{R}^+$  denotes the aspect ratio of the domain. We associate to (4.1) periodic boundary conditions for both vorticity and stream-function. The system (4.1) rewrites then

$$(\mathcal{K})^\epsilon \begin{cases} \partial_t \omega^\epsilon + \frac{1}{\epsilon} \mathbf{u}^\epsilon \cdot \nabla \omega^\epsilon = (\Delta \omega^\epsilon + \cos(x)), & (t, \mathbf{x}) \in \mathbb{R}^+ \times \Omega, \\ -\Delta \Psi^\epsilon = \omega^\epsilon, & \mathbf{u}^\epsilon = {}^\perp \nabla \Psi^\epsilon, \quad (t, \mathbf{x}) \in \mathbb{R}^+ \times \Omega. \end{cases} \quad (4.13)$$

It is trivial to see that  $\omega^\epsilon \equiv \omega_{eq}$  is a steady-state solution of this system. There exists a large literature about problem (4.13). Historically, this flow was introduced by Kolmogorov (see [8] and [7]) as a part of his work concerning hydrodynamic stability problems and turbulence. Its linear stability was first investigated by Meshalkin and Sinai [108], and then completed by Yudovich [78]. They prove, using an elegant continuous fraction method, that when  $K > 1$  the steady-state solution  $\omega_{eq}$  is stable. However, when  $K < 1$ , they show that there exists a critical value of the viscosity below which the solution  $\omega_{eq}$  becomes unstable and then new stationary solutions appear. Furthermore, a global stability result was given in [105] for all forcing amplitudes when  $K = 1$ . There exists a large literature about the bifurcation phenomenon reminded before : in [115], the author investigates this latter for the stationary Navier-Stokes flows with some asymptotic results when the viscosity tends to 0 ; in [26] and [99], the Kolmogorov flow is studied in a more general context. In particular, the authors consider a more general forcing term, namely  $\omega_{eq}(x, y) = \sin(nx)$ , with  $n \in \mathbb{N}^*$ .

In this work, we consider the simplest forcing term, namely  $\omega_{eq} = \cos(x)$  and  $K < 1$ , corresponding to an unstable situation. The scope of this Section is twofold. First, a result about the linear instability of the flow is given result which is new to our knowledge. It consists in a relation between the parameter  $K$  and the instability growth rate. This result will be then validated with our numerical scheme. In the second part we investigate numerically the non-linear Kolmogorov instability. In particular, the AP-properties of the (DAMM)-scheme will be exploited to obtain directly the assumed final pattern of the flow. These numerical results could be a starting point for further analytical studies.

### 4.4.1 Linear instability of the Kolmogorov flow

When  $0 < K < 1$ , there exists a critical value of  $\epsilon$ , denoted by  $\epsilon^*$ , below which an instability appears. Since the linear instability occurs on the short time scale, it is suitable to work with the vorticity equation normalized to the non-viscous time scale :

$$(\mathcal{K}) \begin{cases} \partial_{t'} \omega + \mathbf{u} \cdot \nabla \omega = \epsilon (\Delta \omega + \cos(x)), & (t', \mathbf{x}) \in \mathbb{R}^+ \times \Omega, \\ -\Delta \Psi = \omega, & \mathbf{u} = {}^\perp \nabla \Psi, \quad (t', \mathbf{x}) \in \mathbb{R}^+ \times \Omega. \end{cases} \quad (4.14)$$

Let us set  $\tilde{\omega} := \omega - \omega_{eq}$ ,  $\tilde{\mathbf{u}} := \mathbf{u} - \mathbf{u}_{eq}$  and  $\tilde{\Psi} := \Psi - \Psi_{eq}$ , with  $\Psi_{eq}(x) = \cos(x)$  and  $\mathbf{u}_{eq} = (0, \sin(x))^T$ . Reformulating (4.14) in these fluctuating quantities, and neglecting the small terms of order 2, we obtain the linearized system :

$$(L) \begin{cases} \partial_{t'} \tilde{\omega} + \tilde{\mathbf{u}} \cdot \nabla \omega_{eq} + \mathbf{u}_{eq} \cdot \nabla \tilde{\omega} = \epsilon \Delta \tilde{\omega}, & (t', \mathbf{x}) \in \mathbb{R}^+ \times \Omega, \\ -\Delta \tilde{\Psi} = \tilde{\omega}, & \tilde{\mathbf{u}} = {}^\perp \nabla \tilde{\Psi}, \quad \mathbf{u}_{eq} = {}^\perp \nabla \Psi_{eq}, & (t', \mathbf{x}) \in \mathbb{R}^+ \times \Omega. \end{cases} \quad (4.15)$$

Combining all these equations, we can obtain an equation verified only by  $\tilde{\Psi}$  :

$$\partial_{t'} \Delta \tilde{\Psi} + \sin(x) \partial_y (\tilde{\Psi} + \Delta \tilde{\Psi}) = \epsilon \Delta^2 \tilde{\Psi}. \quad (4.16)$$

We now look for a solution of the equation (4.16) in the form  $\tilde{\Psi}(t', x, y) = \widehat{\Psi}(x) e^{\gamma t'} e^{\mathbf{i} K y}$ , with  $\gamma$  a complex number. Inserting this expression into the previous equation, we obtain an equation for the amplitude function  $\widehat{\Psi}$  : (we have dropped in the following the hat of  $\widehat{\Psi}$  for clarity reasons)

$$\gamma (\Psi''(x) - K^2 \Psi(x)) + \mathbf{i} K \sin(x) (\Psi''(x) + \eta \Psi(x)) = \epsilon (\Psi^{(4)}(x) - 2K^2 \Psi''(x) + K^4 \Psi(x)), \quad (4.17)$$

with  $\eta = 1 - K^2$ . The prime denotes the  $x$ -derivative. The equation (4.17) is a particular case of the so-called Orr-Sommerfeld equation (see [46]), with a sinusoidal velocity profile  $U(x) = \sin(x)$ . In the following, we neglect the viscosity term, (a choice which will be justified in Remark 4.1). Introducing moreover the parameter  $\widehat{\gamma} := \gamma/K$ , the equation (4.17) reads now :

$$-\mathbf{i} \widehat{\gamma} (\Psi''(x) - (1 - \eta) \Psi(x)) + \sin(x) (\Psi''(x) + \eta \Psi(x)) = 0, \quad x \in (-\pi, \pi), \quad (4.18)$$

which is known as the so-called Rayleigh equation, and completed with periodic boundary conditions. Both equations (4.17) and (4.18) with their boundary conditions can be considered as generalized eigenvalue problems of the unknowns  $(\Psi, \widehat{\gamma})$ . Their analysis, in particular the study of the spectrum, permits to have informations about the stability of the considered shear flow. Thus, in what regards equation (4.18), for a given value of  $\eta$ , solutions  $\Psi$  (eigenfunctions) exist only for a certain set of values of  $\widehat{\gamma}$ . In a more general case, namely for any velocity profile  $U(x)$ , the analysis of the Orr-Sommerfeld and of the Rayleigh equation is an old challenging problem. We refer to the book of Drazin & Reid and the references therein for a review of the analysis of these equations [46]. In particular, the analysis of the Rayleigh equation (4.18) can be viewed as the starting point of the stability study of the shear-flow. Naturally, as the Rayleigh equation is an ODE of order 2, it can provide only an approximation of two solutions of the Orr-Sommerfeld equation (4.17) which admits four independent solutions [46].

In our context, namely an unstable case with  $K \in (0, 1)$  and a simple sinusoidal velocity profile, numerical results show that there exists only one real and positive eigenvalue  $\widehat{\gamma}$  for which a solution  $\Psi$  exists in (4.18) (see [11] for more details about these results). Let us start from this knowledge.

We have already mentioned that for  $K > 1$ , that is  $\eta < 0$ , the system (4.17) is unconditionally stable (even for  $\epsilon = 0$ ) whereas for each  $\eta > 0$  there exists a threshold value of the viscosity denoted by  $\epsilon_\star(\eta)$ , below which an instability occurs. The function  $\epsilon_\star(\eta)$  is strictly

increasing, thus invertible with monotonically increasing inverse  $\eta_\star(\epsilon)$ . For fixed  $\epsilon \leq 0$ , an instability occurs for all  $\eta$  above  $\eta_\star(\epsilon)$  and one can speculate that  $\eta_\star(\epsilon) \rightarrow 0$ .

In Figure 4.7, we have plotted a sketch of the expected modulus of the solution  $\Psi$  of (4.18) which satisfies the boundary conditions. One notes that when  $\sin(x)$  can be considered of order unity, the solution  $\Psi(x) = \text{const}$  is sufficiently good. However near  $x = 0$  and  $x = \pm\pi$ ,  $\sin(x)$  tends to zero and the solution may pick up an appreciable value of  $\Psi''(x)$ , so it varies rapidly in the neighbourhood of these points.

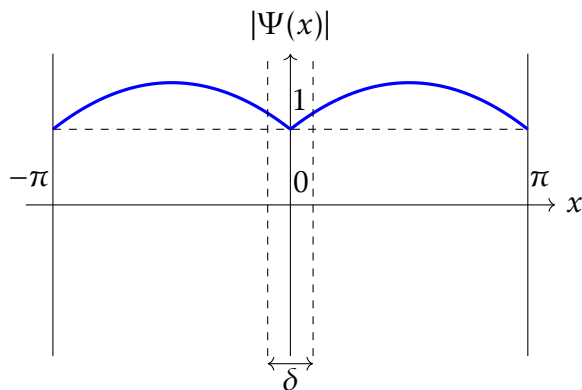


FIGURE 4.7 – Sketch of the expected modulus of the solution  $\Psi(x)$  of (4.18).

One can therefore consider (4.18) as a boundary/internal layer problem. A frequently employed method to deal with such problems is the matched asymptotic expansion method (see [47, 76, 87] for a detailed introduction to these methods). One divides the domain in regions, called the inner and the outer regions. In the outer regions, the variations of the function  $\Psi$  are of order 1, so that one can construct solutions in these regions (called the outer solutions) by solving the equation obtained when neglecting the terms of order  $\widehat{\gamma}$  and  $\eta$  in (4.18). In the inner regions, near the boundary layers at  $x = 0$  and  $x = \pm\pi$ , the variations of the function  $\Psi$  are very rapid, of order  $1/\widehat{\gamma}^\alpha$  for  $\alpha > 0$ . One looks for corresponding approximate solutions called the inner solution. The method can work if the regions of validity of the inner and the outer solutions overlap. In these overlapping regions, one has two asymptotically equivalent solutions which then allow one deriving relations between the arbitrary constants of the solutions of the differential equations in any given region. This matching process can be carried out under the condition that some constraint holds. This constraint constitutes the sought-for relation between the small parameter and the eigenvalue.

We have carried out the above mentioned procedure and the main finding is that solutions can be found if the growth rate  $\widehat{\gamma}$  is proportional to  $\eta$ . This is briefly sketched below.

As far as the outer solutions are concerned, one finds that there exists a small number  $\delta$ , depending on  $\eta$  (so that  $\delta$  tends to zero when  $\eta$  tends to zero) such that, in the intervals  $I_+^\delta = [\delta/2, \pi - \delta/2]$  and  $I_-^\delta = [-\pi + \delta/2, -\delta/2]$ , the outer solution writes :

$$\Psi^{\mathbf{O}}(x) = A_+ \left( 1 - \frac{\eta}{2} x^2 + \frac{\eta \pi}{2} |x| - \frac{\eta}{2} \left( \frac{\pi}{2} \right)^2 - \mathbf{i} \widehat{\gamma} (G(x) + C_1) \right) + \mathcal{R}_+(x, \eta, \delta), \quad x \in I_+^\delta, \quad (4.19)$$

$$\Psi^{\mathbf{O}}(x) = A_- \left( 1 - \frac{\eta}{2} x^2 + \frac{\eta \pi}{2} |x| - \frac{\eta}{2} \left( \frac{\pi}{2} \right)^2 - \mathbf{i} \widehat{\gamma} (G(x) - C_1) \right) + \mathcal{R}_-(x, \eta, \delta), \quad x \in I_-^\delta, \quad (4.20)$$

where  $G(x) = x \ln(|\tan(x/2)|) - 2 \int_0^{\tan(x/2)} \frac{\arctan(t)}{t} dt$ ,  $A_\pm$  is a complex constant to be defined and  $C_1 = 2 \int_0^1 \frac{\arctan(t)}{t} dt$ . The terms  $\mathcal{R}_+$ ,  $\mathcal{R}_-$  are residual parts of higher order in  $\eta$ . These solutions were obtained by expanding  $\Psi$  in powers of  $\eta$  and, assuming symmetry around  $x = \pm\pi/2$ .

As far as the inner equation is concerned, it is customary to work with a stretched variable. For the equation around  $x = 0$ , one chooses  $\xi = x/\widehat{\gamma}$ . By approximating  $\sin(x) \approx x$  and introducing further the complex variable  $z = 1 + \mathbf{i}\xi$ , one can write the inner equation in the form :

$$\frac{d^2 \Psi^{\mathbf{I}}}{dz^2}(z) + \frac{\widehat{\gamma}^2}{z} \Psi^{\mathbf{I}}(z) = 0. \quad (4.21)$$

This equation can be solved in various ways in terms of known functions. As far as asymptotic matching is concerned, to leading order, the simplest way is to expand  $\Psi^{\mathbf{I}}$  in powers of  $\widehat{\gamma}^2$ , with the lowest order solution being a constant. The end result is

$$\Psi^{\mathbf{I}}(\xi) = a + \widehat{\gamma}^2 (\mathbf{i} b \xi - a H(\xi)) + \mathcal{R}(\widehat{\gamma}, \xi, \Delta), \quad (4.22)$$

in terms of the stretched variable, where  $a$  is a real fixed constant,  $b$  a complex constant to be determined,  $H(\xi) = (1 + \mathbf{i}\xi) \ln(1 + \mathbf{i}\xi)$  and  $\mathcal{R}$  a residual term of order higher than  $\eta^2$ . Matching the outer solutions (4.19) for small  $x$  with the inner solution (4.21) for large  $\xi$ , introduces relations between the various coefficients that imply a constraint between  $\eta$  and  $\widehat{\gamma}$ , namely  $\widehat{\gamma} = \eta$  to leading order.

Let us remark here that the common feature of both the inner and outer solutions is that, for the problem at hand, both can be derived as an expansion in powers of  $\eta$  and  $\widehat{\gamma}$ , with the lowest order being a constant. This is consistent with the sketch of Figure 4.7.

*Now, we propose in detail a different, more synthetic method in order to derive the overall solution  $\Psi$  of (4.18) and find again the link between  $\widehat{\gamma}$  and  $\eta$ .*

We begin by recasting (4.18) in the form

$$\Psi''(x) = Q(x)\Psi(x), \quad (4.23)$$

where

$$Q(x) = -\eta - \frac{\mathbf{i}\widehat{\gamma}}{-\mathbf{i}\widehat{\gamma} + \sin(x)}. \quad (4.24)$$

Note that  $Q(x)$  is dependent of  $\eta$  and  $\widehat{\gamma}$  so that it can be considered as small except in a narrow region around  $x = 0$  and  $x = \pm\pi$ .

To the lowest order in  $\widehat{\gamma}$  and  $\eta$ ,  $\Psi(x)$  is a linear function whose slope must be set to zero in order to satisfy the boundary conditions.  $\Psi(0)$  is then a constant which can be fixed to unity in full generality. One can then write :

$$\Psi(x) = 1 + \widetilde{\Phi}(x), \quad (4.25)$$

where  $\widetilde{\Phi}(x)$  is a small correction that embodies the dependence on  $\eta$  and  $\widehat{\gamma}$ . Proceeding iteratively, and neglecting the product  $\widetilde{\Phi}(x)Q(x)$  with respect to  $Q$ , one obtains

$$\widetilde{\Phi}''(x) = Q(x). \quad (4.26)$$

We start by giving a result which allows finding the relation between  $\widehat{\gamma}$  and  $\eta$  :

**Proposition 2** *To satisfy the periodic boundary condition  $\widetilde{\Phi}'(\pi) = \widetilde{\Phi}'(-\pi)$ , it is necessary to have*

$$\frac{\widehat{\gamma}}{(1 + \widehat{\gamma}^2)^{1/2}} = \eta. \quad (4.27)$$

**Proof:** Integrating once (4.26) gives

$$\widetilde{\Phi}'(x) = \int_0^x Q(x') dx' + C. \quad (4.28)$$

Then, the periodic boundary condition  $\widetilde{\Phi}'(\pi) = \widetilde{\Phi}'(-\pi)$  requires :

$$\int_{-\pi}^{+\pi} Q(x') dx' = 0. \quad (4.29)$$

A primitive of  $Q$  can be computed analytically :

$$\int Q(x') dx' = -\eta x + \frac{2\widehat{\gamma}}{\sqrt{1 + \widehat{\gamma}^2}} \arctan\left(\frac{\mathbf{i} + \widehat{\gamma} \tan(x/2)}{\sqrt{1 + \widehat{\gamma}^2}}\right),$$

and then, the constraint (4.29) leads to the desired result. ■

Note that the relation (4.27) is equivalent to  $\widehat{\gamma} \approx \eta$ , to leading order in  $\widehat{\gamma}$ .

Another quantity which permits to validate our numerical method is the phase difference of the solution between the mid points at  $x = \pm\pi/2$  and the reference point  $x = 0$ , where, by construction,  $\Psi(0) = 1$ . The phase is given by :

$$\theta_{\pm} := \arctan\left[\frac{\text{Im}(\Psi(\pm\pi/2))}{\text{Re}(\Psi(\pm\pi/2))}\right]. \quad (4.30)$$

Before giving the expression of  $\theta_{\pm}$ , we state the following Lemma :

**Lemma 2** *Consider  $a > 0$ , one has :*

$$\begin{aligned} \int \frac{x'}{\mathbf{i} \sin(x') + a} dx' &= \frac{1}{\sqrt{1 + a^2}} \left[ \mathbf{i} x \left( \ln \left( 1 + \frac{e^{\mathbf{i}x}}{a + \sqrt{1 + a^2}} \right) - \ln \left( 1 - \frac{e^{\mathbf{i}x}}{-a + \sqrt{1 + a^2}} \right) \right) \right. \\ &\quad \left. + \text{Li}_2 \left( \frac{-e^{\mathbf{i}x}}{a + \sqrt{1 + a^2}} \right) - \text{Li}_2 \left( \frac{e^{\mathbf{i}x}}{-a + \sqrt{1 + a^2}} \right) \right], \end{aligned}$$

where  $\text{Li}_2$  denotes the dilogarithm function [132].

Now, we give the following proposition :

**Proposition 3** *To the leading order in  $\widehat{\gamma}$ , we have :*

$$\theta_{\pm} = \arctan \left[ \frac{\text{Im}(\Psi(\pm\pi/2))}{\text{Re}(\Psi(\pm\pi/2))} \right] = \arctan \left( \frac{\pm C_1 \widehat{\gamma}}{1 + \eta \pi^2/8} \right) + O(\widehat{\gamma}^2 \ln(\widehat{\gamma})) + O(\widehat{\gamma}^2),$$

where  $C_1 = 2 \int_0^1 \frac{\arctan(t)}{t} dt$ .

**Proof:** Let us compute in this proof the values  $\widetilde{\Phi}(\pm\pi/2)$  permitting then the computation of  $\theta_{\pm}$ . By integrating  $\widetilde{\Phi}'$ , given in (4.28) one has :

$$\widetilde{\Phi}(x) = -\eta \frac{x^2}{2} + Cx + \frac{2\widehat{\gamma}}{\sqrt{1+\widehat{\gamma}^2}} \left[ \int_0^x \arctan \left( \frac{\mathbf{i} + \widehat{\gamma} \tan(x'/2)}{\sqrt{1+\widehat{\gamma}^2}} \right) dx' - x \arctan \left( \frac{\mathbf{i}}{\sqrt{1+\widehat{\gamma}^2}} \right) \right] + C_2.$$

The constant  $C_2$  is equal to 0 since  $\widetilde{\Phi}(0) = 0$ .

We focus now on the computation of a primitive of  $\arctan \left( \frac{\mathbf{i} + \widehat{\gamma} \tan(x/2)}{\sqrt{1+\widehat{\gamma}^2}} \right)$ . By integrating by parts, we obtain :

$$\int \arctan \left( \frac{\mathbf{i} + \widehat{\gamma} \tan(x'/2)}{\sqrt{1+\widehat{\gamma}^2}} \right) dx' = x \arctan \left( \frac{\mathbf{i} + \widehat{\gamma} \tan(x/2)}{\sqrt{1+\widehat{\gamma}^2}} \right) - \frac{\sqrt{1+\widehat{\gamma}^2}}{2} \int \frac{x'}{\mathbf{i} \sin(x) + \widehat{\gamma}} dx',$$

and the primitive is obtained from the Lemma 4.2. Thus, the periodic boundary condition  $\widetilde{\Phi}(\pi) = \widetilde{\Phi}(-\pi)$  fixes the value of  $C$  by :

$$C = \frac{\widehat{\gamma}}{2\pi} \int_{-\pi}^{\pi} \frac{x'}{\mathbf{i} \sin(x') + \widehat{\gamma}} dx' + \frac{2\widehat{\gamma}}{\sqrt{1+\widehat{\gamma}^2}} \arctan \left( \frac{\mathbf{i}}{\sqrt{1+\widehat{\gamma}^2}} \right).$$

To the leading order in  $\widehat{\gamma}$ , we have thus

$$C = -\frac{\mathbf{i}\widehat{\gamma}}{\sqrt{1+\widehat{\gamma}^2}} \ln \left( \frac{2}{\widehat{\gamma}} \right) + \frac{2\widehat{\gamma}}{\sqrt{1+\widehat{\gamma}^2}} \arctan \left( \frac{\mathbf{i}}{\sqrt{1+\widehat{\gamma}^2}} \right) + O(\widehat{\gamma}^2).$$

This leads to,

$$\widetilde{\Phi}(\pi/2) = -\eta \frac{\pi^2}{8} - \frac{\mathbf{i}\pi\widehat{\gamma}}{2\sqrt{1+\widehat{\gamma}^2}} \ln \left( \frac{2}{\widehat{\gamma}} \right) + \frac{2\widehat{\gamma}}{\sqrt{1+\widehat{\gamma}^2}} \int_0^{\pi/2} \arctan \left( \frac{\mathbf{i} + \widehat{\gamma} \tan(x'/2)}{\sqrt{1+\widehat{\gamma}^2}} \right) dx' + O(\widehat{\gamma}^2).$$

Now, the integral term can be reformulated as :

$$\begin{aligned} \frac{2\widehat{\gamma}}{\sqrt{1+\widehat{\gamma}^2}} \int_0^{\pi/2} \arctan \left( \frac{\mathbf{i} + \widehat{\gamma} \tan(x'/2)}{\sqrt{1+\widehat{\gamma}^2}} \right) dx' &= \frac{\pi\widehat{\gamma}}{\sqrt{1+\widehat{\gamma}^2}} \arctan \left( \frac{\mathbf{i} + \widehat{\gamma}}{\sqrt{1+\widehat{\gamma}^2}} \right) \\ &\quad - \widehat{\gamma} \int_0^{\pi/2} \frac{x'}{\mathbf{i} \sin(x') + \widehat{\gamma}} dx', \end{aligned}$$

which yields, at leading order in  $\widehat{\gamma}$ ,

$$\begin{aligned} \frac{2\widehat{\gamma}}{\sqrt{1+\widehat{\gamma}^2}} \int_0^{\pi/2} \arctan\left(\frac{\mathbf{i} + \widehat{\gamma} \tan(x'/2)}{\sqrt{1+\widehat{\gamma}^2}}\right) dx' &= \frac{\mathbf{i}\pi\widehat{\gamma}}{2\sqrt{1+\widehat{\gamma}^2}} \ln\left(\frac{2}{\widehat{\gamma}}\right) + \frac{\pi^2\widehat{\gamma}}{4\sqrt{1+\widehat{\gamma}^2}} + \frac{\mathbf{i}\widehat{\gamma}}{\sqrt{1+\widehat{\gamma}^2}} C_1 \\ &+ O(\widehat{\gamma}^2 \ln(\widehat{\gamma})) + O(\widehat{\gamma}^2), \end{aligned}$$

where  $C_1$  is equal to the convergent integral  $2 \int_0^1 \frac{\arctan(t)}{t} dt$ . Finally, we have :

$$\widetilde{\Phi}(\pi/2) = \eta \frac{\pi^2}{8} + \mathbf{i} C_1 \frac{\widehat{\gamma}}{\sqrt{1+\widehat{\gamma}^2}} + O(\widehat{\gamma}^2 \ln(\widehat{\gamma})) + O(\widehat{\gamma}^2) = \eta \frac{\pi^2}{8} + \mathbf{i} C_1 \widehat{\gamma} + O(\widehat{\gamma}^2 \ln(\widehat{\gamma})) + O(\widehat{\gamma}^2).$$

Analogously, we can show that :

$$\widetilde{\Phi}(-\pi/2) = \eta \frac{\pi^2}{8} - \mathbf{i} C_1 \widehat{\gamma} + O(\widehat{\gamma}^2 \ln(\widehat{\gamma})) + O(\widehat{\gamma}^2),$$

which ends the proof. ■

Note that at leading order in  $\widehat{\gamma}$ , we find  $\widetilde{\Phi}'(\pi/2) = \widetilde{\Phi}'(-\pi/2) = 0$  as we see on the sketch of the modulus of  $\widetilde{\Psi}$  in Figure 4.7.

**Summary of the result.** Let us summarize in a few words the last paragraphs. Starting from an unstable situation ( $0 < K < 1$ ), and assuming that  $\eta$  and  $\widehat{\gamma}$  are small parameters of the same order, we proved that it was necessary to have the condition  $\eta = \widehat{\gamma}$  in order to respect the periodic boundary condition. This condition gives the relation between aspect ratio  $1/K$  and growth rate  $\gamma$  for the linear instability phase, namely  $\eta = 1 - K^2 = \gamma/K$ .

### Numerical validation of the analytical instability result

In this section, we validate the results obtained in the previous subsection, namely the fact that in the linear instability phase at leading order,  $\widehat{\gamma} = \eta = 1 - K^2$ . In order to do so, let us consider the following system :

$$\begin{cases} \partial_{t'} \widetilde{\omega} + \mathbf{u} \cdot \nabla \widetilde{\omega} = 0, & (t', \mathbf{x}) \in \mathbb{R}^+ \times (-\pi, \pi) \times \left(-\frac{\pi}{K}, \frac{\pi}{K}\right), \\ -\Delta \widetilde{\Psi} = \widetilde{\omega}, & \mathbf{u} = {}^\perp \nabla \widetilde{\Psi}, \quad (t', \mathbf{x}) \in \mathbb{R}^+ \times (-\pi, \pi) \times \left(-\frac{\pi}{K}, \frac{\pi}{K}\right), \end{cases} \quad (4.31)$$

obtained from (4.14), by setting  $\widetilde{\omega} = \omega - \omega_{eq}$ ,  $\widetilde{\mathbf{u}} = \mathbf{u} - \mathbf{u}_{eq}$ , and  $\widetilde{\Psi} = \Psi - \Psi_{eq}$ , and neglecting the viscous term (as we did in our analytical linear study). We solve (4.31) with the (DAMM)-scheme and we consider the following initial condition  $\omega_{in}^D(x, y) = \cos(x) + \beta_1 \cos(Ky)$ , with  $\beta_1 = 1 \times 10^{-6}$ . In the following, we consider two numerical tests.

#### Phase computation .

From Proposition 4.3, we have :

$$\theta_+ = \text{Arctan}\left(\frac{C_1 \widehat{\gamma}}{1 + \frac{\pi^2 \eta}{8}}\right) \approx \left(\frac{C_1 \widehat{\gamma}}{1 + \frac{\pi^2 \eta}{8}}\right) := \vartheta. \quad (4.32)$$

Furthermore, we have  $\widetilde{\Psi}(t', \pi/2, y) = \widehat{\Psi}(\pi/2) e^{\gamma t'} e^{iKy} = e^{\gamma t'} |\widehat{\Psi}(\pi/2)| e^{i(Ky+\theta_+)}$ . And then, for a fixed  $t'$ , one has

$$\arg \max_{y \in [-\pi/K, \pi/K]} \mathcal{R}e\left(\widetilde{\Psi}(t', \pi/2, y)\right) := y_M = -\theta_+/K \approx -\vartheta/K.$$

This last relation allows us to validate numerically the linear instability result. Indeed, finding  $y_M$  numerically, we can compare the values of  $K y_M$  with the analytical value of  $\vartheta$  obtained from (4.32). In Figure 4.8 (a), we plot the numerical as well as analytical values of  $\vartheta$  as a function of the aspect ratio  $1/K$ . We observe a good correspondance between the two curves, especially for small values of the aspect ratio. In Figure 4.8 (b), the relative error between the values of  $\vartheta$  obtained analytically and numerically is plotted. We remark that this error is mainly due to the error of the analytical estimate of the growth rate and increases with the aspect ratio. Indeed, we have proven that at leading order,  $\widehat{\gamma} = \eta$ . But at the superior order, we have  $\widehat{\gamma} = \eta + f(\eta)\eta^2$  where  $f$  is an unknown function satisfying  $\lim_{\eta \rightarrow 0} f(\eta) < \infty$ . Additional effects are : the truncation error of the numerical scheme (including the error caused by the stabilization parameter  $\sigma$ ) and the error due to the determination of  $y_M$ . Indeed, the value of  $y_M$  is obtained numerically by finding the maximum of  $\widetilde{\Psi}$  in  $y$  for  $x = \pi/2$ , an error of the order of  $\Delta y$  is then present for each point. To minimize this error, we took  $N_y = 256$ . Note that this error worsens when the aspect ratio increases, since  $L_y$  is proportional to this latter.

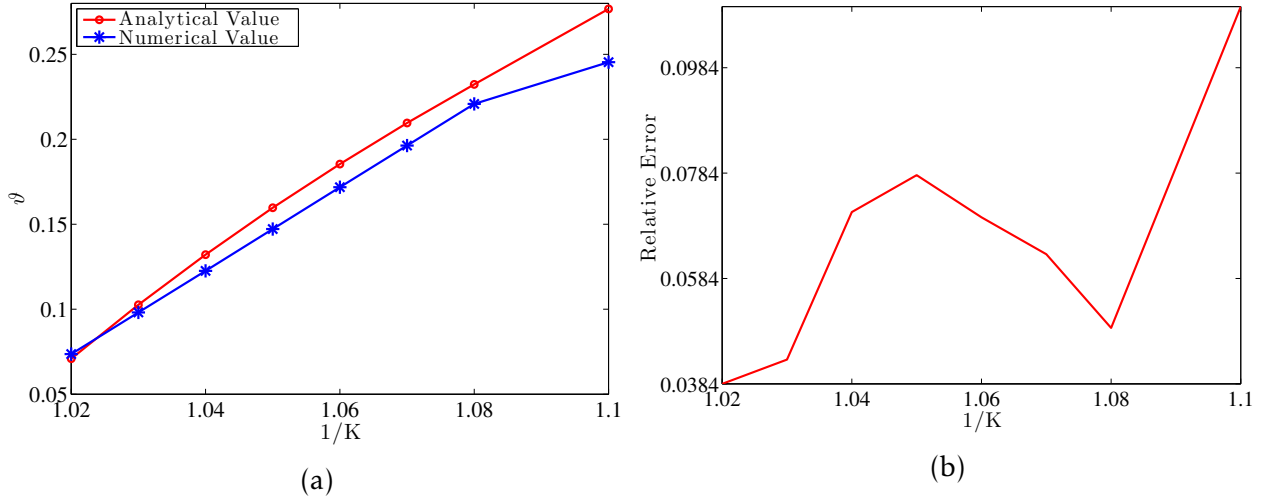


FIGURE 4.8 – (Kolmogorov case with initial condition  $\omega_{in}^D$ ). Evolution of the phase  $\vartheta$  as a function of the aspect ratio  $1/K$  (a), and relative errors between numerical and analytical values of the phase (b). Here,  $N_x = N_y = 256$ ,  $N_t = 200$ ,  $\mathcal{T} = 20$ ,  $\sigma = (\Delta x/L_x)^3$ .

### Growth rate .

In the linear phase, we have  $\widetilde{\Psi}(t', x, y) = e^{\gamma t'} e^{iKy} \widehat{\Psi}(x)$ . Thus,

$$\mathcal{S}(t) := \ln\left(\|\widetilde{\Psi}(t, \cdot, \cdot)\|_{L^\infty(\Omega)}\right) = \gamma t'. \quad (4.33)$$

In Figure 4.9 (a), we plot the evolution of the value  $\gamma$  obtained numerically with the (DAMM)-scheme using (4.33) and the value of  $K\eta$  for several values of the aspect ratio. As we observed for the phase computation, the comparison between the numerical value and the analytical one seems good, especially when the aspect ratio approaches 1. In Figure



4.9 (b), the relative error between the analytical and numerical values is plotted. Note that for values of the aspect ratio  $1/K$  approaching 1, the relative error increases. This is due to the fact that in this limit, the boundary layer at  $x = 0$  becomes sharp. The effect is therefore linked with the truncation error of the numerical scheme. For large values of  $\eta$ , the error is mainly due to the approximation  $\widehat{\gamma} = \eta + f(\eta)\eta^2$  and it worsens when the aspect ratio increases. Nevertheless, the relative error remains limited and consistent with our analytical result.

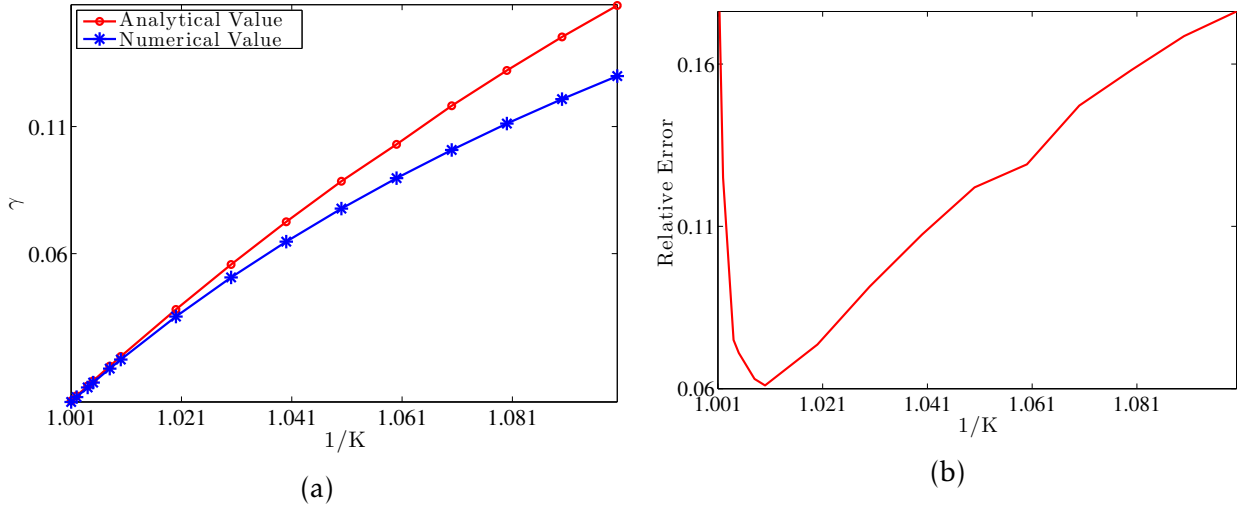


FIGURE 4.9 – (Kolmogorov case with initial condition  $\omega_{in}^D$ ). Evolution of the growth rate  $\gamma$  as a function of the aspect ratio  $1/K$  (a), and relative errors between numerical and analytical values of the growth rate (b). Here,  $N_x = N_y = 128$ ,  $N_t = 500$ ,  $T = 50$ ,  $\sigma = (\Delta x/L_x)^3$ .

**Remark 10** We recall that in order to obtain the equation (4.18), we neglected the viscosity term. To be more precise, this approximation is valid when the viscosity  $\epsilon$  verifies the following condition :  $\epsilon \ll \gamma \rho^2$ , where  $\rho$  denotes the boundary layer thickness. Following our analysis, this condition becomes  $\epsilon \ll \eta^3$ . Since  $\eta$  tends to 0 when the aspect ratio tends to 1, we observe that the viscosity cannot be neglected in this limit. We can argue that for finite viscosity, the mode becomes stable if  $\eta \rightarrow 0$ . This conclusion is consistent with the stability result obtained by Meshalkin and Sinai [108], indicating that the instability region is attained for aspect ratio strictly superior to 1 for a finite value of viscosity.

### Digression : similarities and differences with tearing modes

At this point it is interesting to examine the analogy between the instability in the Kolmogorov flow and the tearing mode occurring in a magnetized plasma [61]. In the 2D MHD model with periodic boundary conditions (see for example [39]), the equivalent of the equilibrium vorticity  $\omega_{eq} = \cos(x)$  is an equilibrium current  $J_{eq} = \cos(x)$  and its associated equilibrium magnetic flux function  $\Psi_{eq} = \cos(x)$ .

Carrying out a linear analysis similar to the one of the previous sections, one ends up with an outer equation for the perturbed flux function  $\Psi$  of the form

$$\sin(x)(\Psi''(x) + \eta \Psi(x)) = 0, \quad x \in (-\pi, \pi). \quad (4.34)$$

This is reminiscent of equation (4.18), when the inertia term is ignored. One can then expect a linear mode structure similar to the one sketched in Figure 4.7. The analogy however stops here. There is a fundamental difference between the instability of the Kolmogorov

flow and the tearing mode instability which is the fact that the former exists in the ideal fluid, whereas the latter requires dissipation. As it is known from the theory, the tearing mode growth rate scales like a fractional power of the resistivity (usually  $\gamma \sim \eta^{3/5}$ , for sufficiently small resistivity), which justifies neglecting the inertia in the limit of small resistivity. In the language of plasma physics this is called a *resistive instability*. Instead, we have seen that the Kolmogorov flow is unstable for arbitrarily small values of the viscosity provided that the aspect ratio exceeds unity, and that its growth rate depends only on a geometrical factor, the aspect ratio. Such a situation would be described as an *ideal instability* in the context of plasma physics.

This difference, resistive versus ideal instability, has a consequence on our understanding of the nonlinear phase of the instability. In the case of the tearing instability, an analytic theory for the nonlinear growth phase [121] and its saturation [49, 109] could be developed by neglecting the plasma inertia. Instead, in the case of the instability of the Kolmogorov flow, no such theory exists, to our knowledge. One can however rely on numerical simulations, as presented in the next sub-sections.

#### 4.4.2 Kolmogorov flow : simulations of the instability

Let us perform in this subsection some simulations of the whole phases (linear, nonlinear and final) of the Kolmogorov flow, starting from an unstable situation. Let us remark that for small perturbation amplitudes, the modes do not interact with each other (at least for some time) and the linear stability theory performed previously can be applied. As the perturbed become however larger, one cannot ignore any more the higher order nonlinear terms, which are responsible for the mode-mode interactions and lead to interesting phenomenon and to the nonlinear saturation. Analytical investigations become however very difficult and we shall accommodate with numerical simulations.

Let  $\Omega_S = (0, 2\pi) \times (0, 2\pi/K)$  and let us supply the system (4.11) with the initial condition  $\omega_{in}^E(x, y) = \omega_{eq}(x) + \beta_3 \mathcal{N}(x, y)$ , where  $\beta_3 = 0.01$  and  $\mathcal{N}$  represents the noise introduced before. In this way, we perturb the steady-state solution of the Kolmogorov flow. We choose an aspect ratio of  $1/K = 1.1$  and a viscosity of  $\epsilon = 1 \times 10^{-3}$ . The final physical time is  $\mathcal{T} = 5000$ . In Figure 4.10, we plot the evolution of the vorticity  $\omega^{\epsilon, \sigma}$  at different times, obtained with our (DAMM)-scheme. We observe again the formation of a variety of flow patterns. Starting from the initial condition (panel (a)), the first stage of the instability is clearly observable in panel (b). Then, we observe during a long time a complex behaviour characterized by the formation of the coherent structures embedded in a sea of filaments (panel (c) and (d)). And finally, in the long-time asymptotics, the vorticity begins to saturate (panel (e)), and a steady state constituted by a pair of opposed vortices is obtained (panel (f)). For this final state, the external forcing term, which adds continuously kinetic energy to the system, is balanced by the viscous dissipation. Note that the wavy pattern obtained during the instability phase is very similar to the one obtained in the Taylor-Green instability (see Figure 4.6).

In order to understand in more details the Kolmogorov instability evolution, we plot in Figure 4.11 three panels corresponding to the evolution of the discrete  $L^\infty$ -norm of  $\tilde{\Psi}$  in three adjacent time-intervals. Dividing the curve in three plots permits to distinguish more clearly all phases of the instability process. In panel (a), we observe the events which occur at short times. The linear phase, characterized by the relation between the aspect ratio and the growth rate, appears first. A nonlinear stage follows immediately, characterized by a strong growth of the norm. This nonlinear phase continues in panel (b) but with a complete different evolution. After the growth of the norm in panel (a), we observe in panel (b)

some damped oscillations and the decrease of the norm. This time-period corresponds to the merging and the filamentation of the vortices observed in Figure 4.10. This oscillating phase is the most interesting one. Indeed, although the filamentation and the merging of the vortices lead to a very chaotic situation, the damped oscillations of the norm of  $\widetilde{\Psi}$  seem to have more ordered variations. Finally, during the third and longest part of the instability, the norm in panel (c) continues to decrease but without any oscillation (or very weak oscillations impossible to see) before it saturates. This saturation phase means that the new equilibrium is attained.

To end this section, let us finally take a cast look at the last panel in Figure 4.10, which suggests to study the relation between the vorticity and the stream function at the final time. Indeed, since the final pattern (panel (f)) in Figure 4.10 does not evolve anymore, the Poisson bracket  $\{\omega^{\epsilon,\sigma}, \Psi^{\epsilon,\sigma}\}$  should be of the same order as the term  $\epsilon(\Delta\omega^{\epsilon,\sigma} + \cos(x))$ . Since we have chosen  $\epsilon = 1. \times 10^{-3}$ , which is small, we should approach in this time asymptotic a functional relation between  $\omega^{\epsilon,\sigma}$  and  $\Psi^{\epsilon,\sigma}$ . In Figure 4.12 (a), we show the final state of the stream function ( $-\Psi^{\epsilon,\sigma}$ ) corresponding to the vorticity field plotted in the Figure 4.10 (panel (f)). As expected, we observe that the vorticity and the stream-function have very similar patterns, with the vorticity more concentrated at the extremal points of the field. In Figure 4.12 (b), the  $\omega^{\epsilon,\sigma} - \Psi^{\epsilon,\sigma}$ -diagram at the final physical time  $\mathcal{T}$  is shown. Despite an accumulation of points around the coordinates  $(-0.25, 0.15)$  and  $(0.25, -0.15)$ , the vorticity and the stream-function seem to be linked with a *sinh* functional relation. We shall investigate in more details this point by studying the limit case  $\epsilon \rightarrow 0$  in the next subsection.

#### 4.4.3 Kolmogorov flow for vanishing values of $\epsilon$

In this section, we consider  $\epsilon = 0$  and the initial condition  $\omega_{in}^E$ . In this limit, the physical time  $\mathcal{T}$  tends to infinity. This allows to study long-time asymptotic of the flow. One of the strengths of the *Asymptotic-Preserving* properties of the (DAMM)-scheme is to allow attaining directly the equilibrium state without computing the intermediate states of the solution. Therefore, we take a large  $\Delta t$ , and obtain the final state without error accumulation. In Figure 4.13, we plot the vorticity field  $\omega^{0,\sigma}$  in the limit regime at the initial time  $n = 0$  and at the final time step  $n = 50$ . We see that the pattern obtained in this way is very similar to that obtained in the non-limit case (see Figure 4.10 (f)). Note that the model that we study has symmetry translation in the  $y$ -direction. As a consequence, the final position of the vortices in the  $y$ -direction depends on the initial condition. In Figure 4.14 (a), we have plotted the  $\omega^{0,\sigma} - \Psi^{0,\sigma}$ -diagram at the initial time. As expected, there is no functional relation between the two fields due to the noise introduced at this time. In Figure 4.14 (b), the same  $\omega^{0,\sigma} - \Psi^{0,\sigma}$  diagram is obtained after 50 time iterations, and we see the *sinh* functional relation already observed in the non-limit case (see Figure 4.12 (b)). In order to validate our impression of a *sinh*-relation between  $\omega^{0,\sigma}$  and  $-\Psi^{0,\sigma}$ , we plotted on the same Figure 4.14 (b) a fitting function  $f_{fit}(\cdot) = -a \sinh(b \cdot)$ , with  $a$  and  $b$  two positive constants. The optimal values  $a = 0.2255$  and  $b = 2.7467$  were obtained by using an optimization Matlab tool. The fitting curve approaches well the numerical curve obtained with the (DAMM)-scheme. Keeping in mind that the limit stream function  $\Psi^{0,\sigma}$  verifies the relation  $-\Delta\Psi^{0,\sigma} = \omega^{0,\sigma} = -a \sinh(b\Psi^{0,\sigma})$ , this permits to find the form of the solution  $\Psi^{0,\sigma}$ . Indeed performing the following change of variables :  $\Phi^{0,\sigma} = b\Psi^{0,\sigma}$ , the function  $\Phi^{0,\sigma}$  verifies then the *sinh*-relation  $\Delta\Phi^{0,\sigma} = ab \sinh(\Phi^{0,\sigma})$ . Inspired by the work of [69], where the previous *sinh*-Poisson relation is investigated, we consider the following fitting

solution  $\Psi_{fit}$  defined by

$$\Psi_{fit}(x, y) := \frac{4}{b} \tanh^{-1} \left( \frac{\sqrt{k} \operatorname{cn}(rx, k) - \sqrt{k_1} \operatorname{cn}(sy + \varphi, k_1)}{1 + \sqrt{k k_1} \operatorname{cn}(rx, k) \operatorname{cn}(sy + \varphi, k_1)} \right),$$

where  $\operatorname{cn}$  denotes the Jacobi elliptic function of the first kind, and  $\varphi \in [0, 2\pi]$  to be fixed. Moreover, the parameters  $k, k_1, r$  and  $s$  verify the following system :

$$s^2(1 - k_1)^2 = ab + 4r^2k, \quad s(1 + k_1) = r(1 + k).$$

In order to close the system, we can remark that  $r$  and  $s$  are linked with the aspect ratio  $1/K$ , as well as the wave numbers  $k$  and  $k_1$  thanks to the periodicity of the  $\operatorname{cn}$ -function. We have then

$$2\pi r = 4I_J(k), \quad 2\pi s/K = 4I_J(k),$$

where  $I_J$  denotes the complete Jacobi integral. From the value of  $ab$  obtained from the fitting of the sinh relation, we are now able to resolve numerically the previous system in order to find the values of  $k, k_1, r$  and  $s$ . In Figure 4.15, we plot in the panel (a) the stream-function obtained with the (DAMM)-scheme at the final time step, and in the panel (b) the steam-function obtained by fitting  $\Psi_{fit}$ . We have set  $\varphi = 0.9$  in order to align the fitting solution with the numerical solution on the  $y$ -axis. Both plots are very similar.

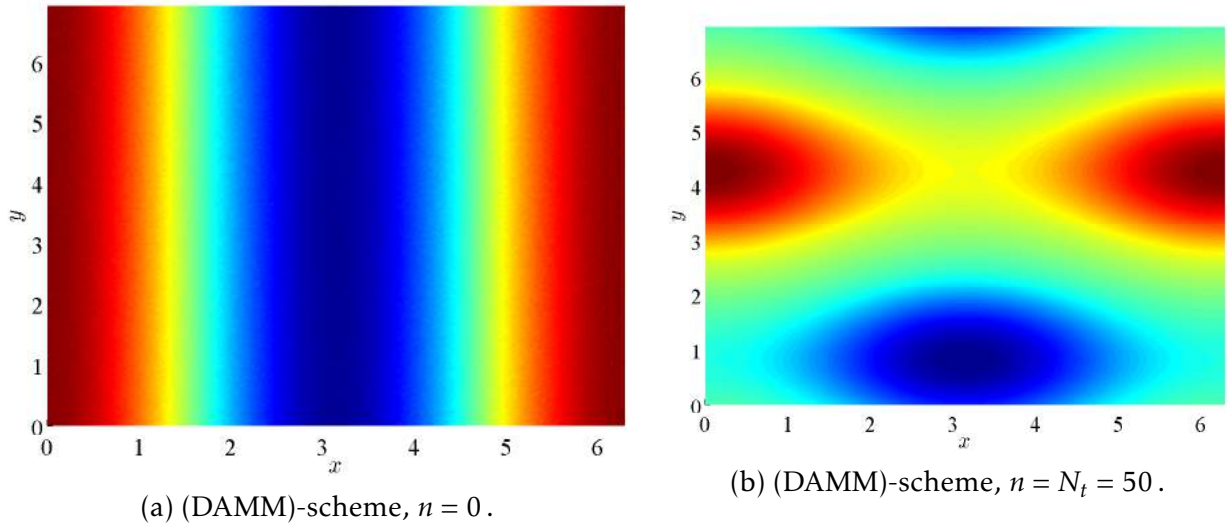


FIGURE 4.13 – (Kolmogorov case for  $\epsilon = 0$  with initial condition  $\omega_{in}^E$ ). Vorticity  $\omega^{0,\sigma}$  at the initial time (panel (a)), and at the final time  $T$  (panel (b)).  $N_x = N_y = 256$ ,  $N_t = 50$ ,  $T = 10$ ,  $\sigma = (\Delta x/L_x)^2$ , and  $1/K = 1.1$ .

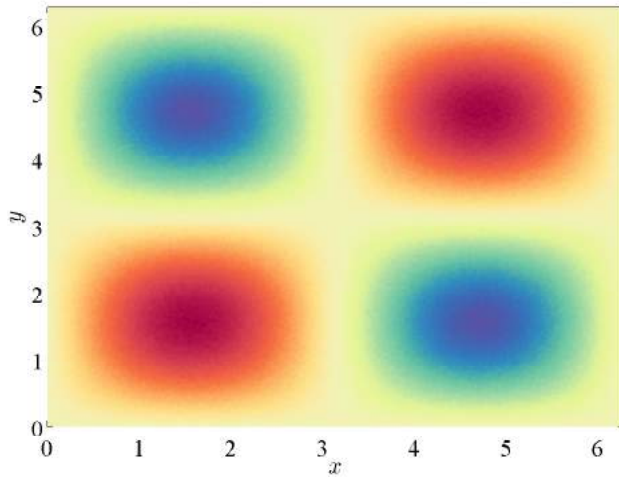
As mentioned in Section 2.1, the (DAMM)-scheme permits to eliminate the small time scales and simulate accurately the desired large scales, fact which is not possible with the (SER)-scheme. Indeed, due to the explicit nature of the (SER)-scheme, it is not feasible to let  $\epsilon$  tend towards zero in order to attain the late state directly. In this case, the CFL-condition requires a very large number of time iterations which becomes rapidly too costly. (Note that for the simulations performed in Figure (4.16), we had to choose  $N_t = 5 \times 10^6$  for  $\epsilon = 1 \times 10^{-3}$ ). If we are just interested in the viscous time scales, the AP-scheme is thus the good method to apply, although it works also for all time scales.

## 4.5 Concluding remarks and perspectives

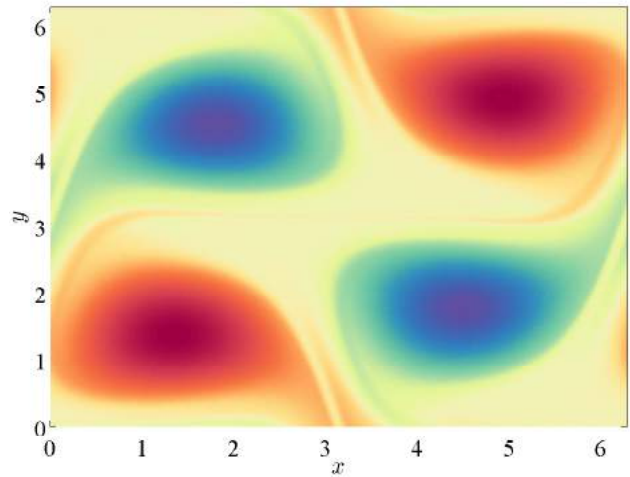
We have proposed in this paper an AP-scheme, mixing micro-macro decomposition and regularization procedure, in order to attain long-time asymptotics for two problems coming from fluid mechanics. In particular, the AP-procedure allowed to obtain very rapidly and accurately the late state of the unstable Kolmogorov flow, avoiding the intermediate states and hence the accumulation of numerical errors.

In this aim, let us mention in this concluding part some words about the difference of our AP-scheme with respect to other more standard schemes. Indeed, we performed some simulations with another scheme, an explicit spectral method, called in the following the (SER)-scheme (SER for Spectral Explicit Reference). In Figure 4.16, we plot in the Kolmogorov flow framework the stream-function obtained with the (SER)-scheme at the physical time  $t' = 5000$  for  $\epsilon = 1 \times 10^{-3}$ . Except for the alignment on the  $y$ -axis, the solutions obtained with both (DAMM) and (SER) schemes seem to coincide (Figure 4.12 (a)). In Figure 4.17, we plot the  $\omega - \Psi$  diagram as we did for the (DAMM)-scheme. The curve obtained is very similar to the one obtained with the (DAMM)-scheme (Figure 4.12 (b)).

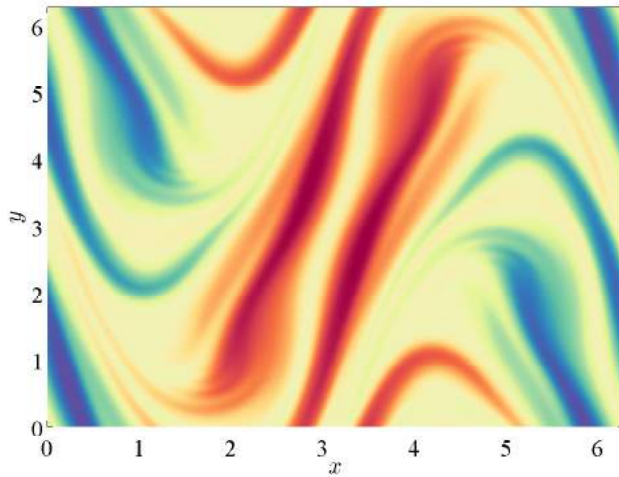
Based on the obtained numerical results, several interesting problems could now be investigated in more details in future works. For example, the functional relation (obtained numerically) between the vorticity and the stream-function need to be further studied by analytical methods. Secondly, the nonlinear as well as the saturation phases of the instability need to be validated by some analytical results as it was done for the tearing modes in MHD [49, 109, 121]. To finish, we would like to remark that our numerical approach could be applied to similar singular-perturbed problems, when one is interested in the long-time scale dynamics.



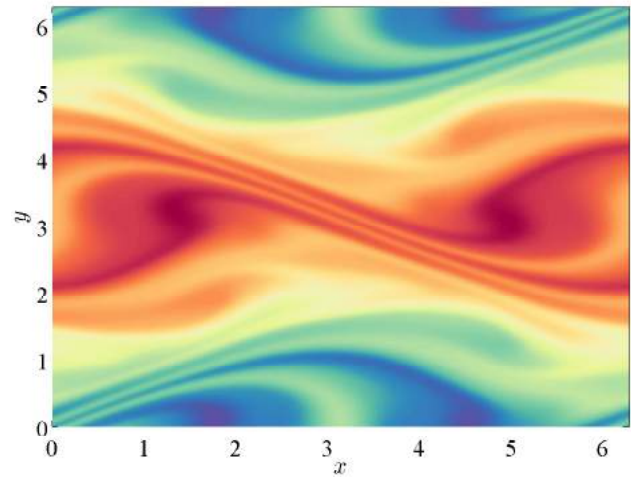
(a) (DAMM)-scheme,  $n = 0$ .



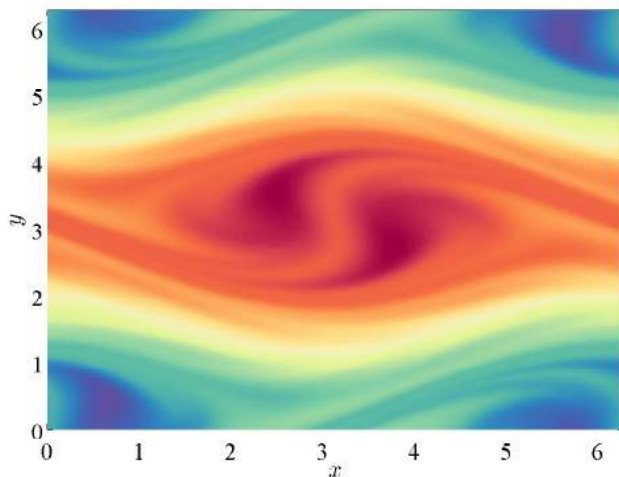
(b) (DAMM)-scheme,  $n = 1100$ .



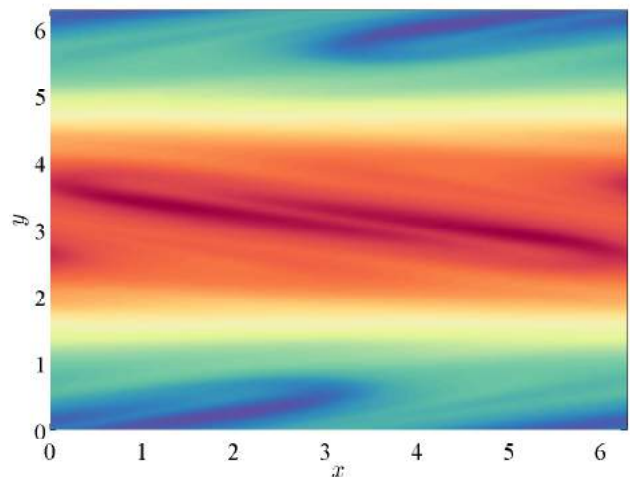
(c) (DAMM)-scheme,  $n = 1350$ .



(d) (DAMM)-scheme,  $n = 1900$ .

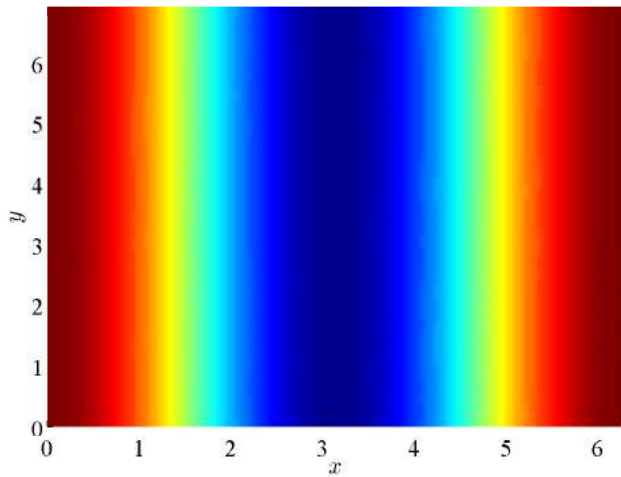


(e) (DAMM)-scheme,  $n = 2250$ .

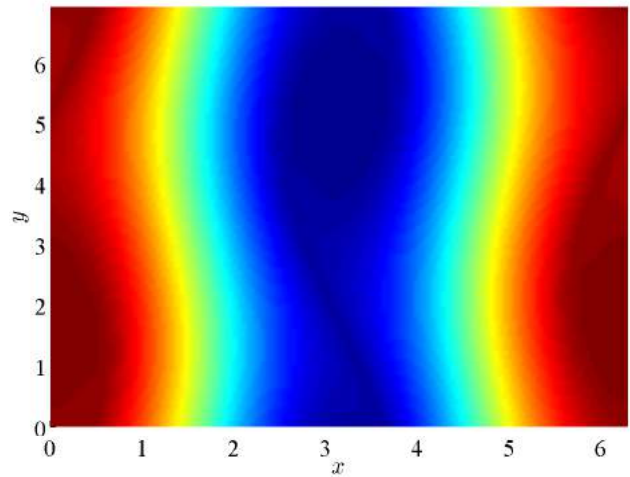


(f) (DAMM)-scheme,  $n = 2500$ .

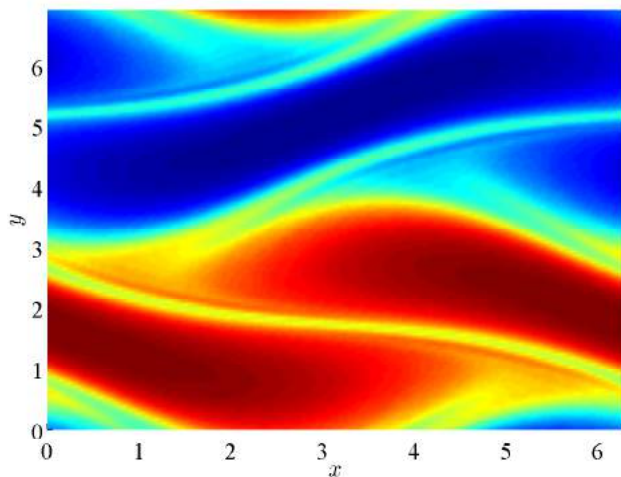
FIGURE 4.5 – (Taylor-Green vortex with initial condition  $\omega_{in}^B$ ). Vorticity field  $\omega^{\epsilon,\sigma}$  at different times. Parameters were  $\epsilon = 5e - 4$ ,  $N_x = N_y = 256$ ,  $N_t = 2500$ ,  $\mathcal{T} = 150$ , and  $\sigma = (\Delta x/L_x)^2$ .



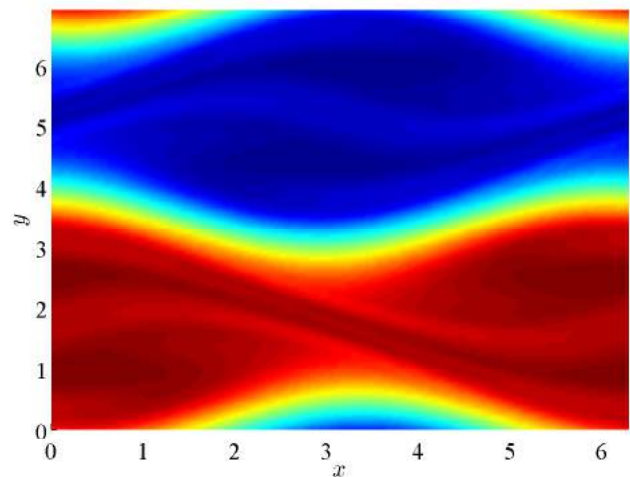
(a) (DAMM)-scheme, Initial condition .



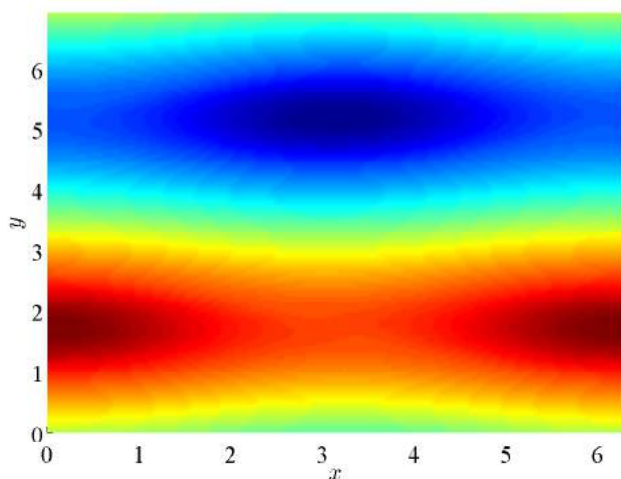
(b) (DAMM)-scheme,  $t = 50$  .



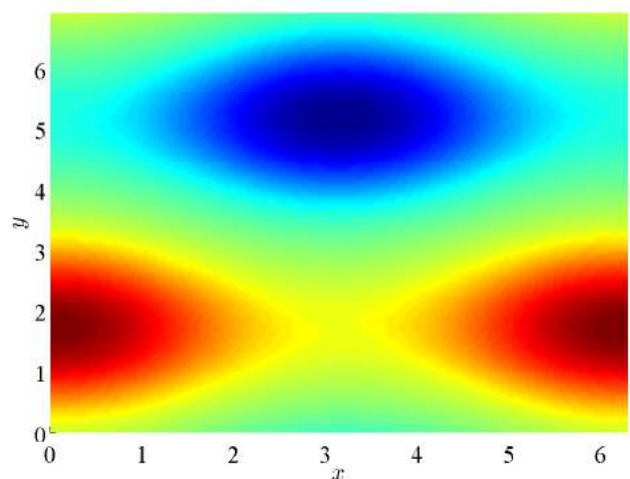
(c) (DAMM)-scheme,  $t = 75$  .



(d) (DAMM)-scheme,  $t = 150$  .

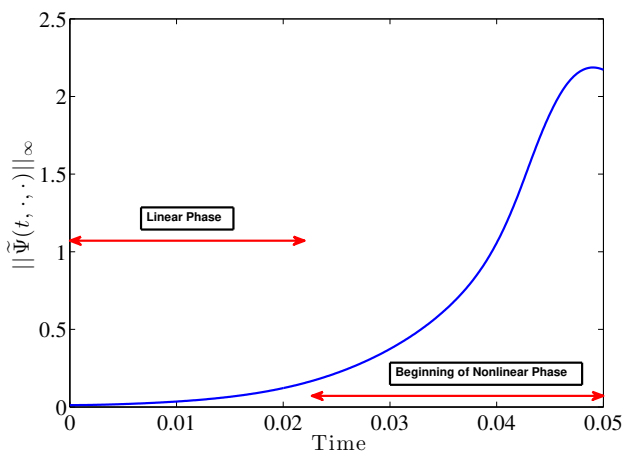


(e) (DAMM)-scheme,  $t = 1000$  .

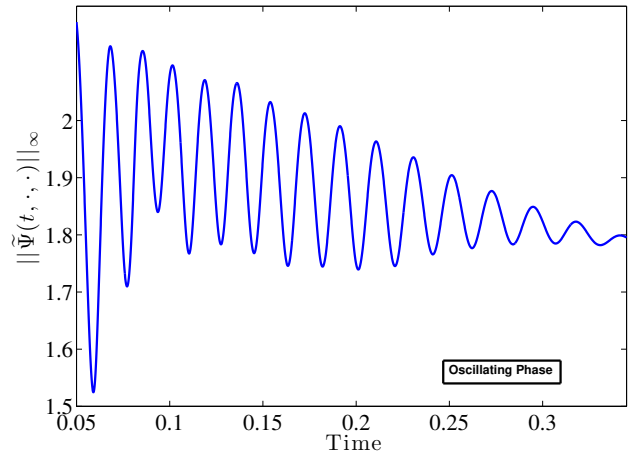


(f) (DAMM)-scheme,  $t = 5000$  .

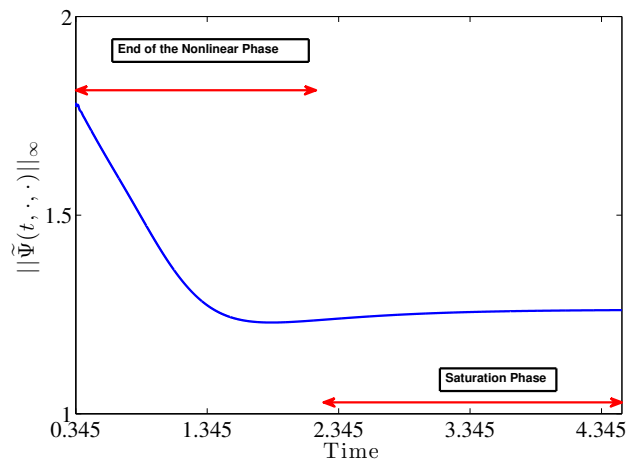
FIGURE 4.10 – (Kolmogorov flow with initial condition  $\omega_{in}^E$ ). Vorticity field  $\omega^{\epsilon,\sigma}$  at different physical times, obtained with a viscosity of  $\epsilon = 1e-3$ .  $N_x = N_y = 128$ ,  $N_t = 50000$ ,  $\mathcal{T} = 5000$ ,  $\sigma = (\Delta x/L_x)^2$ , and  $K = 1/1.1$ .



(a)  $t' \in (0, 0.05)$ .



(b)  $t' \in (0.05, 0.345)$ .



(c)  $t' \in (0.345, 4.5)$ .

FIGURE 4.11 – (Kolmogorov flow with initial condition  $\omega_{in}^E$ ). Discrete  $L^\infty$ -norm of  $\tilde{\Psi}$  versus time with  $\epsilon = 1e-3$ ,  $N_x = N_y = 64$ ,  $N_t = 45000$ ,  $\mathcal{T} = 4.5$ ,  $\sigma = (\Delta x/L_x)^2$ , and  $K = 1/1.1$ .



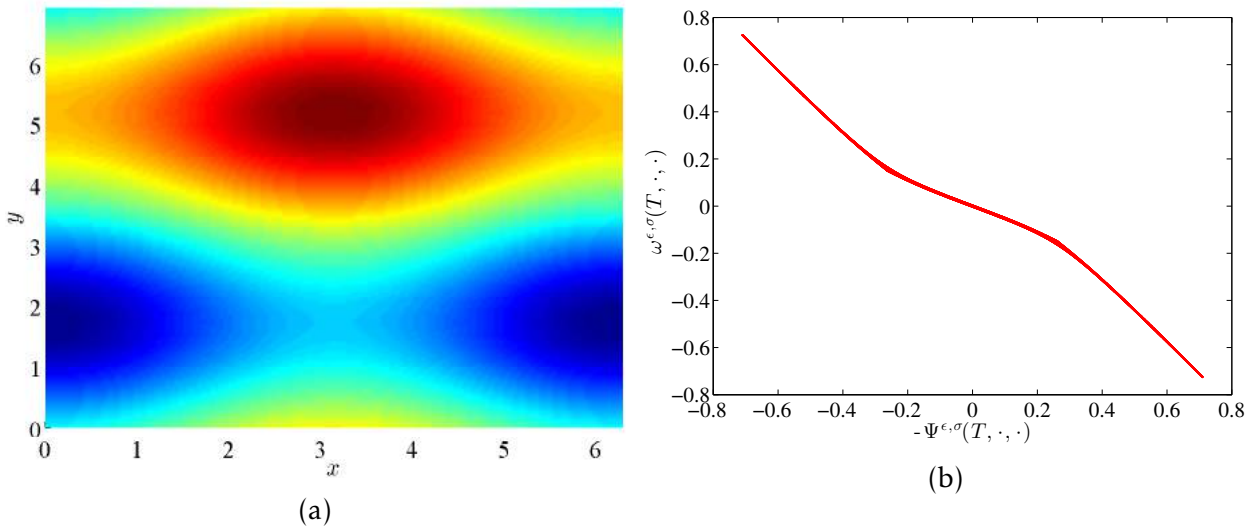


FIGURE 4.12 – (Kolmogorov case with initial condition  $\omega_{in}^E$ ). Final state of the stream-function ( $-\Psi^{\epsilon, \sigma}$ ) (panel (a)), and scattering plot between the vorticity and the stream function (panel (b)) obtained with a viscosity of  $\epsilon = 1e-3$ .  $N_x = N_y = 128$ ,  $N_t = 50000$ ,  $T = 5000$ ,  $\sigma = (\Delta x/L_x)^2$ , and  $K = 1/1.1$ .

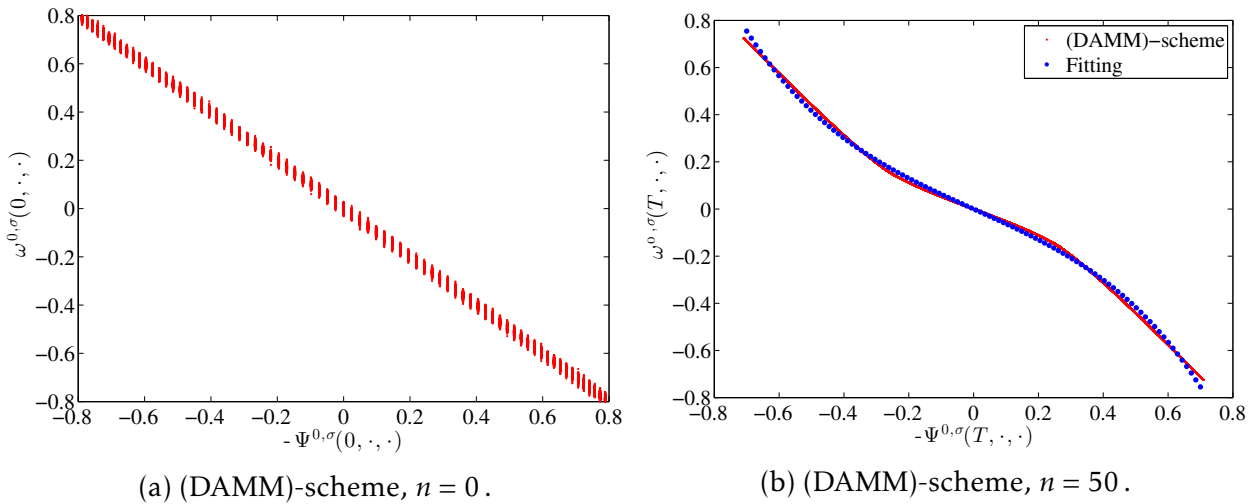


FIGURE 4.14 – (Kolmogorov case for  $\epsilon = 0$  with initial condition  $\omega_{in}^E$ ). Scattering plot between the vorticity and the stream function at the initial time (panel (a)) and at the final time (panel (b)).  $N_x = N_y = 256$ ,  $N_t = 50$ ,  $T = 10$ ,  $\sigma = (\Delta x/L_x)^2$ , and  $1/K = 1.1$ .

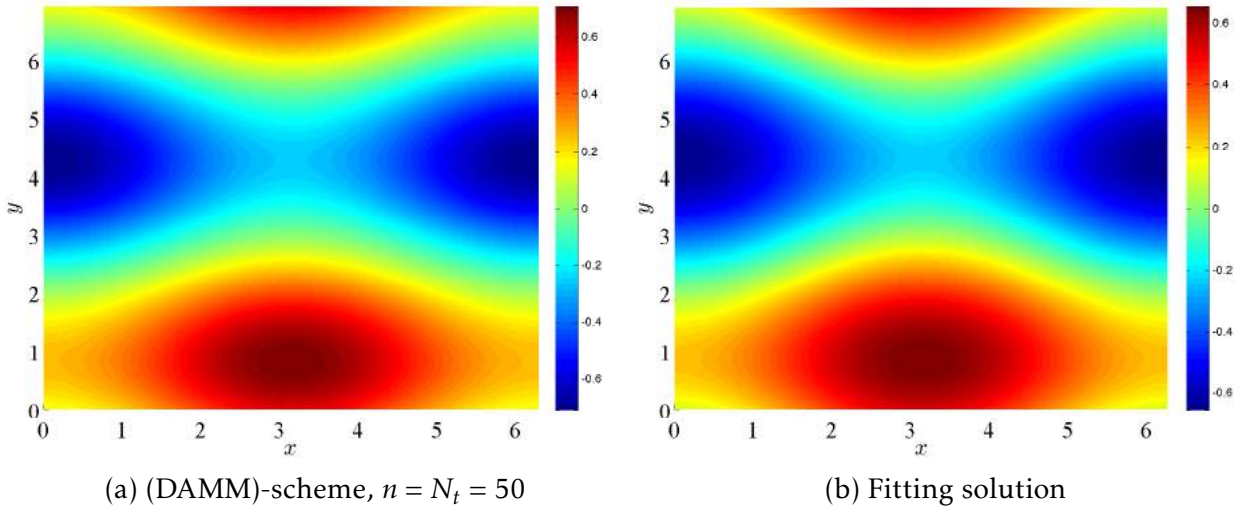


FIGURE 4.15 – (Kolmogorov case for  $\epsilon = 0$  with initial condition  $\omega_{in}^E$ ). Stream function  $(-\Psi^{0,\sigma})$  obtained with the (DAMM)-scheme (panel (a)) and with the fitting  $(-\Psi_{fit})$  (panel (b)).  $N_x = N_y = 256$ ,  $N_t = 50$ ,  $T = 10$ ,  $\sigma = (\Delta x/L_x)^2$ , and  $1/K = 1.1$ .

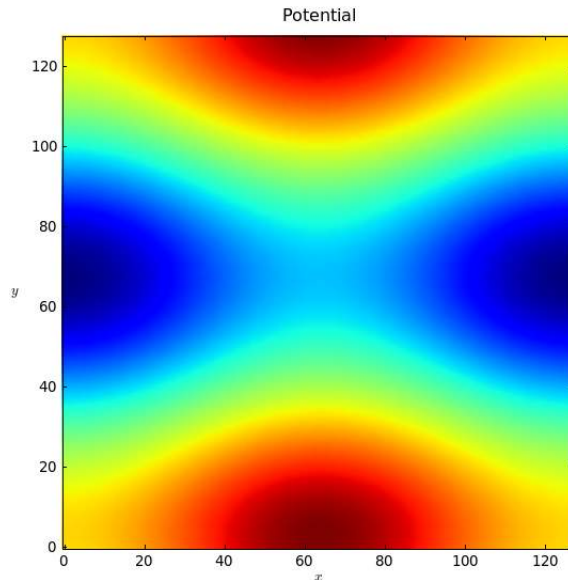


FIGURE 4.16 – (Kolmogorov test case). Stream function  $(-\Psi)$  obtained with the SER-scheme at the final time  $t' = T$ . Parameters were  $N_x = N_y = 128$ ,  $\Delta t = 0.001$ ,  $T = 5000$ ,  $\epsilon = 1 \times 10^{-3}$  and  $1/K = 1.1$ .

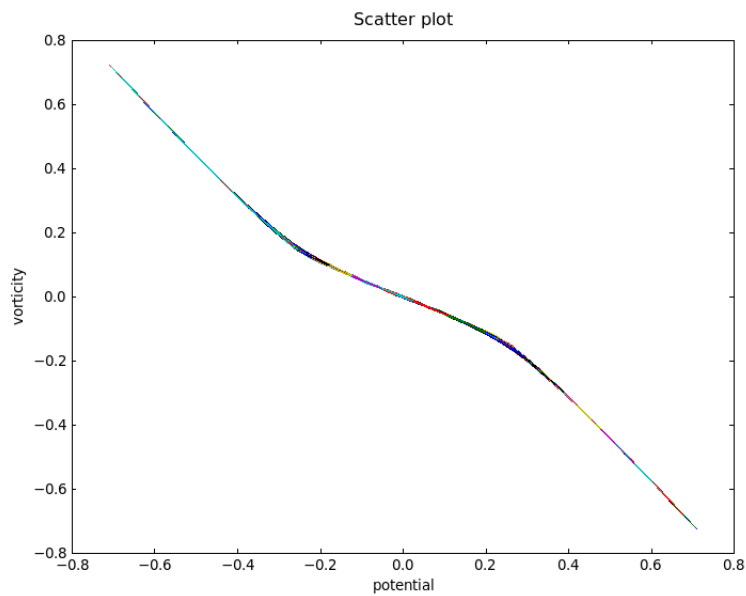


FIGURE 4.17 – (Kolmogorov test case). Scattering plot between the vorticity and the stream function obtained with the SER-scheme at the time  $t' = 10000$ . Parameters were  $N_x = N_y = 128$ ,  $\Delta t = 0.001$ ,  $\mathcal{T} = 10000$ ,  $\epsilon = 1 \times 10^{-3}$  and  $1/K = 1.1$ .

---

# BIBLIOGRAPHIE

- [1] A. Akhiezer, I. Akhiezer, R. Polovin, A. Sitenko, , and K. Stepanov. *Plasma Electrodynamics. Vol. I : Linear Theory. Vol. II : Non-Linear Theory and Fluctuations*. Pergamon Press, 1975.
- [2] R. Alexander. Diagonally implicit runge-kutta methods for stiff o.d.e.'s. *SIAM Journal on Numerical Analysis*, 6 :1006–1021, 1977.
- [3] G. Allaire. Homogenization and two-scale convergence. *SIAM J. Math. Anal.*, 23 :1482–1518, 1992.
- [4] G. Allaire. Two-scale convergence. *Esaim Proceedings*, 38 :1–35, 2012.
- [5] D. Anderson, R. Fedele, and M. Lisak. A tutorial presentation of the two-stream instability and landau damping. *American Journal of Physics*, 69 :1262–1266, 2001.
- [6] A. Arakawa. Computational design for long-term numerical integration of the equations of fluid motion : two dimensional incompressible flow. *Journal of Computational Physics*, 135 :119–143, 1966.
- [7] V. I. Arnol'd. Kolmogorov's hydrodynamic attractors. *Proc. Royal Soc. London A*, pages 19–22, 1991.
- [8] V. I. Arnol'd and L.D. Meshalkin. Seminar led by an kolmogorov on selected problems of analysis. *Usp. Mat. Nauk*, 15 :20–24, 1960.
- [9] A. A. Arsen'ev. Existence in the large of a weak solution of vlasov's system of equations. *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki*, 15 :136–147, 1975.
- [10] H. Ashley and M. Landahl. *Aerodynamics of Wings and Bodies*. Dover Publications, 1985.
- [11] L. Belenkaya, S. Friedlander, and V. Yudovich. The unstable spectrum of oscillating shear flows. *SIAM J. Appl. Math.*, 59 :1701–1715, 1965.
- [12] P. M. Bellan. *Fundamentals In Plasma Physics*. Cambridge University Press, 2006.
- [13] A. Bensoussan, J. L. Lions, and G. Papanicolaou. Asymptotic analysis for periodic structures. *Studies in Mathematics and its Applications*, 1958.
- [14] I. B. Bernstein, J. M. Greene, and M. D. Kruskal. Exact nonlinear plasma oscillations. *Physical Review*, 108 :546–550, 1957.
- [15] P. Bertrand, F. Filbet, and E. Sonnendrücker. Conservative numerical schemes for the vlasov equation. *J. Comput. Phys.*, 172 :166–187, 2001.

- [16] P.B Bochev and R. B Lehoucq. Regularization and stabilization of discrete saddle-point variational problems. *Electronic Transactions on Numerical Analysis*, 22 :97–113, 2006.
- [17] W. J. Bos, S. Futatani, S. Benkadda, M. Farge, and K. Schneider. The role of coherent vorticity in turbulent transport in resistive drift-wave turbulence. *Physics of Plasmas*, 15, 2011.
- [18] S. Boscarino, F. Filbet, and G. Russo. High order semi-implicit schemes for time dependent partial differential equations. *Electronic Transactions on Numerical Analysis*, 68 :975–1001, 2016.
- [19] M. Bostan. Transport equations with disparate advection fields. application to the gyrokinetic models in plasma physics. *SIAM J. Sci. Comp.*, 31 :334–368, 2008.
- [20] M. Bostan. The vlasov-poisson system with strong external magnetic field. finite larmor radius regime. *Asymptot. Anal.*, 61 :91–123, 2009.
- [21] T. J. M. Boyd and J. J. Sanderson. *The Physics Of Plasmas*. Cambridge University Press, 2003.
- [22] C. Brezinski, M. Redivo-Zaglia, G. Rodriguez, and S. Seatzu. Multi-parameter regularization techniques for ill-conditioned linear systems. *Numer. Math.*, 94 :203–228, 2003.
- [23] C. Le Bris. *Systèmes Multi-échelles. Modélisation et Simulation*. Springer-Verlag, 2005.
- [24] D. Calvetti, S. Morigi, L. Reichel, and F. Sgallari. Tikhonov regularization and the l-curve for large discrete ill-posed problems. *J. Comput. Appl. Math.*, 123 :423–446, 2000.
- [25] A. De Cecco, C. Negulescu, and S. Possanner. Asymptotic transition from kinetic to adiabatic electrons along magnetic field lines. *SIAM Multiscale Modeling and Simulation*, 15 :309–338, 2017.
- [26] F. F. Chen. *Plasma Physics and Controlled Fusion*, volume 1. Springer-Verlag, 2006.
- [27] Z. M. Chen and W. G. Price. Stability and instability analyses of the dissipative quasi-geostrophic equation. *Nonlinearity*, 21 :765–782, 2008.
- [28] Y. Cheng, I.M. Gamba, and P.J Morrison. Study of conservation and recurrence of runge-kutta discontinuous galerkin schemes for vlasov-poisson systems. *Journal of Scientific Computing*, 56 :319–349, 2013.
- [29] P. C. Clemmow and J. P. Dougherty. *Electrodynamics Of Particles And Plasmas*. Addison-Wesley Publishing Company, 1969.
- [30] A. Crestetto, N. Crouseilles, and M. Lemou. Asymptotic-preserving scheme based on a finite volume/particle-in-cell coupling for boltzmann-bgk-like equations in the diffusion scaling. *FCVA VII - Elliptic, Parabolic and Hyperbolic Problems, Springer Proceedings in Mathematics and Statistics*, 78 :827–835, 2014.
- [31] N. Crouseilles and M. Lemou. An asymptotic preserving scheme based on a micro-macro decomposition for collisional vlasov equations : diffusion and high-field scaling limits. *Kinet. Relat. Models*, 4 :441–477, 2011.
- [32] N. Crouseilles, M. Lemou, and F. Méhats. Asymptotic-preserving schemes for oscillatory vlasov-poisson equations. *Journal of Computational Physics*, 248 :287–308, 2013.
- [33] N. Crouseilles, M. Mehrenberger, and F. Vecil. Discontinuous galerkin semi-lagrangian method for vlasov-poisson. *ESAIM : Proceedings*, 32 :211–230, 2011.

- [34] P. A. Davidson. *Turbulence, An Introduction For Scientists and Engineers*. Oxford University Presse, Oxford, 2004.
- [35] P. Degond, F. Deluzet, A. Lozinski, J. Narski, and C. Negulescu. Duality based asymptotic-preserving method for highly anisotropic diffusion equations. *Communications in Mathematical Sciences*, 10 :1–31, 2012.
- [36] P. Degond, A. Lozinski, J. Narski, and C. Negulescu. An asymptotic-preserving method for highly anisotropic elliptic equations based on a micro-macro decomposition. *Journal of Computational Physics*, 231 :2724–2740, 2012.
- [37] P. Degond and M. Tang. All speed scheme for the low mach number limit of the isentropic euler equation. *Communications in Computational Physics*, 10 :1–31, 2011.
- [38] J. L. Delcroix and A. Bers. *Physique Des Plasmas*. CNRS Editions, 1994.
- [39] F. Deluzet, C. Negulescu, M. Ottaviani, and S. Possanner. Numerical study of the plasma tearing instability on the resistive time scale. *J. Comp. Phys.*, 280 :602–635, 2015.
- [40] L. Demeio and J. P. Holloway. Numerical simulations of bgk modes. *Journal of Plasma Physics*, 46 :63–84, 2009.
- [41] L. Demeio and P. F. Zweifel. Numerical simulations of perturbed vlasov equilibria. *Phys. Fluids B*, 2 :1252–1255, 1990.
- [42] G. Dimarco, R. Loubère, and M-H. Vignal. Study of a new asymptotic preserving scheme for the euler system in the low mach number limit. *SIAM J. Sci. Comput.*, 39 :2099–2128, 2016.
- [43] R. J. DiPerna and P.-L. Lions. Global weak solutions of vlasov-maxwell systems. *Comm. Pure Appl. Math.*, 12 :655–658, 1988.
- [44] R. J. DiPerna and P.-L. Lions. Solutions globales d'équations du type vlasov-poisson. *C. R. Acad. Sci. Paris Sér. I Math*, 6 :729–757, 1989.
- [45] R. J. DiPerna and P.-L. Lions. Global weak solutions of kinetic equations. *Sem. Mat. Univ. Politec. Torino*, 3 :259–288, 1990.
- [46] P. G. Drazin and W. H. Reid. *Hydrodynamic Stability*. Cambridge University Presse, 2nd Edition, 2010.
- [47] M. Van Dyke. *Perturbation Methods In Fluid Mechanics*. Academic Press INC, 1964.
- [48] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Academic Publishers, 1996.
- [49] D. Escande and M. Ottaviani. Simple and rigorous solution for the nonlinear tearing mode. *Phys. Letters*, 323 :278–284, 2004.
- [50] D.F. Escande, D. Bénisti, Y. Elskens, D. Zarzoso, and F. Doveil. Basic microscopic plasma physics from n-body mechanics. *submitted*.
- [51] B. Fedele and C. Negulescu. Numerical study of an anisotropic vlasov equation arising in plasma physics. *Kinetic and Related Models*, 11 :1395–1426, 2018.
- [52] B. Fedele, C. Negulescu, and M. Ottaviani. Numerical and analytical study of a forced vorticity equation and its asymptotic-preserving resolution. *in preparation*.
- [53] B. Fedele, C. Negulescu, and S. Possanner. Asymptotic-preserving scheme for the resolution of evolution equations stiff transport terms. *SIAM Multiscale Model. Simul.*, 17 :307–343, 2019.
- [54] F. Filbet. *Contribution à l'analyse et la simulation numérique de l'équation de Vlasov*. PhD thesis, Université Henri Poincaré (Nancy), 2001.

- [55] F. Filbet and S. Jin. A class of asymptotic-preserving schemes for kinetic equations and related problems with stiff sources. *Journal of Computational Physics*, 229 :7625–7648, 2010.
- [56] F. Filbet and S. Jin. An asymptotic preserving scheme for the es-bgk model of the boltzmann equation. *J. Sci. Computing*, 46 :204–224, 2011.
- [57] F. Filbet and L. Pareschi. A numerical method for the accurate solution of the fokker-planck equation in the non-homogenous case. *Journal of Computational Physics*, 179 :1–26, 2002.
- [58] J. P. Freidberg. *Plasma Physics And Fusion Energy*. Cambridge University Press, 2008.
- [59] E. Frénod and E. Sonnendrücker. Homogenization of the vlasov equation and the vlasov-poisson system with a strong external magnetical field. *Asymp. Anal.*, 18 :193–214, 1998.
- [60] E. Frénod and E. Sonnendrücker. Long time behavior of the two dimensionnal vlasov equation with a strong external magnetic field. *Math. Models Methods Appl. Sci.*, 10 :539–553, 2000.
- [61] H. Furth, J. Killeen, and M. N. Rosenbluth. Finite-resistivity instabilities of a sheet pinch. *Phys. Fluids*, 6 :459–484, 1963.
- [62] X. Garbet, Y. Idomura, L. Villard, and T. H. Watanabe. Gyrokinetic simulations of turbulent transport. *Nuclear Fusion*, 50 :043002, 2010.
- [63] P. Ghendrih, M. Hauray, and A. Nouri. Derivation of a gyrokinetic model. existence and uniqueness of specific stationary solutions. *Kinetic and Related Models*, 2 :707–725, 2010.
- [64] R. J. Goldston and P. H. Rutherford. *Plasma Physics*. Taylor & Francis Group, 1995.
- [65] F. Golse and L. Saint-Raymond. The vlasov-poisson system with strong magnetic field. *J. Math. Pures Appl.*, 78 :791–817, 2001.
- [66] V. Grandgirard, M. Brunetti, P. Bertrand, N. Besse, X. Garbet, P. Ghendrih, G. Manfredi, Y. Sarazin, O. Sauter, E. Sonnendrücker, J. Vaclavik, and L. Villard. Gyrokinetic simulations of turbulent transport. *Journal of Computational Physics*, 217 :395–423, 2006.
- [67] V. Grandgirard, Y. Sarazin, X. Garbet, G. Dif-Pradalier, Ph. Ghendrih, N. Crouseilles, G. Latu, E. Sonnendrücker, N. Besse, and P. Bertrand. Gysela, a full-f global gyrokinetic semi-lagrangian code for itg turbulence simulations. *Theory of Fusion Plasmas*, 871 :100–111, 2006.
- [68] A. Gupta, R. Ganesh, and Ashwin Joy. Kolmogorov flow in two dimensional strongly coupled dusty plasma. *Physics of plasmas*, 21, 2014.
- [69] D. Gurarie and K.W. Chow. Vortex arrays for sinh-poisson equation of two-dimensional fluids : equilibria and stability. *Physics of fluids*, 16 :3296–3305, 2004.
- [70] W. Hackbusch. *Multigrid methods and applications*. Springer-Verlag, 1985.
- [71] T. S. Hahm. Nonlinear gyrokinetic equations for turbulence in core transport barriers. *Physics of Plasmas*, 3 :4658–4664, 1996.
- [72] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II : Stiff And Differential-Algebraic Problems, Second Edition*. Springer-Verlag, 1996.
- [73] R.D. Hazeltine and J.D. Meiss. *Plasma Confinement*. Dover Publications, Inc. Mineola, 2003.

- [74] R.E. Heath, I.M Gamba, P.J. Morrison, and C. Michler. A discontinuous galerkin method for the vlasov-poisson system. *Journal of Computational Physics*, 231 :1140–1174, 2012.
- [75] M. H. Holmes. *Introduction to Numerical Methods in Differential Equations*. Springer-Verlag, 2007.
- [76] M. H. Holmes. *Introduction To Perturbation Methods*. Springer-Verlag, 2012.
- [77] Y. W. Hou, M. X. Chen, M. Y. Yu, and B. Wu. Linear and nonlinear behaviour of two-stream instabilities in collisionless plasmas. *J. Plas. Phys.*, 81 :1–10, 2015.
- [78] V. I. Yudovich. Example of the generation of a secondary stationary or periodic flow when there is loss of stability of the laminar flow of a viscous incompressible fluid. *Journal of Applied Mathematics and Mechanics*, 29 :527–544, 1965.
- [79] P. E. Jabin. The vlasov-poisson system with infinite mass and energy. *J. Statis. Phys.*, 103 :1107–1123, 2001.
- [80] S. Jin. Asymptotic-preserving (ap) schemes for multiscale kinetic and hyperbolic equations : a review. *Rivista di Matematica della Universita di Parma*, 3 :177–216, 2012.
- [81] S. Jin and L. Wang. An asymptotic preserving scheme for the vlasov-poisson-fokker-planck system in the high field regime. *Acta Math. Sci. Ser. B Engl.*, 31 :2219–2232, 2011.
- [82] S. Kaplun and A. P. Lagerstrom. Asymptotic expansions of navier-stokes solutions for small reynolds numbers. *J. Math. Mech.*, pages 585–593, 1967.
- [83] M. Kikuchi and M. Azumi. *Frontiers In Fusion Research II*. Springer-Verlag, 2015.
- [84] A. Klar. An asymptotic-induced scheme for non-stationary transport equations in the diffusive limit. *SIAM Journal of Numerical Analysis*, 35 :1073–1094, 1998.
- [85] A. J. Klimas. Numerical method based on the fourier-fourier transform approach for modeling 1-d electron plasma evolution. *J. Comput. Phys.*, 50 :270–306, 1983.
- [86] G. Knorr. Plasma simulation with few particles. *J. Comput. Phys.*, 13 :165–180, 1973.
- [87] P. A. Lagerstrom. *Matched Asymptotic Expansions, Ideas And Techniques*. Springer-Verlag, 1988.
- [88] P. A. Lagerstrom, J. D. Cole, and L. Trilling. Problems in the theory of viscous compressible fluids. *unpublished*, 1949.
- [89] L. Landau. On the vibration of the electronic plasma. *English translation in J. Phys. (USSR)*, 10, 1946.
- [90] I. Langmuir. Oscillations in ionized gases. *Proc. Nat. Acad. Sci. U.S.*, 14, 1928.
- [91] E. W. Larsen, J. E. Morel, and W. F. Miller. Asymptotic solutions of numerical transport problems in optically thick, diffusive regime. *J. Differential Equations*, 69 :287–324, 1987.
- [92] E. W. Larsen, J. E. Morel, and W. F. Miller. The convergence of numerical transfer schemes in diffusive regimes. i. discrete-ordinate method. *SIAM, J. Numer. Ana.*, 36 :1333–1369, 1999.
- [93] M. Lemou and L. Mieussens. A new asymptotic preserving scheme based on micro-macro formulation for linear kinetic equations in the diffusion limit. *SIAM J. Sci. Comput.*, 31 :334–368, 2008.
- [94] M. Lesieur. *Turbulence in Fluids*. Springer-Verlag, 2008.



- [95] R. J. LeVeque. *Finite Difference Methods for Ordinary and Partial Differential Equations*. Siam, 2007.
- [96] R. G. Littlejohn. A guiding center hamiltonian : A new approach. *Journal of Mathematical Physics*, 20 :2445–2458, 1979.
- [97] B. Louvet. Ces moteurs à plasma qui pourraient un jour nous emmener dans l’espace. <https://sciencepost.fr/2017/05/moteurs-a-plasma-pourraient-jour-emmener-lespace/>.
- [98] A. Lozinski, J. Narski, and C. Negulescu. Highly anisotropic temperature balance equation and its asymptotic-preserving resolution. *M2AN (Mathematical Modelling and Numerical Analysis)*, 48 :1701–1724, 2014.
- [99] D. Lucas and R. Kerswell. Spatiotemporal dynamics in two-dimensional kolmogorov flow over large domains. *Journal of Fluid Mechanics*, 750 :518–554, 2014.
- [100] A. Lucque and H. Schamel. Electrostatic trapping as a key to the dynamics of plasmas, fluids and other collective systems. *Physics Reports*, 415 :261–359, 2006.
- [101] H. J. Lugt. *Vortex Flow in Nature and Technology*. John Wiley & Sons, 1983.
- [102] A. J. Majda and A. L Bertozzi. *Vorticity and Incompressible Flow*. Cambridge University Press, 2002.
- [103] G. Manfredi. Long-time behavior of nonlinear landau damping. *Phys. Rev. Lett.*, 79 :2815–2818, 1997.
- [104] G. Manfredi and P. Bertrand. Stability of bernstein-greene-kruskal modes. *Journal of Plasma Physics*, 7 :2425–2431, 2000.
- [105] C. Marchioro. An example of absence of turbulence for any reynolds number. *Commun. Math. Phys.*, 105 :99–106, 1986.
- [106] C.J. McKinstrie, R. E. Giacone, and E. A. Startsev. On the vibration of the electronic plasma. *English translation in J. Phys. (USSR)*, 10, 1946.
- [107] A. Mentrelli and C. Negulescu. Asymptotic-preserving scheme for highly anisotropic non-linear diffusion equations. *Journal of Comp. Phys*, 231 :8229–8245, 2012.
- [108] L. D. Meshalkin and Ia. G. Sinai. Investigation of the stability of a stationary solution of a system of equations for the plane movement of an incompressible viscous liquid. *PMM*, 6 :1140–1143, 1961.
- [109] F. Militello and F. Porcelli. Simple analysis of the nonlinear saturation of the tearing mode. *Phys. Plasmas*, 11 :L13–L16, 2004.
- [110] D. Montgomery and T. Armstrong. Asymptotic state of the two-stream instability. *J. Plasma. Physics*, 1 :425–433, 1967.
- [111] L. Navoret. *Méthodes asymptotico-numériques pour des problèmes issus de la physique des plasmas et de la modélisation des interactions sociales*. PhD thesis, Université Paul-Sabatier (Toulouse), 2010.
- [112] C. Negulescu. Asymptotic-preserving schemes. modelling, simulation and mathematical analysis of magnetically plasmas. *Rivista di Matematica della Università di Parma*, 4 :265–343, 2013.
- [113] C. Negulescu. Kinetic modelling of strongly magnetized tokamak plasmas with mass disparate particles. the electron boltzmann relation. *SIAM MMS (Multiscale Model. Simul.)*, 16 :1732–1755, 2018.
- [114] G. Nguetseng. A general convergence result for a functional related to the theory of homogenization. *SIAM J. Math. Anal.*, 20 :608–623, 1989.

- [115] H. Okamoto. A study of bifurcation of kolmogorov flows with an emphasis on singular limit. *Documenta Mathematica*, Extra Volume ICM :513–522, 1998.
- [116] Y. Omura, H. Matsumoto, T. Miyake, and H. Kojima. Electron beam instabilities as generation mechanism of electrostatic solitary waves in the magnetotail. *Journal of Geophysical Research*, 101 :2685–2697, 1996.
- [117] T. O’Neil. Collisionless damping of nonlinear plasma oscillations. *Phys. Fluids*, 8 :2255–2262, 1965.
- [118] L. Prandtl. Fluessigkeiten bei sehr kleimer reibung. *International Math. Kongress*, pages 484–491, 1905.
- [119] J. M. Rax. *Physique Des Tokamaks*. Les Éditions de l’École Polytechnique, 2011.
- [120] M. N. Rosenbluth, W. M. MacDonald, and D. I. Judd. Fokker-planck equation for an inverse square force. *Physics Review*, 107, 1957.
- [121] P. Rutherford. Nonlinear growth of the tearing mode. *Phys. Fluids*, 16 :1903–1908, 1973.
- [122] T.K Sengupta, N. Sharma, and A. Sengupta. Non-linear instability analysis of the two-dimensional navier-stokes equation : the taylor-green vortex problem. *Physics of Fluids*, 30, 2018.
- [123] M. Shoucri and G. Knorr. Numerical integration of the vlasov equation. *J. Comput. Phys.*, 14 :84–92, 1974.
- [124] P. Le Tallec. Domain decomposition methods in computational mechanics. *Computational mechanics advances*, pages 121–220, 1994.
- [125] L. Tartar. Cours peccot. *Collège de France*, 1977.
- [126] L. Trefethen and D. Bau. *Numerical Linear Algebra*. Siam, 1997.
- [127] C. Villani. A review of mathematical topics in collisional kinetic theory. *Handbook of mathematical fluid dynamics, Vol. I.*, pages 71–305, 2002.
- [128] C. Villani and C. Mouhot. On landau damping. *Acta Math.*, 207 :29–201, 2011.
- [129] E. Weinan. *Principles of Multiscale Modeling*. Cambridge University Press, 2011.
- [130] S. Wollman and E. Ozizmir. Numerical approximation of the one-dimensional vlasov-poisson with periodic boundary conditions. *SIAM, J. Num. Anal.*, 33 :1377–1409, 1996.
- [131] E. Zeidler. *Nonlinear Functional Analysis and its Applications, II-A*. Springer-Verlag, 2011.
- [132] D. Zwillinger. *Standard Mathematical Tables and Formulae*. CRC Press, 2012.

# Étude mathématique et numérique d'équations cinétiques et fluides multi-échelles pour la description d'un plasma de fusion

## RÉSUMÉ

Cette thèse de doctorat a pour thématique la modélisation mathématique et la simulation numérique de plusieurs équations d'évolution anisotropes qui modélisent des phénomènes issus de la physique des plasmas et de la mécanique des fluides. Les plasmas de fusion thermonucléaires sont un milieu très instable et anisotrope, d'où l'apparition de plusieurs problèmes mathématiques intéressants et complexes.

La première partie porte sur des modèles jouets issus de l'équation de Vlasov anisotrope. L'objectif étant de développer des schémas numériques (en particulier des schémas préservant l'asymptotique) qui résolvent ces modèles de manière efficace en vue de les implémenter ultérieurement sur des modèles plus physiques et plus complexes. En particulier, ce travail a permis de dégager les avantages et les inconvénients de nos schémas numériques en fonction de la nature du problème considéré.

La seconde partie est dédiée à l'étude de modèles plus complexes, notamment le système de Vlasov-Poisson. D'un point de vue numérique, un seul schéma préservant l'asymptotique, basé sur une décomposition Micro-Macro couplé avec une méthode de régularisation est développé. Grâce à ce schéma, il sera possible d'atteindre les états d'équilibres BGK du système de Vlasov-Poisson en quelques itérations temporelles, en évitant ainsi une importante accumulation d'erreurs numériques.

La dernière partie s'attache à étudier un système de Vorticité-Poisson, issu de la mécanique des fluides. En particulier, deux écoulements caractéristiques de ce système seront étudiés : les écoulements de Taylor-Green et de Kolmogorov. Le premier permettra principalement de valider notre procédure numérique, qui est similaire à la procédure déjà évoquée dans la partie précédente. En revanche, nous étudions plus en détail l'écoulement de Kolmogorov qui peut conduire à une instabilité sous certaines conditions. Un résultat analytique est donné pour la phase linéaire de cette instabilité, reliant le taux d'instabilité et la rapport d'aspect du domaine. Les phases non-linéaire et de saturation sont ensuite étudiées numériquement. En particulier, les propriétés AP de notre schéma permettront d'atteindre en quelques itérations en temps un nouvel équilibre issu de l'instabilité.

## ABSTRACT

This thesis deals with the mathematical modeling and the numerical simulation of several evolution equations with a stiff term which model phenomena coming from plasma physics and fluid mechanics. Thermonuclear plasmas are a highly unstable and anisotropic medium, where the occurrence of several interesting mathematical problems.

The first part concerns toy-models obtained from the anisotropic Vlasov equation. The aim is to develop numerical methods (in particular asymptotic-preserving schemes) which resolve efficiently these problems, in the aim to pass then to more complex physical models. In particular, this work has permitted to highlight both advantages and drawbacks of the developed numerical schemes depending on the considered problem.

The second part is dedicated to the study of more complex models, in particular to the Vlasov-Poisson system. From a numerical point of view, an AP scheme is developed, based on a Micro-Macro decomposition coupled with a regularization technique. Thanks to this scheme, it is possible to attain the BGK waves, solution of the Vlasov-Poisson equation, in few time iterations, avoiding thus an important accumulation of numerical errors.

The last part focus on the study of a Vorticity-Poisson system, coming from fluid mechanics. In particular, two characteristics flows of this system are investigated : the so-called Taylor-Green and Kolmogorov flows. The first one permits mainly to validate our numerical procedure, similar to that evocated in the last part. However, the Kolmogorov flow is more deeply studied. It can lead to an unstable flow under certain conditions. An analytical result for the linear instability is given, linking the growth rate of the instability phase with the aspect ratio of the domain. Then, both non-linear and saturation phases are numerically investigated. In particular, the special AP-properties of our scheme permit us to attain in only few iterations a new equilibrium of the instability.