Wavelet-based study of dissipation in plasma and fluid flows

Romain Nguyen van Yen

N° d'ordre: 10061

THÈSE

Présentée pour obtenir

LE GRADE DE DOCTEUR EN SCIENCES DE
L'UNIVERSITÉ PARIS-SUD XI

Spécialité: Physique

par

Romain Nguyen van yen

Wavelet-based study of dissipation in plasma and fluid flows

Soutenue le 8 décembre 2010 devant la Commission d'examen:

Pr.	Claude BARDOS	(Président du jury)
Pr.	Olivier TALAGRAND	(Examinateur)
Pr.	Jean-Marcel RAX	(Examinateur)
Pr.	Rupert KLEIN	(Rapporteur)
Pr.	Marie FARGE	(Directrice)
Pr.	Kai SCHNEIDER	(Co-directeur)

Rapporteurs:

Pr.	Rupert Klein
Pr.	Yanick Sarazin

Thèse préparée à
l'École Normale Supérieure de Paris
Laboratoire de Météorologie Dynamique (UMR 8539)
24 rue Lhomond
75 231 Paris CEDEX 5

Abstract

The topic of dissipation by macroscopic flows is approached by considering two of its most representative occurences, namely dissipation by plasma flows in the vanishing collisionality limit and dissipation by fluid flows in the vanishing viscosity limit. It is argued that dissipation can arise either due to the residual effect of a microscopic coupling parameter, or due to purely macroscopic nonlinear mixing effects. The combination of these two phenomena puts the problem out of reach of the most successful statistical methods that have been developed in the context of conservative systems, and also raises fundamental mathematical questions. Moreover, explicit computations that would resolve all scales of such flow are still unfeasible in this context, because of the present limitations in memory size and number of operations. It is thus widely recognized that new ways have to be found to make progress.

For that purpose, we explore the applicability of a multiscale wavelet framework. First, the partial differential equations which describe the flow must be recast into a discrete wavelet representation, while preserving consistency with the dissipative mechanisms we have outlined. This step, which we call regularization, is the subject of two chapters in this thesis, concerning the special cases of the one-dimensional Vlasov-Poisson equations on the one hand, and of the two-dimensional incompressible Euler equations on the other hand. The possibilities to develop these schemes for the practical simulation of flows is assessed, and they are compared with other existing regularizations mechanisms.

To proceed further, the origin of the residual dissipation must be tracked down and linked to mathematical properties of the solutions. We obtain some elements in this direction by studying the collision of a vorticity dipole with a wall in the vanishing viscosity limit. If the solutions are known behave well mathematically, one can readily move to the next step, which is the definition of macroscopic dissipation in the wavelet representation. This is the case for two-dimensional homogeneous turbulent flows which we subsequently address. Finally, as a perspective for future work, we perform a preliminary wavelet analysis of a three-dimensional turbulent boundary layer flow.

Keywords : wavelets, turbulence, Navier-Stokes, Vlasov, boundary layer.

Étude en base d'ondelettes de la dissipation par les écoulements dans les plasmas et dans les fluides

Résumé

Le problème de la dissipation par les écoulements macroscopiques est abordé par l'entremise de deux de ses manifestations les plus représentatives, la dissipation par les écoulements plasmas dans la limite de faible collisionalité, et la dissipation par les écoulements fluides dans la limite de faible viscosité. On part du principe que la dissipation peut avoir deux causes distinctes, soit l'effet résiduel d'un paramètre de couplage au niveau microscopique, soit l'effet purement macroscopique du mélange non-linéaire. La combinaison de ces deux phénomènes rend le problème impossible à traiter par les méthodes habituelles qui ont été appliquées avec succès aux systèmes conservatifs, et soulève des questions mathématiques fondamentales. De plus, le calcul explicite à toutes les échelles de tels écoulements n'est pas encore envisageable dut fait des limitations actuelles de la taille mémoire et de la vitesse des opérations. Il est donc communément admis que de nouvelles méthodes doivent être développées.

Dans ce but, on explore le potentiel d'une approche multi-échelles en ondelettes. Tout d'abord, les équations aux dérivées partielles décrivant l'écoulement doivent être reformulées dans le cadre d'une représentation discrète en ondelettes qui reste compatible avec les mécanismes dissipatifs soulignés plus haut. Cette étape, appelée régularisation, fait l'objet de deux chapitres de cette thèse, concernant les cas particuliers des équations de Vlasov-Poisson d'une part, et des équations d'Euler bi-dimensionelles incompressibles d'autre part. On évalue la faisabilité d'un calcul des écoulements en utilisant les schémas ainsi développés, et on compare ces derniers à d'autres schémas proposés précédemment.

Pour aller plus loin, il faut remonter aux origines de la dissipation résiduelle et la relier aux propriétés mathématiques des solutions. On obtient quelques éléments allant dans cette direction en étudiant numériquement le phénomène de collision d'un dipôle de vorticité avec une paroi dans la limite de faible viscosité. Lorque les solutions se comportent bien mathématiquement, comme c'est le cas pour les écoulements turbulents bi-dimensionels homogènes que nous abordons ensuite, on peut dors et déjà passer à l'étape suivante qui est la définition de la dissipation macroscopique dans la représentation en ondelettes. Finalement, on présente une analyse en ondelettes d'un écoulement turbulent tri-dimensionel dans une couche limite, ce qui ouvre des perspectives pour l'extension de la méthode.

Mots-clefs : ondelettes, turbulence, Navier-Stokes, Vlasov, couche limite.

Contents

Remerciements

En 2007 je finissais mon master sur la production d'énergie et les énergies renouvelables lorsque j'ai assisté au cours de Pascale Hennequin sur la fusion par confinement magnétique. Cela a été mon premier contact frontal avec ce mot qui fait peur: la turbulence. Alors j'ai cherché sur internet un stage sur la turbulence dans les plasmas magnétisés, et de fil en aiguille j'ai rencontré Marie Farge et Kai Schneider, qui ont bien voulu par la suite encadrer mon travail de thèse sur un sujet qui a beaucoup évolué, mais est resté dans l'esprit qui m'avait motivé au départ. Deux appartements, deux vélos, trois cartes bleues, beaucoup de kilos de châtaignes et de tartes plus tard, la fin est arrvivée. Ces années de thèse ont été intenses et enrichissantes sur tous les plans, surtout grâce aux personnes qui m'ont entouré et que je voudrais ici remercier.

J'ai le sentiment que Marie et Kai ont été des encadrants hors du commun, par leur ouverture d'esprit, leur organisation éclectique, et surtout leur passion pour la recherche et l'attrait sans fard qu'ils ont su donner aux questions les plus fondamentales soulevées par leurs travaux. Je leur serai toujours reconnaissant de la très grande confiance qu'ils m'ont faite et de tout ce qu'ils m'ont apporté, humainement et scientifiquement.

Je suis très reconnaissant également à Yanick Sarazin et Rupert Klein d'avoir bien voulu être rapporteurs de cette thèse, et d'avoir lu mon travail avec intérêt, et à Jean-Marcel Rax et Olivier Talagrand pour avoir eu la gentillesse de bien vouloir participer au jury. Mes remerciements vont également à Claude Bardos, non seulement pour sa contribution au jury, mais surtout pour sa présence continue pendant la thèse et pour ses conseils mathématiques qui ont été déterminants. Sur ce plan je suis aussi redevable à Alex Grossmann, Laure Saint-Raymond et Thierry Paul.

Xavier Garbet a été présent dès ma première visite à Cadarache début 2007, et je lui suit reconnaissant pour sa gentillesse et son soutien. L'aspect fusion de ma thèse a aussi été incarné par Nicolas Fedorczak et Pascale Monnier-Garbet, confrontés à la réalité des expérimentations sur Tore Supra, qui ont donné vie à toute une partie de mon sujet de thèse, et je les remercie pour leur gentillesse. Je remercie aussi Gérard Bonhomme et Pascale Hennequin pour leurs encouragements, et Éric Sonnendrücker dont l'intérêt pour le dévelopement de la méthode PIW a été indéfectible.

Durant cette thèse j'ai eu la chance d'être accueilli ici et là par des personnes remarquables. Je remercie Nick Kingsbury et Keith Moffatt pour leur hopitalité à Cambridge et pour leur aide, Diego del Castillo-Negrete pour une visite agréable à Oak Ridge et une collaboration fructueuse, Bill Dorland et Steve Cowley pour l'école d'hiver sur la fusion à Los Angeles en 2008, Mani, Vivek, Anant et Harsha pour un mémorable séjour à Dehli, le Wissenschaftkollegs zu Berlin où j'ai fait plusieurs rencontres intéressantes, et bien sûr le CIRM où je retournerai avec plaisir. Il faut aussi mentionner le Café Chéri(e) où plusieurs passages de cette thèse ont été rédigés.

J'ai été très content de mes nombreuses visites à Marseille et des rencontres de Salah, Benjamin, Wouter. Merci à Dmitry et Erwan pour les discussions interminables que nous avons eu sur de nombreux sujets. Au LMD j'ai aussi bénéficié d'un excellent environnement de travail en compagnie de Victor Duval, Max Bolot et Lionel Guez. J'ai aussi été très content de faire partie pendant ces quelques années de la jeune génération du LMD qui est vraiment cool. Je voudrais remercier le personnel de l'ENS et tout particulièrement

Marie-Christine Roos, Marie-Christine Goury et Matthieu Perrault pour leur gentillesse et leur disponibilité, et aussi Lionel Guez et l'équipe d'Idris. George Khuzadje a pris la peine d'appliquer la librairie de transformée en ondelettes que j'ai développé et je le remercie pour son intérêt.

J'ai certes appris beaucoup de choses pendant cette thèse, mais j'ai aussi désappris en grande partie la langue chinoise, malgré toutes les conversations que nous avons eues avec Yin Qi Zheng et Qiu Yan Qi. L'aspect mathématique de ces conversations doit sans doute être incriminé, mais il a heureusement nourri bon nombre d'éléments de cette thèse. Salut à LiMa, Martin, Jérémie, et tous les autres avec qui nous l'avons partagé l'été 2010 studieux à Luminy à l'occasion du CEMRACS. Ces années de thèse paraîtraient aussi incomplètes sans les joies théâtrales vécues en compagnie des Physiciens de la petite troupe, Yann, Pauline M, Pauline K, Antoine S, Antoine F, JB, Soline et Jeanne. Je voudrais également rappeler le souvenir de Cédric Bassilekin qui m'a appris plein de belles choses dans le temps. Et finalement rien n'aurait été possible sans tous mes amis proches que je ne vais pas tous citer ici, et bien sûr ma famille et apparentés, Raphaëlle qui a eu la patience d'écouter beaucoup de mes théories et l'impertinence de proposer les siennes, Lou, Poloche et Greg qui sont toujours là, mes frères et soeurs Dora, Émilie et Benjamin et mes parents, et bien sûr Lonia dont la porte a toujours été grande ouverte. Quant à mes grands-parents, leur honnêteté et leur grandeur d'âme m'ont procuré une énergie qui m'a été indispensable jusqu'ici et que j'espère garder longtemps encore.

Introduction

The long term goal underlying this work is the identification and description of macroscopic dissipative processes in classical flows. By macroscopic, we mean that which does not depend on the microscopic (i.e. molecular) properties of the system, for example on coupling coefficients such as fluid viscosity or plasma collisionality. The achievements presented here are extremely modest in regard to the ambition of this program. We have focused on two well known systems, turbulent flows in fluids, and filamentary flows in hot plasmas [1]. Both have been studied experimentally and theoretically for a long time, the medal going undoubtedly to turbulent flows in this respect. There has been a long term and fruitful connection between them on the theme of "wave turbulence", which appears in plasma as well as in fluids. However, the connection between the study of "strong turbulence" (or simply "turbulence", as we refer to it from now on) in fluid flows and phenomena occurring in hot plasmas has been fully realized only in the last few years. We have directly benefited from this new state of mind. Indeed, the collaborations going on at the moment between the communities working on the dynamo problem in astrophysics and geophysics, on turbulent fluids, and on the controlled nuclear fusion program is extremely active and stimulating, and have inspired a lot of what we are going to present.

Maybe it is time to slip a word or two about the scientific motivations from which this work has originated. The first comes from the atmospheric sciences, a field which features as a recurring concern the frontier between the predictable and the unpredictable. It is the wish of long range weather forecasters to push back this frontier as far as possible. For this we must bring under our scrutiny the phenomena that are currently beyond the frontier and understand what makes us unable to predict them. Turbulence is of course presumed guilty here, and this justifies the study of the loss of information which is entailed by dissipation in turbulence.

Another important motivation is the development of the research program on controlled nuclear fusion, a five decades long undertaking in terms of physics and engineering. The priority up to now was to write down the equations governing each part of the system, and to study their linear stability, in order to help the experimentalists reach situations that remain stable for a few seconds. This is now more or less complete. At present the tools for the analysis of these equations are still being developed, and some basic phenomena remain unexplained. But the main concern is now shifting towards an integrated system view, of a reactor that has to produce energy for years. This has been common practice in the fluid mechanics world for years. When the system is fully nonlinear and sustained, new types of questions emerge. What are the main structures that play a role in the disordered steady state? What reliable cycles can be implemented around the steady state to operate the reactor, for fueling, exhausting, etc. ? What will be the state of the vessel after a few months of operation? Understanding the matter of dissipation is essential to approach all these questions.

Let us now briefly mention some recurrent technical themes of this thesis. The first theme is multiscale analysis, and it comprises the orthogonal functions known as wavelets. Thanks to the unique properties of wavelets, flows can be at the same time discretized in conservative ways, and decomposed into superposition of multiple scales. Another theme that will pop up all the time is statistics. It will allow us to define dissipation, thanks to the

[1] The notion of filamentary flow will be introduced in Chapter 1

Bayesian – or "Jaynesian" – interpretation of statistical physics, and then to study it using wavelets and other tools. Finally, a third theme that we would like to mention is numerical experimentation, i.e. the explorations of the properties of a mathematical model by numerical integration using a computer. The results that we shall obtain with this approach are tied to the model from which they are obtained, and their purpose is not so much to directly learn about the underlying physical systems, than to understand the models better.

We dedicate a first chapter to a historically oriented presentation of the context of the present study. At the onset, a few open physical questions are stated, and the words of the title are defined. Then we split the presentation into three sections, dedicated to different and complementary approaches to the same problem: statistical, mathematical, and numerical.

In a second chapter, we set ourselves to give the elements of wavelet theory that are required to understand the sequel. A unified notation is proposed, that we specialize in the next chapters when working on more specific problems. In addition, we propose a complete formulation of an adaptive wavelet transform framework.We describe some choices that were made in the implementation, since we believe that they are crucial to actually make things work. To conclude this chapter, we report a few mathematical results that helped us on the way, and an application to experimental data processing.

The third chapter contains the results we have obtained in the field of plasma simulation. The first step was to study how wavelets could be used to obtain a regularized – or even denoised – estimate of the particle distribution function, starting from the position of many discrete markers. The second step was to actually implement in a code the estimation method that came out of the first step, and to solve the 1D Vlasov-Poisson problem in order to demonstrate the capability of the approach.

A closely connected topic is treated in the fourth chapter, namely the regularization of inviscid fluid equations. The first attempt that we made, reported in the first section of this chapter, was restricted to a toy model, the 1D Burgers equation, for which we showed the regularizing properties of a wavelet-based filter, relying on Kingsbury wavelets. The success of this approach drove us to generalize it to the 2D Euler equations, and to systematically compare it with several other regularization methods. The results are reported in the next section. In particular, we obtained the surprising result that simple Galerkin-truncation was sufficient to regularize the 2D incompressible Euler equations, in contrast to what happens for the 1D Burgers toy model. The final section of chapter 4 recalls some mathematical results that help understand the results of the two previous ones.

In the fifth and final chapter we start directly addressing physical questions that were only implicit in the preceding ones. A preliminary section is necessary to describe in some detail the penalization method that we apply in the second section to explore the issue of energy dissipation in 2D flows with solid boundaries. Then we come back to 2D wall-less flows and focus on conditional statistical modeling and the associated notion of macroscopic dissipation Finally, we report results on turbulent boundary layer analysis, obtained in a collaboration with a group at Technische Universität Darmstadt (Germany), that constitute a first step towards the extension of our approach to three dimensions.

Part I

Context and open questions

> Many years ago, on the basis of some "statistical assumptions" and reasoning that I have been unable to follow, Kolmogorov proposed the value $n = 5/3$, which does seem to have some experimental support. But its theoretical justification and range of validity do not seem at all well understood as yet. This is a problem for the future, on which I think we have a good chance of succeeding.
>
> E.T. Jaynes
> *Where do we go from here?* (1985)

> The circular arrangement of the axes of the Difference Engine 'round large central wheels led to the most extended prospects. The whole of artithmetic now appeared within the grasp of mechanism. A vague glimpse even of an Analytical Engine opened out, and I pursued with enthusiasm the shadowy vision.
>
> Lord Charles Babbage
> *Passages from the life of a philosopher* (1864)

Summary

I.1 Physical concepts

I.1.1 Turbulent fluid flows

Many types of fluid flows appear very disordered while possessing some reproducible properties. Two grassroot examples are the flow of air past an airplane wing, and the flow in the lower layers of the atmosphere on a windy day. Enough is known about the lift and drag forces exerted by the air on a wing to allow thousands of planes to fly safely everyday. Weather forecasting techniques can already predict the average wind speed with some confidence, and our own senses have recorded many subtle properties, most of which are even difficult to formulate rigorously, but such that we can immediately tell the difference between a strong breeze and a gusty storm precursor.

Striking descriptions of such flows are found in da Vinci's notebooks, but the first precise mechanical studies seem to date from the time of Newton (see, e.g., Truesdell, 1968). In the 18th century, rational mechanics was eagerly tackling many long standing issues in all domains of engineering, and flows were an important part of that huge undertaking. Two of the most important contributors of the time were Euler and d'Alembert. They were both prolific writers, to say the least, and reviewing their work would take us well beyond our scope. Let it suffice to give some elements that matter the most for for our purpose. Elaborating on the general principles set by Newton, and on the work of Jean Bernoulli, Euler worked out the conservation of linear momentum, introduced the general notion of pressure field and derived differential equations for the motion of fluids (Truesdell, 1968). He realized that certain fluids were effectively incompressible, and that the pressure field would then have to adjust itself so as to locally preserve the volume of fluid elements. The modern formulation of the equations he arrived at – the incompressible Euler equations – can be written as:

$$\begin{cases} \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \boldsymbol{\nabla})\mathbf{u} = -\boldsymbol{\nabla} p \\ \boldsymbol{\nabla} \cdot \mathbf{u} = 0 \end{cases} , \tag{I.1.1}$$

where $\mathbf{u}(\mathbf{x}, t)$ is the velocity field, $\boldsymbol{\nabla}$ is the gradient operator acting on $\mathbf{x}$, and p is the pressure field. The density of the fluid is assumed to remain constant and has been fixed to 1 by an appropriate choice of mass unit. In the following, we consider only this incompressible case, and shall therefore mostly refer to (I.1.1) simply as "the Euler equations". Note that we have also omitted the external body forces, which would appear on the right hand side of the first equation in (I.1.1).

D'Alembert (1768), working with the potential flow model that he had developed and that can also be seen as a special case of (I.1.1), pointed out that when a flow arrives from infinity at constant velocity and encounters a solid body, the fluid does not exert any force on the body in the flow direction, i.e., that the drag force vanishes. This conclusion, that has become known as the d'Alembert paradox, seemed in contradiction with many natural phenomena, like the flying of birds, and also with the everyday experience of engineers. For that matter, this was not the last occurrence of a contradiction between theory and experiment in the history of fluid mechanics, far from it. The required breakthrough was made in the 1820s, first by Navier (1823) with important contributions of Cauchy (1823), and then by Saint-Venant (1843) and Stokes (1845) using a different approach. The resulting theory was enriched by allowing for non-conservative contact forces interior to the flow,

leading to the now well established incompressible Navier-Stokes equations:

$$\begin{cases} \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \boldsymbol{\nabla})\mathbf{u} = -\boldsymbol{\nabla} p + \nu \Delta \mathbf{u} \\ \boldsymbol{\nabla} \cdot \mathbf{u} = 0 \end{cases}, \tag{I.1.2}$$

where ν is the kinematic viscosity of the fluid. Due to the occurrence of the Laplace operator Δ in (I.1.2), additional boundary conditions are required in the presence of solid obstacles, compared to (I.1.1) for which it is sufficient to impose the obvious non penetration of the flow. Navier (1823) proposed that the velocity component u_τ parallel to a wall should be proportional to the shear stress $\nu \partial_n u_\tau$ across the plane tangent to the wall at the same location, i.e.,

$$u_\tau + \alpha \partial_n u_\tau = 0, \tag{I.1.3}$$

where α is called the "slip length". The relation (I.1.3) is now called the Navier boundary condition. In rarefied gases, it can be justified from kinetic theory, as envisioned by Maxwell, and proved rigorously by (Coron, 1989). However, in dense gases and in liquids, it is more customary to use the "no-slip" boundary conditions, which correspond to the case $\alpha = 0$ above, and were proposed by Coulomb (1800) and advocated by Stokes (1845) after some hesitations. A detailed historical account of the controversies on this topic in the XIX-th century has been put together by Goldstein (1965, p. 676).

During the same period, thanks to the development of differential calculus, formal computations based on equations (I.1.1-I.1.2) allowed new interesting results to be derived. Helmholtz (1858) pointed out an important quantity, the vorticity

$$\boldsymbol{\omega} = \boldsymbol{\nabla} \times \mathbf{u}, \tag{I.1.4}$$

and used it to show that the circulation of the velocity field $\mathbf{u}$ along any closed loop advected with the flow was conserved for solutions to the incompressible Euler equations. This is also contained in a famous theorem of Kelvin (1869), which gave birth to a whole field of research now known as topological fluid dynamics. For flows which depend only on two space coordinates, ω is a scalar and it is conserved along the trajectories of infinitesimal particles advected by the flow, called Lagrangian markers. This is a first and instructive manifestation of the important differences between two-dimensional and three-dimensional flows, on which we return later.

Parallel to the theoretical work on the equations, fluid mechanics evolved into a mature experimental field during the 19th century. A milestone on this road was the study by Reynolds (1883) of flows in pipes for varying inflow velocities and diameters. He discovered the phenomenon of transition to turbulence, and showed that, for a given flow geometry, this transition was controlled by a single dimensionless parameter, now called the Reynolds number Re, corresponding to the ratio between inertia and viscous forces. The same phenomenology was reproduced in various flow geometries by Couette (1890), and later by G.I. Taylor (1936) and many others. It was hence understood that disordered flows corresponded to the case $\mathrm{Re} \gg 1$, that is, when nonlinear effects due to inertia are dominant. For example, the Reynolds number for a car on a motorway typically takes values between 10^6 and 10^7, while it reaches above 10^9 in the atmospheric boundary layer. We shall refer to this regime as "fully developed turbulence", which we shall most of the time abbreviate by the single word "turbulence". We thus stick to the strict definition of turbulence as the disordered motion of a fluid occurring at high Re. Note that the Reynolds number can be

used rigorously as a scaling parameter when the geometry, boundary conditions, and initial data are fixed, whereas more informal comparisons require a lot of caution.

Major progress towards a resolution of d'Alembert's paradox using the Navier-Stokes equations was made by Prandtl (1904). Given a flow in contact with a solid obstacle, and assuming $\mathrm{Re} \gg 1$, Prandtl focuses on a thin neighborhood of the obstacle of thickness scaling like $\mathrm{Re}^{-\frac{1}{2}}$, called the boundary layer. By a suitable rescaling of variables, he shows that under certain hypotheses the flow inside this layer is a solution to some simplified equations, now called the Prandtl equations (see Schlichting, 1979 and Sec. I.3.3 below). To describe the flow outside the layer, the Euler equations are to be used, considering that since $\mathrm{Re} \gg 1$ the effects of viscous forces are negligible far from the wall. If everything goes well in the limiting process, the scaling of the drag force with Re is constrained to be $\mathrm{Re}^{-\frac{1}{2}}$. However in the 1930s it became clear from the results of many aerodynamic experiments including those by the groups of Prandtl, Burgers, von Kármán, and Dryden that this scaling breaks down for sufficiently large Re. Visual observation of the flow in the new regime reveals that in some regions the boundary layer tends to "detach" from the wall, in the sense that some fluid particles that are close to the wall at some time get entrained far from it at later times.

This new regime has fascinating consequences on the behavior at large Re of the drag force F_D, which can be written by dimensional analysis as

$$F_D = C_D \rho U^2 L^2,$$

where ρ is the fluid density, U is some characteristic velocity, L is some characteristic length, and C_D is a non-dimensional number called the drag coefficient. It turns out that the most simple approximation is that the drag coefficient depends only on the geometry of the obstacle and not on the Reynolds number. Most drivers know that the fuel consumption of their vehicle is roughly proportional to the square of its velocity. This empirical behavior of C_D, already reported by Newton in his study of the pendulum, was familiar to civil engineers at the beginning of the XIX-th century, as mentioned also by Coulomb (1800), and recalled by Darcy (1858) in his 260 pages mémoire on pipe flows. Since then much more empirical knowledge has been accumulated, concerning for example the drag crisis (the sudden reduction in the drag coefficient first observed for flows around spheres at $\mathrm{Re} \sim 10^5$ by Eiffel, and studied by Nikuradse and Prandtl in the Göttingen wind tunnel, see Darrigol, 2005), logarithmic dependency with Re in confined flows and the influence of rugosity (see e.g. Schlichting, 1979). Further theoretical investigations were conducted by von Kármán (1921, 1930) and Prandtl (1925, 1932) and have led to a few verifiable statements known today as the law of the wall. This theory is essentially based on similarity arguments and as such contains several unknowns whose behavior remains to be explained. Moreover even its most basic predictions are still challenged by comparison with modern experimental data, as we shall see below. We should also mention that in a largely forgotten work, Burgers (1923) worked out a simple 2D kinematic model for a fully developed turbulent channel flow involving vortices of size scaling like Re^{-1} in the neighborhood of the walls, whose presence is sufficient to qualitatively explain the observed drag force. The difficulty to incorporate dynamical effects in such a model may explain its lack of popularity. Overall, it appears that the explanation of the behavior of C_D in the fully developed turbulence regime is still an open problem. It is one of the most basic, long standing, and yet central questions in the field of turbulence.

The drag question has at least one reassuring quality, namely that it is clearly formu-

lated, and that any proposed theory is therefore liable to immediate experimental falsification. Another important question concerns the energy dissipation rate of homogeneous isotropic turbulence (HIT), but it is arguably less clearly cut than the drag question, because the definitions of the words "homogeneous" and "isotropic" rest on non trivial statistical concepts. For the moment, let us simply imagine a region of turbulent flow that is left to itself and far from any forcing device. By averaging the kinetic energy contained in this region, one can measure an energy dissipation rate ε. Then by repeating this experiment at different Reynolds numbers, while keeping the same preparation protocol for the flow, one can study the dependency of ε as a function of Re. In practice, this has been done in wind tunnels using so called "grid generated turbulence" (Burgers, 1926; Dryden and Kuethe, 1929; Batchelor and Townsend, 1948; Comte-Bellot and Corrsin, 1966) and the outcome showed that ε becomes almost independent on Re for Re sufficiently large (Sreenivasan, 1984). Coming back to the car example, we can only admire the great consistency of Nature: for large Re, energy is injected in the flow due to the non-vanishing drag at the surface of the vehicle (an apparently very local phenomenon), and the flow arranges itself so that it manages to dissipate this energy thanks to the non-vanishing ε in its bulk! (Frisch, 1995).

Other open questions in the field of turbulence suffer from a less enticing guise. Among them, one finds almost everything that has to do with localized flow events. It is thus an open question to even formulate such notions as a "predictable local event" in a turbulent flow. Far from being purely academic, these topics are of central concern for many applications, among which weather forecasting and the study of pollution spreading.

I.1.2 Filamentary plasma flows

Before coming to plasmas, we are compelled to make a brief return to the kinetic theory of gases, which will serve as a transition. Around the same period of time when the mechanical aspects of fluids were studied, the successes of thermodynamics motivated some physicists to propose microscopic models with the hope that it would allow them to directly compute the values of some thermodynamic coefficients that had been measured experimentally [2]. Their starting hypothesis was that the properties of matter resulted from simple laws of motion followed by very many molecules (Maxwell, 1867). In this way, Maxwell explained the stability of Saturn's rings and computed the velocity distribution of particles in a gas. This distribution, usually denoted f, is the central object of kinetic theory. It is defined by the property that for any $\Omega \subset \mathbb{R}^6$, $\int_\Omega f(\mathbf{x},\mathbf{v},t)\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{v}$ is the probability of finding at time t a particle whose position $\mathbf{x}$ and velocity $\mathbf{v}$ are such that $(\mathbf{x},\mathbf{v}) \in \Omega$. Boltzmann (1872) then derived the equation now bearing his name, which describes the time evolution of the aforementioned distribution in the case of a dilute gas:

$$\frac{\partial f}{\partial t} + \mathbf{v}\cdot\boldsymbol{\nabla}_{\mathbf{x}} f + \frac{\mathbf{F}_{ext}}{m}\cdot\boldsymbol{\nabla}_{\mathbf{v}} f = C(f), \tag{I.1.5}$$

where $\mathbf{F}_{ext}$ is an external force acting on each particle, m is the mass of a particle, and $C(f)$ models the effect of binary collisions. He proved the inevitable increase of the quantity $S_B(f) = -\int f\log f$, which we refer to as the Boltzmann entropy, and made the bold conjecture that this was the reason for the second principle of thermodynamics. A major

[2] Here and in the following, the term microscopic is strictly reserved to phenomena occurring at a molecular scale, independently of the description chosen for the system.

argument supporting his view was that the only distribution of molecular velocities maximizing his entropy corresponded to the distributions previously obtained by Maxwell, and which could be deduced in an independent manner using the notion of statistical ensemble that Boltzmann had introduced earlier, and that had been further developed by Gibbs. The increase of Boltzmann entropy, although it does not prove the second principle of thermodynamics as is now well understood (see e.g. Jaynes, 1965), concerns us nevertheless because it justifies the passage from the molecular description to the fluid description. The process, known as the Chapman-Enskog expansion, consists in a derivation of the fluid flow equations by taking the low order moments of (I.1.5) with respect to $\mathbf{v}$. For its success, it is essential that f is close to a local Maxwellian distribution, a state called local thermodynamic equilibrium which is established when the mean free path of molecules is small with respect to the macroscopic length scale.

Plasmas are like gases except that they contain charged particles, i.e., ions and electrons, instead of only neutral molecules. The word plasma was coined by Langmuir (1928) when he first described the waves that now bear his name. Langmuir used the essential concept of shielding, introduced earlier by Debye and Hückel (1923), which roughly stated means that a particle effectively experiences forces only from other particles closer than a certain length, the Debye length. In plasmas, the Boltzmann collision operator cannot be used, and the matter of collisions become altogether very complicated (see Villani, 2002 for a review). Fortunately, there exists a regime where interesting macroscopic phenomena occur at length and time scales *smaller* than the mean free path and mean free time. This regime, corresponding to the opposite of the fluid regime, is usually called "hot", because it is most commonly triggered by raising the temperature of the plasma. The effects of the collision term $C(f)$ can then be replaced by a mean field approximation, where the electromagnetic forces are added to $\mathbf{F}_{ext}$. The Vlasov equation is thus obtained:

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \boldsymbol{\nabla}_{\mathbf{x}} f + \frac{\mathbf{F} + \mathbf{F_{ext}}}{m} \cdot \boldsymbol{\nabla}_{\mathbf{v}} f = 0 \tag{I.1.6}$$

where $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$, with q the charge of the considered particle, and $\mathbf{E}$ and $\mathbf{B}$ respectively the electric and magnetic fields due to the particles themselves. In the following, and in the title of this thesis, "plasma flow" refers to the distribution function f, which plays the same role in a hot plasma as does the velocity field in a fluid. To obtain a closed system of equations, one writes down the Maxwell equations for $\mathbf{E}$ and $\mathbf{B}$, thus yielding a nonlinear system called the Vlasov-Maxwell equations, or Vlasov-Poisson in the special case when the magnetic field vanishes or can be neglected.

Contrary to the Boltzmann equation which relies on strong hypotheses concerning the post-collisional velocities, the Vlasov equation can be derived directly from the Newton equations of molecular dynamics, although at the moment the derivation has been made rigorous only for a mollified interaction potential (Braun and Hepp, 1977). The approximation leading to the Vlasov equation is controlled by the largeness of the so-called "plasma parameter", usually denoted Λ, which is the number of particles contained in a sphere whose radius is the Debye length. The Vlasov equations have been used to derive many important results concerning plasmas, including the explanation of Landau damping, i.e., the transfer of energy between waves and particles due to resonant interactions. Unsurprisingly, the main shortcoming of the Vlasov equation is that it does not describe the effect of collisions.

As for turbulent fluid flows, there are many open questions concerning hot plasma flows. A number of interesting ones have applications in the field of magnetically confined nuclear

fusion, which we take as an illustrative example. Due to Coulomb repulsion, the fusion between two nuclei can occur only if their kinetic energy is very large, for example compared to the one of molecules in a gas at a normal temperature. To achieve an appreciable amount of fusion events, it is thus necessary to reach temperatures of several million Kelvins, at which a gas has long since turned into a plasma. By surrounding the plasma with suitably arranged coils, the Lorentz force can be tailored as a confinement device, and the system can, in principle, function as a stationary source of energy. In practice, the situation is made extremely complex by the wild variety of instabilities that feed on the free energy present in the system. The resulting fluctuations are often referred to as "plasma turbulence", because of the stochastic nature of the flow and of the waves observed in the plasma. But it should be stressed that this has little to do, in terms of the quantities involved, with fluid turbulence as we have seen it in the previous section. The rotational motion occurring generically in incompressible fluid flows, and in which the word "turbulence" has its etymological roots, is observed in plasmas only in special cases. In a recent proposition to extend the concept of "cascade" to magnetized plasmas (Schekochihin et al., 2008), turbulence is defined roughly as a process transferring energy from large to small scales, which seems at the same time too restrictive – since flows are typically forced locally in space by the effect of boundaries – and too vague – since there are non-turbulent, yet very efficient ways of transferring energy to fine scales, e.g. shock formation in compressible fluids. We would like to avoid these semantic issues altogether by using the word turbulence, for the time being, only when it applies to fluid flows. The use of the term "hot plasma" is common to refer to a plasma in the low collisionality regime, but it would be more satisfactory to qualify f rather than the plasma, in the same manner that the adjective "turbulent" qualifies the flow and not the fluid. In the following, we use the adjective "filamentary" to refer to such flows, since it has been observed that f develops filaments in phase space, which get more numerous as the collisionality is lowered. As the corresponding noun we choose "filamentation".

One of the goals of the research program on magnetically confined fusion is to understand the fluctuations in hot plasmas in order to improve the energy confinement time of the device, and thus to open the possibility of using fusion as an economically sustainable source of energy. In the currently most successful configuration, called the tokamak, the plasma has the shape of a torus. The coils produce a toroidal magnetic field, and an additional poloidal field is generated by a strong current circulating in the plasma itself. For lack of a cheaper solution, the most utilized manner of increasing the energy confinement time has been to increase the size of the device. The largest one is called the Joint European Torus (JET) and is located in Culham (UK). Its major and minor radius are respectively 3 meters and 90 centimeters. The International Tokamak Experimental Reactor (ITER), currently under construction in Cadarache (France), will be about twice as large. For reasons that we do not detail, the largest the size of the device, the smallest the collisionality of the plasma. Hot plasmas are thus at the heart of the fusion program. They are also well represented in astrophysics. But the understanding of hot plasmas is far from complete from a theoretical point of view. In a sense the situation is even more dire than what we have mentioned concerning turbulence, since for plasmas the relevant collision model, the Landau operator, is extremely complicated and poorly understood, and in addition the respective status of other approximate models is not clear from a mathematical point of view.

I.1.3 Dissipation

Now that we have separately introduced turbulence and filamentation, and summarized some open problems of general interest concerning these two regimes, we would like to explain what brings them together in this thesis, and to formulate the specific questions we are seeking to answer. As suggested by the title, it turns out that the key concept is dissipation.

In the common language, the word dissipation can take many connotations, from the negative judgment addressed to an "intemperate, dissolute, or vicious way of living" (Oxford English Dictionary, 1931) to the the resolution of a misunderstanding which leaves a vague impression still floating in the air. In the following, we concentrate on the precise concept of dissipation in modern physics, defined by Balian (2006) as an increase of the entropy. That brings us to the definition of entropy, the central concept of statistical physics. Fascinating historical accounts on the subject are to be found in (Brush, 1976; Truesdell, 1980; Gallavotti, 1995). We sketch here only the modern view, with a special emphasis on the previously introduced physical systems.

For simplicity we consider an isolated (or thermostated) fluid or plasma. To any probability density ρ over microscopic states corresponds a microscopic entropy, which we denote by $S_G(\rho)$, the Gibbs entropy:

$$S_G(\rho) = -\int_\Gamma \rho \log \rho, \tag{I.1.7}$$

where Γ denotes the phase space. Because of the Liouville theorem, S_G is conserved by the Hamiltonian microscopic dynamics of the fluid or plasma. At this level, it is therefore basically a useless quantity. However, it turns out that S_G is maximal over all admissible ρ's when ρ is an equilibrium distribution, where the macroscopic flow vanishes. The value of S_G realized by such an equilibrium state may be called the equilibrium entropy, S_E (experimental entropy in (Jaynes, 1965), sometimes called Clausius entropy). Hence S_G is well suited to the analysis of equilibrium problems, or of perturbations thereof.

But a moving fluid is by definition out of thermodynamic equilibrium, and therefore its properties – including dissipation – can not be computed using S_G. Nevertheless, since parcels of fluids are assumed to be in local thermodynamic equilibrium, an entropy density s_L can be defined and integrated over the fluid volume to yield a time dependent non-equilibrium entropy S_L of the moving fluid. S_L can be shown to increase monotonically in time, because of molecular friction, which makes the fluid relax towards a global equilibrium characterized by $S_L = S_E$. In general, one has only the inequality $S_L \leq S_E$. The increase of S_L in time corresponds to what is usually referred to as "molecular dissipation", a contraction for "dissipation due to molecular friction forces". There have been numerous attempts to base predictive theories of turbulence on hypotheses concerning S_L. The "maximum entropy production principle" and the "minimum entropy production principle" fall into that category, which we do not develop further.

In a plasma described, for example, by the Boltzmann equation, the role of S_L is played by the Boltzmann entropy S_B, which takes into account the absence of local thermodynamic equilibrium. Hence the increase of S_B in a plasma is a form of dissipation. Note that the Boltzmann equation can also be used to describe a dilute gas, which may happen to be in local thermodynamic equilibrium, so that the relevant entropy could well be S_L. This apparent conflict allows us to pinpoint the essential notion that the entropy depends on the description of the system currently being considered. For a fluid, it is natural to speak of

"energy dissipation", because the increase of entropy is equivalent to the decrease of kinetic energy. With that terminology in mind, we are strongly driven to wonder what is the quantity that is being dissipated in a plasma, and we feel that there ought to be such a quantity, and that the use of the word dissipation does not make sense without it. Sometimes, the Boltzmann entropy is defined without the minus sign and one speaks of "entropy dissipation", a strange expression which seems at odds with the simple definition of dissipation recalled above.

Before we can help soothe the linguistic itch, we need to say a little bit more about the interpretation of the entropy itself. Since the work of Jaynes (1957), it is clear that S_G should be interpreted as a quantifier of the *spreading* of the distribution ρ over Γ. For example, if Γ has a finite volume, S_G will reach its maximum for the uniform distribution. The entropy is hence seen as a lack of information. Complete information about the system would mean that ρ is concentrated in a single point of Γ, in which limit S_G tends towards $-\infty$. Now coming back to the word "dissipation", we see that it comes directly from the Latin verb "dissipare", constructed by adding the prefix "dis" – apart – to "supare" – to throw away, to spread. It is therefore legitimate to speak of dissipation of the probability density ρ, whose spreading indeed corresponds to an entropy increase, and also to a rejection of information.

In our view, the link between turbulence and filamentation lies in the interplay of macroscopic mixing with microscopic dissipation. In turbulent fluid flows, macroscopic mixing arises because of the advection term in the Euler equations, and microscopic dissipation because of molecular friction. In filamentary plasma flows, macroscopic mixing can be attributed to the nonlinear term in the Vlasov equation, and microscopic dissipation to residual collisional effects. The main theoretical prospect that motivates our study on the long term is to bring these macroscopic and microscopic phenomena under the scrutiny of a unified set of concepts. But first, we present some of the lines of thought that have been followed up to now to tackle such problems.

I.2 Statistical models

I.2.1 (Un-)Predictability

Following the trend at the beginning of the XX-th century, it was legitimate to attack the issue of turbulence by the deterministic edge, that is, to try to predict the future evolution of the flow from the knowledge of its present state. This innocent wish has opened a Pandora box of questions pertaining to predictability, through which we want to make a small detour before coming to the statistical models per se. The predictions we are interested in are announcements of future events, or more generally assignments of probabilities to future events, using the rational knowledge of causes and effects. Predictability studies thus naturally share a strong bond with numerical weather forecasting, a field which can be traced back to the famous book by Richardson (1922). The declared goal of the field is to predict the evolution of flows and other components of the atmospheric systems as long as can be achieved given a certain array of measurement devices. This problem was discussed at length by Lorenz (1969a,b, 1982), who introduced the notion of "range of predictability" in the context of weather forecasting. He popularized the notion that accurate predictions

were possible for a short time – of the order of one or two weeks in the atmosphere – and impossible for longer times.

Beyond weather forecasting, crude approximations to the Navier-Stokes equations are now used on an everyday basis in an attempt to answer questions that depend on the local properties of the flow, for example: will an oil spill in the gulf of Mexico affect the fishing industry in the North Atlantic? And even more questions seem completely outside of the range of our predicting tools, for example: will the tree in your garden survive the next storm? These questions are very difficult to answer because the flow evolution is very sensitive to perturbations in its initial conditions. That is, two flows initially very similar to each other are likely to look very different after some time. This property defines a class of dynamical systems that are called "chaotic". For the weather forecast problem and others, it seems that our eagerness to predict needs to be somewhat tempered as soon as chaos is at work. However, this does not mean that the deterministic description of Nature is irrelevant. For any system, chaotic or not chaotic, attempts at predicting the future are prone to some amount of uncertainty. It is thus largely irrelevant to speak of unpredictable *systems*, but it is a key question to find out *what is predictable* about a system, given a certain knowledge about its present state and environment.

Perhaps the first studies that were related to this issue concern the stability properties of known stationary solutions to the Euler equations, as introduced by Rayleigh (1879). The topic continues today with the study of transition to turbulence, or more generally transition to chaos. However, in the fully developed turbulent regime, flows are non stationary and it is very difficult to relate their properties to those of stationary or quasi-stationary solutions, although there has been some amount of work along this line (Robert and Sommeria, 1992). Other tools like eigenmode expansions do not seem well suited to address predictability, and even generalized concepts like Lyapunov exponents fail to capture its essential features (Lorenz, 1996). In particular, predictability could depend on the type and amplitude of the initial perturbation, as was shown by Robert and Rosier (2001) for the 2D incompressible Euler equations, and by Zaliapin and Ghil (2010) in the broader context of the Earth climate system. In fact, to this day most practically relevant issues pertaining to predictability are still largely open, as for example the prediction of prediction errors, or the characterization of the perturbations to which the flow is the most sensitive on the long term.

I.2.2 The K41 and KBL67 theories

The large number of active degrees of freedom observed in turbulent flows, and the phenomenological arguments supporting their unpredictability, seem to call for a statistical theory. This was also a natural route to follow if one strove to interpret the somewhat peculiar experimental observations on turbulent flows, and was advocated for that purpose by Burgers (1929) and Taylor (1935). The idea soon attracted the interest of mathematicians specialized in probability theory, like Kolmogorov (1941) (hereafter K41) who, inspired by experimental results and ideas of Obukhov (1941), took as central quantity the statistical distribution F of the difference $\delta\mathbf{u}(\mathbf{x}, t, \mathbf{r})$ between the velocities at two points in a 3D flow separated by the vector $\mathbf{r}$. To define this distribution, they modeled the flow by a random variable living in some ensemble of realizations, and indexed by some probability space, which they had to leave unspecified for lack of a better way. They then made several hypotheses. The first was that the flow was statistically stationary, homogeneous and isotropic, i.e. that F did not

depend on t and $\mathbf{x}$, and was covariant under change of coordinates by rotation and translation. The second hypothesis was that F depended only on $\mathbf{r}$, on the Reynolds number, and on the average energy dissipation rate, denoted $\overline{\varepsilon}$. Finally, the third hypothesis was that for $|\mathbf{r}|$ larger than a certain length $\eta = \nu^{\frac{3}{4}}\overline{\varepsilon}^{-\frac{1}{4}}$ – now called the Kolmogorov length – F ceased to depend on Re, and depended only on r and on $\overline{\varepsilon}$.

From these three hypotheses, K41 derived in particular that:

$$\left\langle \frac{|\delta\mathbf{u}\cdot\mathbf{r}|^2}{|\mathbf{r}|^2} \right\rangle = C\overline{\varepsilon}^{\frac{2}{3}}|\mathbf{r}|^{\frac{2}{3}}, \tag{I.2.1}$$

where C is a constant. This prediction has shown great success in describing many experimental results. It is perhaps better known as the Kolmogorov $5/3$ law because it is equivalent to

$$\langle E(k)\rangle = C_K\overline{\varepsilon}^{2/3}k^{-\frac{5}{3}},$$

where $E(k)$ is the energy spectrum of the flow, defined as the energy content of a shell $k \leq k < k+1$ in Fourier space, and C_K is called the Kolmogorov constant. Later on, followers of Kolmogorov realized that under the same hypotheses as K41 much stronger constraints on the overall flow statistics could be deduced, in particular concerning the higher order moments of the velocity difference distribution, which are known as structure functions:

$$\left\langle \frac{|\delta\mathbf{u}\cdot\mathbf{r}|^p}{|\mathbf{r}|^p} \right\rangle = C\overline{\varepsilon}^{\frac{p}{3}}|\mathbf{r}|^{\frac{p}{3}} \tag{I.2.2}$$

for any positive p. And unfortunately, such generalized predictions, commonly referred to as an extended "K41 theory", typically disagree with experiments. In particular, it was shown that structure functions increased more slowly with r, than predicted by (I.2.2), see for example (Anselmet et al., 1984) for landmark results on this. To explain this discrepancy it can be argued that the idealized hypotheses underlying K41 do not correspond to the situation encountered in experiments. But the same discrepancy was observed in numerical experiments constructed to be compatible with K41 (see e.g. Ishihara et al., 2009). In that case the flaw is likely due to large deviations of typical individual realizations, for example because activity is more and more localized in regions occupying a vanishingly small fraction of space when going to fine scales, a phenomenon known as intermittency. Another hypothesis which is often associated to K41 is that the flow does not have any dynamically important features at scales finer than the Kolmogorov scale (i.e. that the evolution will be the same if those scales are filtered out). Given that η is proportional to $L\mathrm{Re}^{-\frac{3}{4}}$ (where L is the length scale used to define the Reynolds number), this hypothesis leads to an estimate of the number of active degrees of freedom of order $\mathrm{Re}^{\frac{9}{4}}$.

The case of 2D flows was later studied by Kraichnan (1967), Leith (1968) and Batchelor (1969) (hereafter KBL67) who made use of the same kind of arguments. The situation is made a little bit more complex by the existence of two independent quadratic quantities conserved by the 2D Euler equations, namely the energy and the enstrophy $Z = \frac{1}{2}\int_\Omega \omega^2$. As a result, two different behaviors for the structure functions are possible in such a theory. KBL67 conjectured that the enstrophy dissipation rate $\overline{\eta}$ would play the same role in 2D as the energy dissipation rate in 3D, while in 2D the energy dissipation rate would vanish in the fully developed turbulence regime. This implies in particular that, if L denotes the "energy injection scale", then for $r \leq L$, the structure functions should behave like $\overline{\eta}^{\frac{p}{3}}r^p$. The corresponding energy spectrum is proportional to k^{-3}.

The derivation of theories "à la Kolmogorov" to describe plasma flows is currently a hot research topic. Indeed, it has been noticed that the same phenomenology was ongoing in hot plasmas, albeit with a number of new features due to the fact that plasmas have their own characteristic length and time scales so that their flows are not approximately scale invariant like fluid flows, but tend to develop different regimes depending on the scale that is being considered (Schekochihin et al., 2008). Despite these additional difficulties, Gürcan et al. (2009) were able to derive an expression for the statistically stationary spectrum of drift wave turbulence, and showed that it agreed reasonably well with experimental measurements in the scrape-off layer (SOL) of the tokamak Tore Supra.

To take into account the observed intermittency of turbulent flows, it is possible to allow the scaling exponent of the velocity increments to depend on space. This leads to the multifractal framework (Frisch and Parisi, 1985; Benzi et al., 1984; Meneveau and Sreenivasan, 1987). However, in its present formulation, the approach still lacks explanatory power, and it still cannot be derived from the NSE. Even when leaving aside the matter of intermittency the main shortcoming of this sort of statistical theory is that their range of application is very limited. Indeed the underlying hypotheses are rarely satisfied in practical situations. We come back in some detail on these points in Sec. V.3.2.

I.2.3 Turbulence models

To make theoretical progress, and at the same time to help engineers confronted to practical problems, there is a strong need of reduced models of turbulence. Since it would be unreasonable to attempt to review even all the different kinds of turbulence models, we limit ourselves to a few illustratory examples (see Schiestel, 1993 for more).

One of the most natural approaches to derive statistical models is to start by averaging the Navier-Stokes equations over an ensemble of realizations. As remarked already by Reynolds (1995), the resulting equation is not closed, since it involves averages of products that can not be reduced to products of averages. Nevertheless a second equation can be obtained to describe the time evolution of these products, which in turn depends on rank three tensors, etc. In the end an infinite hierarchy of moment equations is obtained, and turbulence models attempt to truncate this hierarchy to only a few equations by imposing a model for the moment at some predefined order.

Such closed models can be divided into two categories, one point closures and two point closures. One of the most popular two point closures is the eddy-damped quasi normal Markovian (EDQNM) closure, introduced by Orszag (1970), and for which closure is imposed at the level of fourth order moments. Another approach starts from the equations written in Fourier space and uses them to propose closed evolution equations for the energy spectrum of the flow, called "shell models". This practice is well established in the fluid turbulence community, and has recently been applied to some equations relevant for hot plasmas (Gürcan et al., 2010). On the other hand, one point closure models are usually referred to as Reynolds-averaged Navier-Stokes equations (RANS), since their main objective is to model the Reynolds stresses, i.e. the second order centered moments of the velocity field. RANS is widely used in engineering applications, since its mathematical simplicity allows for a large range of effects to be included in some phenomenological way. See the review by Speziale (1991) for more on RANS.

Instead of averaging over a statistical ensemble, another idea that has been pursued is to coarse-grain the velocity field by convolving it with a smoothing kernel, by truncating its Fourier expansion, or by any other more sophisticated means. Once again the resulting equations are not closed, and this time a model has to be introduced to describe the so-called "sub-grid scales", i.e. the scales that have been lost to coarse-graining. This idea was put forward by Prandtl (1925) and Richardson and Gaunt (1930) who introduced what is now called "eddy viscosity", which consists in modeling the effects of the sub-grid scales simply by increasing the viscosity coefficient. The latter authors also proposed a statistical interpretation on which we shall come back in Sec. V.3.2. More advanced numerical implementations of the same approach are now called large eddy simulations (LES), as pioneered by Smagorinsky (1963) and Lilly (1967) for modeling the general circulation of the atmosphere, and developed by Deardorff (1970) and later authors for application to engineering related problems. The simple eddy viscosity is there replaced by a coefficient depending on the local properties of the flow. Later Kraichnan (1976) advocated the use of a spectral eddy viscosity depending on wave-number, and thus shed light on the link between LES and statistical theories of turbulence. In the context of atmospheric modeling, the Smagorinksy model has been superseded by other approaches like hyper-dissipation (Basdevant et al., 1981), which consists in replacing the term $\nu\Delta\mathbf{u}$ by $-\nu_\alpha(-\Delta)^\alpha\mathbf{u}$ with $\alpha > 1$, and that we shall encounter again in Sec. IV.2. Today LES has become a very popular mean of simulating complex engineering flows (see e.g. Ferziger and Perić, 2002; Lesieur et al., 2005), but it still suffers from severe shortcomings.

I.3 Mathematical view

In addition to the statistical theories outlined above, there have been a number of attempts to derive results on turbulence from first principles. This section recalls a few issues that were encountered in doing so.

I.3.1 Weak solutions and well-posedness

The first question coming to the mind of a mathematician when given a initial value problem from physics concerns the existence and uniqueness of solutions. She may also wonder if the solutions depends continuously on the initial data. These questions were first formulated explicitly by Hadamard (1902) and they are at the basis of the well-posedness theory for differential equations. For partial differential equations (PDEs), it was possible to start fulfilling the program proposed by Hadamard only after Leray (1934) introduced the concept of weak solution. In the modern formulation, one may simply say that a weak solution satisfies a PDE "in the sense of distributions", and we do not reproduce the definition here. In contrast, a solution is said to be "classical" when it satisfies the PDE in the point-wise sense. We feel that it is important to review a few important results concerning the well-posedness of the Navier-Stokes, Euler and Vlasov-Poisson equations.

I.3.1.1 Navier-Stokes equations

The strong formulation of the initial value problem for the Navier-Stokes equations (NSE) is as follows (Foiaš et al., 2001). Given an initial vector field $\mathbf{u}_0$ defined on an open subset Ω of $\mathbb{R}^d$ ($d = 2$ or 3), taking values in $\mathbb{R}^d$, and such that its divergence $\boldsymbol{\nabla} \cdot \mathbf{u}_0$ vanishes, and given a final time $T > 0$, find sufficiently regular fields $\mathbf{u} : \Omega \times [0, T] \to \mathbb{R}^d$ and $p : \Omega \times [0, T] \to \mathbb{R}$ satisfying (I.1.2) such that $\mathbf{u}(\cdot, 0) = \mathbf{u}_0$ and $\mathbf{u}(\mathbf{x}, \cdot) = 0$ on $\partial\Omega$ (no-slip boundary conditions). Sometimes, instead of $\Omega \subset \mathbb{R}^d$, we shall encounter the case $\Omega \subset \mathbb{T}^d$, where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ is the unit torus, or even $\Omega = \mathbb{T}^d$. The latter corresponds to a periodic flow filling the whole space, which is quite non physical but very convenient mathematically and numerically. In the following we shall always assume that the frontier of Ω is as regular as necessary, since the presence of corners or other boundary singularities introduces additional difficulties which we want to avoid.

The well-posedness results on this initial value problem are divided into two main categories: long time results, and short time results. The first kind holds on arbitrary time intervals $[0, T]$, whereas the second kind holds only on sufficiently short time intervals $[0, T_\star]$, where $T_\star$ may depend on the initial data. Long time existence of weak solutions, for initial data in L^2, has first been established by Leray (1934) in the whole space, and then by Hopf (1951) in bounded domains. Solutions constructed using their approach are now called Leray-Hopf weak solutions. In two dimensions (2D), thanks to results of Lions and Prodi (1959) and Ladyzhenskaya (1959), the uniqueness of these weak solutions is established, as well as their continuity with respect to initial data. Therefore the well-posedness in the sense of Hadamard is granted.

In three dimensions (3D), only short time uniqueness is proved, for strong solutions and smooth initial data. The possible lack of uniqueness is related to the occurrence of singularities in the Leray-Hopf solutions, which can be shown to require that the velocity becomes infinite in some points of space. The set of points in space-time where this occurs is at the moment best controlled by a theorem of Caffarelli et al. (1982) asserting that its Hausdorff dimension is strictly less than one. Note that the apparition of singularities in a regular solution is often referred to as "blow-up" of the solution. Although the blow-up of the 3D NSE is still an important open problems in mathematical physics (Fefferman, 2000), it is widely believed that solving it will not help much in the answering of physical questions, mostly because it involves non physical infinite velocities.

I.3.1.2 Euler equations

Recent nice reviews on the mathematical theory of the incompressible Euler equations may be found in (Bardos and Titi, 2007) and (Constantin, 2007). For the Euler equations, there is no general and well established existence theory for weak solutions. The difficulties encountered when attempting to build such a theory are revealed by a series of striking results by Scheffer (1976), Shnirelman (1998) and de Lellis and Székelyhidi (2009, 2010), who constructed so-called "wild solutions" of the Euler equations, that can suddenly start to oscillate when they were equal to zero a moment before, and whose energy can vary in an almost arbitrary way!

In 2D, the best long time existence and uniqueness results for weak solutions were proved by Yudovich (1963) in the case of bounded vorticity ($\omega \in L^\infty$). Delort (1991) proved long

time existence in the case where ω is a measure with simple changes of sign. Note that this last case is especially interesting since it is closely related to the problem of the evolution of a vortex sheet, which has important physical applications.

In 3D, short time existence and uniqueness is proved for sufficiently smooth initial data, but for long times both existence and uniqueness are open. A crucial theorem of Beale, Kato, and Majda (1984) (BKM) asserts that if a smooth solution exists on the time interval $[0, T_\star]$, and if it is such that

$$\int_0^{T_\star} \max_\Omega |\boldsymbol{\omega}| < +\infty,$$

then there exists $T > T_\star$ such that the solution can be extended to $[0, T]$. In particular, blow-up of Euler solutions is possible only if their vorticity diverges to infinity. The BKM theorem was followed by several similar results belonging to the class of blow-up criteria for the Euler equations, and it has now become possible to confront numerical approximations of the solutions to these criteria in order to accumulate some evidence for or against blow-up (see Gibbon, 2008 for a recent review on this topic).

The research on well-posedness of the Euler equations is arguably more crucial for progress on turbulence theory than the one for the NSE[3]. Indeed, the program stemming from the "Onsager conjecture" (see Eyink and Sreenivasan, 2006) aims to describe turbulent flows by weak solutions to the Euler equations containing singularities, and which should have a non-zero energy dissipation rate, in order to agree with experimental evidence (see Sec. I.1.1). Should that goal be achieved, one would have a Re-independent theory of turbulence, an attractive perspective (which is however repeatedly criticized, to say the least, see e.g. Kholmyansky and Tsinober, 2009). The idea of Onsager relied on the notion of Hölder regularity of a solution $\mathbf{u}$, defined by

$$\forall \mathbf{x}, \mathbf{y}, |\mathbf{u}(\mathbf{x}) - \mathbf{u}(\mathbf{y})| < C|\mathbf{x} - \mathbf{y}|^\alpha$$

where C and α are real constants, and α is called the order of regularity. He proved that energy would be conserved for weak solutions with Hölder regularity of order $\alpha > \frac{1}{3}$, and conjectured the existence of dissipative weak solutions with $\alpha \leq \frac{1}{3}$. Up to now, there are two main kinds of results along this line. On the one hand, upper bounds have been derived on the regularity of putative dissipative weak solutions to the Euler equations (Constantin et al., 1994; Cheskidov et al., 2008), by showing that solutions whose regularity is large enough (in some sense more complicated than Hölder regularity) necessarily conserve energy. On the other hand, there have been systematic attempts to construct dissipative weak solutions to the Euler equations (Shvydkoy, 2009, 2010). Despite these considerable advances, this central mathematical challenge related to turbulence theory remains open, both in 2D and in 3D.

I.3.1.3 Vlasov-Poisson equations

Restricting ourselves to the periodic case for simplicity, the initial value problem for the Vlasov-Poisson equations in non-dimensional form is as follows. Given a positive initial condition $f_0(\mathbf{x}, \mathbf{v})$ defined on $\mathbb{T}^d \times \mathbb{R}^d$, and a final time T, find a positive function $f(\mathbf{x}, \mathbf{v}, t)$,

[3] In particular, the occurence of singularities in Euler solutions does not require that the velocity becomes infinite

defined on $\mathbb{T}^d \times \mathbb{R}^d \times [0,T]$, such that $f(\cdot,\cdot,0) = f_0$, satisfying the Vlasov equation (I.1.6), with $m = 1$ and $\mathbf{F} = \boldsymbol{\nabla}\phi$ where ϕ – the electric potential – solves the Poisson equation:

$$-\Delta\phi = 1 - \int_{\mathbb{R}^d} f(\mathbf{x},\mathbf{v},t)\mathrm{d}v, \tag{I.3.1}$$

the constant 1 coming from the neutralizing background of ions, which are supposed to remain immobile in this description.

Existence and uniqueness of classical solutions was obtained under fairly optimal conditions by Lions and Perthame (1991), building on many earlier results. Pfaffelmoser (1992) gave a completely different proof of existence of classical solutions. All this is concisely reviewed in (Bouchut et al., 2000). Another mathematical question concerns the rigorous derivation of the Vlasov equations from the microscopic Newton equations governing the charged particles For a smooth potential this is known (Braun and Hepp, 1977) as we have previously seen, but it is an open problem in the case of the actual Coulomb potential (see, e.g., Hauray and Jabin, 2007). The long time existence seems to be also an open problem when the Landau collision operator is added to the Vlasov-Poisson equations (Villani, 2002). On the other hand, when using the much simpler Fokker-Planck operator, long time existence and uniqueness are granted (Bouchut, 1993).

I.3.2 Dynamical systems and attractors

Following a long period of lethargy[4], the mathematical theory of dynamical systems initiated by Poincaré at the end of the XIX-th century was revived in the second half of the XX-th century following the seminal work of Lorenz (1963) and Smale (1967). It is now being actively developed, in particular in its application to turbulence. A detailed account is to be found in (Foiaš et al., 2001), and the collection of papers by Ruelle (1995) offers a panorama on the subject. A global attractor of a dynamical system such as the forced NSE is a set of states close to which any trajectory ends up being. Associated to it, there may exist invariant measures, which quantify the average fraction of time that the system spends in every subset of the attractor. The geometry of a global attractor and the behavior of an invariant measure, if they exist, contain much information about the long time behavior of a system. The landmark paper of Ruelle and Takens (1971) led to the realization that such attractors were, in most cases of interest, not smooth manifolds but more complicated sets. In particular, their Hausdorff dimension is typically not an integer.

We report some of the rigorous results on this topic here because they are important to understand the present limitations of the numerical approach. In a sense such results aim to give rigorous ground to the qualitative estimates obtained from the K41 and KBL67 theories, that the number of "active degrees of freedom" of a turbulent flow should be proportional to $\mathrm{Re}^{\frac{9}{4}}$ in 3D and to Re^{1} in 2D. However, since the results are formulated in the case of the forced NSE, they are stated in terms of the Grashof number $\mathcal{G}$ which characterizes the amplitude of the forcing after the equations have been non-dimensionalized. A possible definition of $\mathcal{G}$, which we have adopted in the discussion below, is $\mathcal{G} = \mathrm{Re}^2\mathcal{F}$, where $\mathcal{F}$ is the long time average of the energy injected into the flow by the forcing.

[4] see however the historical accounts by Dahan Dalmedico et al. (1992) emphasizing in particular the early contributions of the Russian school

A first category of results are concerned with the minimal number of modes of a solution of the forced NSE that have to be known in order to unambiguously determine its behavior in the limit $t \to \infty$. This approach was started by Foiaš and Prodi (1967). Roughly stated, the subsequently proved theorems tell us that for 2D periodic wall-less flows, the number of determining modes is of order $O(\mathcal{G})$ or less (Jones and Titi, 1993), and for flows in contact with walls, it is $O(\mathcal{G}^2)$ or less (Foiaš et al., 1983). A second category of results relates to properties of the attractor of the forced NSE. A global attractor is a subset of the allowed solution space close to which all the solutions end-up being in the limit $t \to \infty$. The main theorem that concerns us is that the Hausdorff dimension of the global attractor for the 2D NSE is bounded from above by $O(\mathcal{G})$, with or without walls (see Foiaš et al., 2001).

To put this last result in the context of what we shall discuss below it is essential to understand the link between $\mathcal{G}$ and Re. The point is that for the energy of a solution to remain bounded, the forcing characterized by $\mathcal{G}$ has to be balanced by a sufficient amount of energy dissipation. That is, denoting by ε_∞ the non-dimensional long time average energy dissipation rate, we must have $\varepsilon_\infty \sim \mathcal{F}$. Hence if $\varepsilon_\infty = O(1)$, we get that $\mathcal{G} = O(\mathrm{Re}^2)$, whereas if $\varepsilon_\infty = O(\mathrm{Re}^{-1})$ we get that $\mathcal{G} = O(\mathrm{Re})$. Hence roughly stated, the global attractor for the 2D NSE has dimension bounded by $O(\mathrm{Re})$ if $\varepsilon_\infty = O(\mathrm{Re}^{-1})$, and by $O(\mathrm{Re}^2)$ if $\varepsilon_\infty = O(1)$. The reason for singling out these two cases will be explained in the conclusion to this thesis.

One of the hopes behind the study of the attractor for the NSE is to obtain rigorous justifications of some numerical methods used to compute the flow. At the moment, it is still an open question to show in the general case that the solutions to finite dimensional projections of the Navier-Stokes equations accurately approximate the solutions to the full equations uniformly in time. Nevertheless, a heuristic estimate of the required number of modes can be proposed based on the dimension of the attractor. Momentarily denoting by N_A this dimension, and assuming that the problem is isotropic, this can be translated into an estimate for the numerical resolution Δx:

$$\Delta x \propto {N_A}^{\frac{1}{d}}. \tag{I.3.2}$$

I.3.3 Singular limits

The study of the attractor associated to a forced and dissipative PDE gives information about the long time behavior of ensemble of solutions. The features of the attractors are functions of the parameters of the PDE, like the viscosity or the collisionality, and the properties of these functions can tell us something about the various possible regimes for the PDE, as we have seen in the previous paragraph. A complementary approach consists in focusing on a single solution, studying its dependency on the PDE parameters, and possibly introducing some statistical measure at a later stage of the analysis. We emphasize this approach in this section, by recalling some results on turbulent flows seen as functions of their Reynolds numbers, and on filamentary plasma flows seen as functions of their collisionalities. A review on singular limits in fluid dynamics may be found in (Masmoudi, 2007).

I.3.3.1 Vanishing viscosity limit

The mathematical formulation of the vanishing viscosity limit problem for the NSE is as follows: what is the behavior when $\nu \to 0$ of the family of solutions $(\mathbf{u}_\nu)$ of the NSE on a time interval $[0, T]$, all corresponding to the same initial condition, but each one to a different viscosity ν ? Note that by suitable rescaling the family can be equivalently parametrized by Re^{-1}, which is more suited to the comparison with experimental results. The case where the initial condition itself depends on ν is also of interest but we leave it aside for simplicity.

More precisely, one may first wonder in what norms $(\mathbf{u}_\nu)$ is bounded when $\nu \to 0$, and then if it converges in some sense to some limit velocity field $\mathbf{u}$. If this is the case, it is natural to ask whether this limit satisfies the Euler equations, since the latter are obtained by letting $\nu = 0$ in the NSE. We distinguish between the wall-less case, where the fluid is not in contact with any solid, and the opposite wall-bounded case.

Wall-less case. The vanishing viscosity limit of classical solutions to the Navier-Stokes solutions in the absence of walls was studied by Golovkin (1966) in 2D, and by Swann (1971) and Kato (1972) in 3D. It is shown that the Navier-Stokes solutions converges to the smooth Euler solution as long as the latter exists, that is, for all time in 2D and possibly for a finite time in 3D. Note that the smoothness requirements for these results are higher than the minimal smoothness of classical solutions. The convergence rate is $O(\nu)$ in any Sobolev norm compatible with the smoothness (Beale and Majda, 1981; Constantin, 1986). However, the constant in the O may increase very fast in time. This explains why in the 2D case, the rate $O(\nu)$ is observed numerically in the L^2 norm, but is very difficult to observe for higher order Sobolev norms (see Sec. IV.2).

The more abstract case of Leray-Hopf weak solutions was considered by DiPerna and Lions (see Lions, 1996; Bardos and Titi, 2007). They showed that any weak vanishing viscosity limit of a family of such solutions was a so-called "dissipative solution" to the Euler equations. In addition, they proved that as long as a smooth solution to the Euler equations exists, any dissipative solution on the same time interval coincides with it. Therefore the DiPerna and Lions convergence results include the above results concerning classical solutions, however, in the weak setting convergence rates are not available. Bardos and Titi (2007) explore further the general case where the existence of a smooth Euler solution is not assumed. They show that a weak limit $\overline{\mathbf{u}}$ of a sequence $(\mathbf{u}_\nu)$ of Leray-Hopf weak solutions satisfies a modified Euler equation, with additional terms that can heuristically be linked with the turbulence models discussed above. A connection with the statistical energy spectrum of turbulence is also proposed.

The specific case of vortex patches, that is, when the initial vorticity field is the indicator function of a compact set, was explored by Constantin and Wu (1995, 1996). Abidi and Danchin (2004) proved that the convergence rate was $O(\nu^{\frac{3}{4}})$ if the patches have smooth boundaries (see also Masmoudi, 2007). The nice thing with vortex patches is that the initial condition has vorticity in L^∞, so that it is regular enough to fit within the Yudovich existence and uniqueness theory (see Sec. I.3.1.2). The situation is more delicate for vortex sheets, that is, when the initial vorticity field is a measure supported on a regular curve. This last case concerns us since we shall see in the next paragraph and in Sec. V.2 that it is closely related to the existence of a non-vanishing molecular dissipation rate in the vanishing viscosity limit of wall-bounded 2D flows. The problem was briefly discussed by Delort (1991) in the same

paper were he studied the well-posedness issue, and he concluded that the vanishing viscosity limit of solutions to the NSE with positive vortex sheet initial data would be a weak solution to the Euler equations. However it is not clear whether or not such solutions conserve energy.

Remark that the convergence rate $O(\nu)$ proved in the case of smooth solutions has important numerical consequences. Indeed one of the most common ways of approximating solutions to the Euler equations is by computing solutions to the Navier-Stokes equations with very small viscosity. We have seen in Sec. I.3.2 that a rough estimate of the numerical resolution required to resolve these solutions was $\Delta x \propto N_A^{\frac{1}{d}}$, where N_A is the attractor dimension. We have also recalled that, in the case $d = 2$ and when the energy dissipation rate scales like $O(\mathrm{Re}^{-1})$ (i.e., like $O(\nu)$) N_A is bounded from above by Re^1. Overall these results suggest that Δx should be proportional to $\mathrm{Re}^{-1/2}$ (i.e., to $\sqrt{\nu}$). Equivalently, one may say that the minimum viscosity for which the flow can be correctly computed for a given resolution Δx is proportional to Δx^2. That means that the error with respect to the inviscid flow is also proportional to Δx^2, or in other words that a 2D Navier-Stokes solver can be used as an order 2 solver for the 2D Euler equations.

Another practically important field of research concerns other ways of approaching the Euler solution. We have already mentioned some of them in Section I.2.3, where they were introduced as turbulence models. That category includes the hyper-dissipative Navier-Stokes equations (Basdevant et al., 1981), the Euler-α equations (Holm et al., 1998), the Smagorinsky model as well as other LES models, contour dynamics (Dritschel, 1989), and vortex methods (Cottet and Koumoutsakos, 2000). We shall come back to this issue in more detail in Sec. IV.2.

Wall-bounded case. The first formal asymptotic analysis of the vanishing viscosity limit in the presence of walls goes back to Prandtl (1904). It consists in splitting the problem into two parts by introducing a boundary layer of thickness proportional to $\sqrt{\nu}$ along the solid:

$$\Gamma_P = \{\mathbf{x} \in \Omega \mid d(\mathbf{x}, \partial\Omega) < c\sqrt{\nu}\}$$

where c is an arbitrarily chosen constant. Under the appropriate scaling (so called parabolic scaling), the flow is then shown to asymptotically satisfy the Euler equations in $\Omega \backslash \Gamma_P$, and new equations – the Prandtl equations – are derived that are asymptotically satisfied inside Γ_P. In the simplified case where the boundary of Ω is assumed to be horizontal, the Prandtl equations can be written

$$\begin{cases} \partial_t v_1 + v_1 \partial_1 v_1 + v_2 \partial_2 v_1 - \partial_2^2 v_1 = \alpha(x_1) \\ \boldsymbol{\nabla} \cdot \mathbf{v} = 0 \\ v(\cdot, 0) = 0, \quad v(\cdot, X_2) \xrightarrow[X_2 \to \infty]{} \beta(x_1) \end{cases}, \tag{I.3.3}$$

where $v_1(X_1, X_2) = u_1(X_1, \varepsilon X_2)$, $v_2(X_1, X_2) = \varepsilon^{-1} u_2(X_1, \varepsilon X_2)$, $\varepsilon = \sqrt{\nu}$, and the functions α and β depend on the solution outside the boundary layer. Prandtl realized that his approach would break down in some cases. We now would like to state a few rigorous results from the literature about this problem since they are essential for the understanding of the work presented in Section V.2.

The first result we are presenting justifies the Prandtl approach in the restricted setting where the initial data are analytic (i.e. when $\mathbf{u}_0(\mathbf{x})$ can be extended to a holomorphic

function $\mathbf{u}_0(\mathbf{z})$ of a complex vector $\mathbf{z}$, defined for each coordinate of $\mathbf{z}$ close enough to the real axis). Moreover, the domain Ω is assumed to be a half-space. The results are believed to hold more generally when $\partial\Omega$ has analytic regularity. The theorem can be informally stated as follows:

Theorem 1 (Sammartino and Caflisch, 1998a,b)**.** *Suppose that $\Omega \subset \mathbb{R}^d$ is a half-space, and that the initial condition $\mathbf{u}_0$ is analytic. Then there exists a time T_d, independent of ν, such that for $t \in [0, T_d]$,*

$$\|\mathbf{u}_\nu(t) - \mathbf{u}_0(t)\| \underset{\nu\to 0}{=} O(\sqrt{\nu}),$$

where the norm can be, for example, any Sobolev norm. Moreover the restriction of $\mathbf{u}_\nu$ to Γ_P approximates a rescaled solution to the Prandtl equations (I.3.3).

Which means that in the context of the initial value problem, the Prandtl approach is basically valid up to T_d, but may break down at later times. The importance of the analyticity condition is still not fully understood, but it is conjectured to be essential. Recent results indeed support the fact that the Prandtl equations are ill-posed in Sobolev spaces (Gerard-Varet and Dormy, 2009). The next and maybe most important result concerns the link between vanishing viscosity limit and energy dissipation.

Theorem 2 (Kato, 1984)**.** *Assume that there exists a smooth Euler solution $\mathbf{u}(t)$ on $[0, T]$. Then the following assertions are equivalent:*

(i) $\mathbf{u}_\nu(t) \to \mathbf{u}(t)$ *in* $L^2(\Omega)$ *uniformly with respect to* $t \in [0, T]$,

(ii) $\mathbf{u}_\nu(t) \rightharpoonup \mathbf{u}(t)$ *weakly in* $L^2(\Omega)$ *for each* $t \in [0, T]$,

(iii) $\nu \displaystyle\int_0^T \int_\Omega |\boldsymbol{\nabla}\mathbf{u}|^2 \xrightarrow[\nu\to 0]{} 0$,

(iv) $\nu \displaystyle\int_0^T \int_{\Gamma_K(\nu)} |\boldsymbol{\nabla}\mathbf{u}|^2 \xrightarrow[\nu\to 0]{} 0$, *where* $\Gamma_K(\nu) = \{x \in \Omega \mid d(x, \partial\Omega) < c\nu\}$.

The boundary layer $\Gamma_K(\nu)$ is reminiscent of the one which was introduced for physical reasons by Burgers (1923) and which was discussed in Sec. I.1.1. We thus refer to $\Gamma_K(\nu)$ as the Burgers-Kato boundary layer.The quantity $|\boldsymbol{\nabla}\mathbf{u}|^2$ (where $\boldsymbol{\nabla}\mathbf{u}$ is the velocity gradient tensor, not to be confused with $\boldsymbol{\nabla} \cdot \mathbf{u}$, the divergence of $\mathbf{u}$) can be seen as the rate of dissipation of kinetic energy by the flow, even though in physical terms, it would be more adequate to define the latter by $|\mathbf{S}|^2$, where $\mathbf{S}$ is the symmetric part of the tensor $\boldsymbol{\nabla}\mathbf{u}$. In any case, the interpretation of the Kato theorem is that the convergence to the Euler solution in the vanishing viscosity limit is equivalent to the convergence to zero of the energy dissipation in the Burgers-Kato layer on the same time interval. Interpreted in the light of Kato's theorem, the results that we will present in Sec. V.2 suggest that $\mathbf{u}_\nu \nrightarrow \mathbf{u}$ in general. In contrast, for flow satisfying certain symmetries, for example 2D circularly symmetric flows (Lopes Filho, Mazzucato, and Nussenzveig Lopes, 2008; Lopes Filho, Mazzucato, Nussenzveig Lopes, and Taylor, 2008), $\mathbf{u}_\nu \to \mathbf{u}$ can be proved rigorously. Some consequences of the Kato theorem in terms of numerics will be discussed in Sec. I.4.3.1.

In the event that $\mathbf{u}_\nu \nrightarrow \mathbf{u}$, it can still be shown that $\mathbf{u}_\nu$ converges weakly to some limit $\overline{\mathbf{u}}$ (Bardos and Titi, 2007), and one may wonder if $\overline{\mathbf{u}}$ is another weak solution to the Euler

equations, or a solution of some modified equations including additional terms to account for the dissipative effects. This question is completely open, and even some clue about it would be a remarkable breakthrough in the field of mathematical fluid dynamics.

I.3.3.2 Vanishing collisionality limit

Up to now the mathematical community has focused mostly on the opposite limit of high collisionality, for which convergence to Navier-Stokes solutions is now proved. According to Villani (2002), no rigorous result on the approximation of the Vlasov equation in the presence of small collisional effects, modeled for example by the Landau operator, is available. Even the simpler case of the Fokker-Planck operator does not seem to have drawn much mathematical attention.

Fortunately, a few very interesting results related to this issue are to be found in the physics literature. Lenard and Bernstein (1958), Auerbach (1977) and Ng et al. (2006) studied Landau damping in the vanishing collisionality limit of the linearized Vlasov-Poisson equations augmented with simplified collision operators, and showed that entropy production occurs in a localized region of phase space, of thickness proportional to $\nu^{\frac{1}{3}}$. Strikingly, the entropy production rate seems to converge to a value independent of ν when $\nu \to 0$. Hazeltine (1998) studied transport theory, i.e. the derivation of flux-gradient relationships for stationary solutions, in the vanishing collisionality limit under the simplified Krook collision model. studied weakly collisional Landau damping. Krommes and Hu (1994) made a detailed analysis of the "entropy paradox", a concern that had been raised by earlier studies. They arrived at the conclusion that collisions are essential for the eventual establishment of steady states but did not comment on finite time collisional effects.

We now briefly recall the properties of Landau damping which, although it is a purely collisionless effect, could reinforce collisions, as we explain below. The phenomenon was predicted by Landau (1946) using the Vlasov-Poisson equations, and confirmed experimentally in the 1960s, see (Malmberg and Wharton, 1964) and references therein. In the simplified setting where the plasma is excited by a cosine wave, damping manifests itself by an exponential decay of the electric energy $\mathcal{V} = \frac{1}{2}\int_\Omega |\mathbf{E}|^2$. Since the total energy $\mathcal{E} = \mathcal{V} + \frac{1}{2}\int_{\mathbb{R}} \mathrm{d}\mathbf{v} \int_\Omega \mathrm{d}\mathbf{x} |v|^2 f$ is constant for solutions of the Vlasov-Poisson equations, it follows that the kinetic energy $\mathcal{E} - \mathcal{V}$ increases. In this framework, all this happens at constant Boltzmann entropy. We present a numerical example of Landau damping in Section III.2. Note that a recently proved mathematical theorem, the main arguments behind which are summarized in (Mouhot and Villani, 2010), rigorously establishes Landau damping for the nonlinear Vlasov-Poisson equation starting from an initial condition which is very close to a stationary solution satisfying certain stability conditions. In particular, the proof lifts all the doubts on the fact that the mechanism underlying Landau damping is the conversion of the variations of f with respect to $\mathbf{x}$ into fast oscillations of f with respect to $\mathbf{v}$. This raises the questions of how Landau damping is modified in the presence of small collisional effects. Indeed, all collision operators tend to smear out oscillations of f in the $\mathbf{v}$ direction, and this is accompanied by an increase of the Boltzmann entropy. The trend to equilibrium for the Fokker-Planck and Boltzmann problems was studied by Desvillettes and Villani (2001, 2004), but in the framework of neutral gases where Landau damping is absent. The rigorous mathematical study of Landau damping coupled with weak collisional effects is therefore an open topic of great physical interest.

The results from the physical literature quoted above suggest that Landau damping could be the plasma analog to the missing link between turbulence and molecular dissipation in the case of the Euler equations. In the situations that have been studied up to now, the effect of Landau damping on the electric field seems to be robust to weak collisional effects, or in other words the vanishing collisionality limit seems to be dissipative but non-singular. More general cases remain open for exploration.

I.4 Numerical approach

I.4.1 Foundations

Computational fluid dynamics (CFD) aims to take advantage of some information about the present state of a fluid in order to predict its future state (stationary or not), by computing approximate solutions to differential equations that can be derived from physical considerations. Here, we do not come back on the meaning of "predict" which has been covered in Sec. I.2.1, and we would like to focus on the "computing" part.

The historical roots of CFD can be traced back to the very beginning of the XX-th century with Sheppard and Richardson, but its full potential was recognized only after the invention of digital computers. The events of this early period lay somewhat buried under the weight of World War II, and are tied to the Manhattan project and the development of the hydrogen bomb in the 1950s. A very brief and instructive summary has been published in Farge (1988b). Much more information can be found in the complete works of John von Neumann (1963), who is considered as one of the founders of the field. Today CFD has evolved into a multi-disciplinary tool belonging to the wide class of numerical experimentation methods, which is sometimes attributed the status of an independent field of research under the name computational science.

CFD can also be interpreted in a slightly different way, as part of the broader field named "numerical mathematics", i.e. the detailed analysis of numerical approximations to solutions of mathematical problems in order to understand their properties and provide indications of what could maybe be proved or what is likely to be false. This game has been played by mathematicians basically for ever (think of number theory, geometry, etc.), but its rules are now rapidly changing due to the increase in computer power. The technical difficulties to take advantage of this increased power are becoming so high that there is now a tendency towards specialization in numerical mathematics, giving birth to new fields of research like computational number theory, computational geometry, etc. Given the host of mathematically open questions concerning fluid dynamics, of which we have reviewed but a small part above, it seems wholly justified to include CFD inside numerical mathematics. However the overlap between the physical, mathematical, and engineering aspects of CFD is at the same time fruitful and severely confusing. The current theories being still quite incomplete, it is very difficult to dissipate this confusion, and we have not succeeded in doing so, as will appear to the reader in the upcoming chapters.

We now briefly recall the technical ingredients of CFD. The first one is discretization, i.e. the reduction of the problem to a finite number of operations realizable in a finite time on a computer. The study of discretization is one of the goals of the mathematical field known as numerical analysis. The discretization involves two steps: space discretization,

and time discretization. For both steps many options are available, each one having its pros and its cons. Concerning spatial discretization, we shall be mostly concerned with so-called Galerkin approaches, where the equations are projected onto a finite dimensional linear space spanned by a set of orthogonal functions. The Fourier-Galerkin approach for the Euler equations is described in some detail in Sec. IV.3, and the Wavelet-Galerkin discretization for the Poisson equation is presented in Sec. III.2. Time discretization is commonly done using time marching algorithms, that "advance" the solution in time in a causal fashion. We shall make use of two such algorithms, third order Runge-Kutta and Verlet. In addition to accuracy, an essential property to consider in the case of time marching algorithms is their stability. Indeed, even a very accurate approximation breaks down sooner or later if the error grows exponentially in time. For the explicit time discretization schemes that we are using, stability imposes a restriction on the time step δt which is dictated by the spatial resolution. In the case of linear transport equations, of the type

$$\partial_t \theta + \mathbf{u} \cdot \boldsymbol{\nabla} \theta = 0,$$

the stability criterion was rigorously obtained by Courant, Friedrichs, and Lewy (1928) and is now referred to as the CFL condition.

I.4.2 Technological requirements

The persistent increase in the power of computers has been an important factor in the development of CFD since its infancy. In this section we briefly mention how it is possible to take advantage of this technological progress to tackle new problems.

For technical and commercial reasons, the potential of recent computers can only be fully exploited if many central processing units (CPUs) are simultaneously utilized during the computation. Most of the time, this requires that they communicate through a network, and these communications take time. The more CPUs participate to the resolution of a given computational problem, the more communications are necessary. The dependency of the total time required to finish it on the number of CPUs is called the "scaling" of the method. A commonly used measure of scaling is the parallelization efficiency, which is the ratio between the CPU time consumed for the same task before and after parallelization. New algorithms, and sometimes even new programming methods, are needed to obtain a good scaling. This has given rise to yet another emerging field of study, namely high performance computing (HPC). We have chosen to include in the body of this thesis some technical points that we have developed concerning the parallel wavelet transform, and that are essential to the viability of our approach in the long term. These may be found in Sec. II.2. Note that recently, the use of graphics processing units (GPUs) as a complement to large parallel codes has been achieved with some degree of success, but the viability of the approach is not clear on the long term, and we have therefore chosen not to engage in it.

We feel that the often overlooked matter of software engineering also deserves to be mentioned here. In the last few decades the establishment of a global software industry has stimulated the invention of modern programming techniques that aim to improve efficiency in terms of development, verification, debugging and code reuse. The paradigm of object oriented programming, embodied by languages such as Java and C++, is at the heart of these efforts. Although this evolution has been up to now mostly overlooked by the CFD community, other scientific domains already take advantage of it. We have chosen

to work with the C/C++ language, mostly because it benefits from a much wider user community than FORTRAN, but also because generic programming techniques have allowed us to implement very general wavelet transform algorithms that can now be adapted to many contexts, a feature that would have been much more difficult to obtain using FORTRAN. Moreover, for complicated problems involving parallelization, it is essential to take advantage of libraries that have been well documented and well tested by the community. Concerning for example the Fourier transform algorithms that we make heavy use of, a huge effort has been spent by a team at the Massachusetts Institute of Technology (Frigo and Johnson, 2005) to develop the "fastest Fourier transform in the west" (FFTW). This library relies on advanced software engineering techniques to select a way of computing the Fourier transform which is actually optimal for the specific computer, and for the specific configuration at the time when the program is being run. Input/output operations in a parallel environment are also a very delicate thing to do, and it has been essential for us to use the HDF5 library (developed by a consortium of US government agencies), which provides a portable and efficient data storage format. The analysis of the huge amount of data resulting from HPC will probably require even more advanced storage interfaces in the future. FORTRAN interfaces are also available for these libraries, but they may become unsupported in the future.

I.4.3 Examples of current achievements

I.4.3.1 Turbulent wall bounded flows

It is instructive to briefly describe some state of the art computer experiments of turbulent wall-bounded flows in contact with solids found in the CFD literature.We also take this opportunity to introduce the classical concepts of wall turbulence that are essential to follow the discussion. The basic theory and vocabulary can be found for example in the textbook (Tennekes and Lumley, 1972), but we utilize mostly the recent review article by Marusic et al. (2010), and the clear statements of open problems by McKeon and Sharma (2010).

To understand the scope of the problem, some background must first be given. The study of wall-bounded flows has focused on three canonical configurations, the boundary layer, the pipe flow and the channel flow, and we focus here on the channel flow. Strictly speaking, one calls channel flow the flow between two parallel plates extending indefinitely in the directions labeled x and z which occurs when there a constant pressure gradient is applied in the x direction. x,y and z are respectively called the stream-wise, the wall-normal, and the span-wise directions. The walls are assumed to be located at $y = 0$ and $y = 2h$. The classical theory has focused mostly on the description of the mean flow U in the stream-wise direction, which by symmetry depends only on y. By averaging the Navier-Stokes equations, U can be shown to satisfy the equation

$$\nu \frac{\mathrm{d}U}{\mathrm{d}y} = \overline{uv} + U_\tau^2 \left(1 - \frac{y}{h}\right), \tag{I.4.1}$$

where u and v are the components of the velocity fluctuations respectively in the x and y directions, $\bar{\cdot}$ denotes averaging along x, z and time, and $U_\tau = \sqrt{\nu \left.\frac{\mathrm{d}U}{\mathrm{d}y}\right|_{y=0}}$ is called the friction velocity. The term $\overline{uv}$ is called diagonal Reynolds stress. Since the work of von Kármán (1921), the classical approach has been to look for a solution of this equation under

the form

$$U = U_\tau f\left(\frac{yU_\tau}{\nu}\right), \tag{I.4.2}$$

where f is a universal function. In various ways, dimensional arguments and (I.4.1) can be used to predict the existence of a region, called the log-layer, where

$$f(y^+) \approx \frac{1}{\kappa}\ln(y^+) + B, \tag{I.4.3}$$

with κ a possibly universal constant called the Kármán constant, and B a geometry dependent constant. The experimental support for this attractively simple and general "law of the wall" is at the same time overwhelming and shaky. Indeed, a log-layer is observed in virtually all high Reynolds number experiments, and even in some of the most recent numerical simulations (see below), but only for a restricted range of y^+. Moreover, the value of the "universal" constant κ is not agreed upon. Due to these difficulties, many alternative theories have been proposed, some of which attempt to describe f using power laws instead of a logarithm. From this brief discussion, it appears that the theory of wall bounded flows is very incomplete even concerning the average velocity, leaving aside the matter of fluctuations which are even less understood.

However, one particular feature of the existing theory that is very interesting to us has to do with its scaling at high Reynolds number Re, where Re is defined by $hU(h)/\nu$. We would like to make a digression about this here, although it is slightly out of place. We assume that (I.4.3) still holds as a very rough order of magnitude when $y = h$, i.e. at the center of the channel. Then from (I.4.2), and due to the logarithmic dependence of f on y^+ (I.4.3), it appears that the ratio between $U(h)$ and U_τ can have only a logarithmic dependence on the Reynolds number. By definition of U_τ, this implies that in non-dimensional units:

$$\left.\frac{\mathrm{d}U}{\mathrm{d}y}\right|_{y=0} \propto \mathrm{Re}$$

up to logarithmic factors. By a naive gradient length argument, this scaling can be seen as the statistical signature of the existence of a boundary layer of thickness Re^{-1} in the neighborhood of the wall. Hence we see that the log-law, as an experimental result, is consistent with the model of Burgers (1923) that was mentioned earlier, and also with Kato's theorem. This connection can be made, as we just did, in a purely phenomenological way without invoking the Kolmogorov scale and the notion of cascade. In fact the essential point is that the stream-wise velocity everywhere in the channel scales with U_τ, and this scaling hypothesis was introduced by von Kármán (1921) precisely to account for the behavior of the drag coefficient at high Reynolds number, which was the essential issue at the time. From this discussion is appears that the Kato theorem is a good starting point to start investigating rigorous foundations of the phenomenological Kármán theory.

After this digression, we now review some CFD techniques that are currently being applied to channel flows. Numerical results on the channel flow problem were recently reviewed by Jiménez and Moser (2007). Numerically, periodic boundary conditions with a sufficiently large period are used in the x and z directions. The standard approach to the problem is strongly influenced by the work reported by Kim et al. (1987), who worked at a Reynolds number 3300 (based on the channel width and mean velocity), and used a resolution of $192 \times 129 \times 160$. The unknown velocity field is expanded onto a mixed basis constructed from Fourier modes in the x and z directions, and Chebychev polynomials in the y direction.

Moser and Moin (1984) gave more detail on the numerical methods, and concluded that "The energy spectra [...] illustrate that the grid resolution is adequate, since the energy density associated with the high wave-numbers is several decades lower than the energy density corresponding to low wave-numbers, and there is no evidence of energy pile-up at high wave-numbers. It should be noted, however, that the drop-off of the computed spectra of high wave-numbers is not sufficient evidence that the computed results are unaffected by the small-scale motions neglected in the computations. It is not clear what significant dynamical roles, if any, these small scales would play if included in the computations. Numerical experiments with much finer resolutions than those used here would presumably clarify this issue. Such computations are very difficult and time consuming to carry out with the present computers." Today, impressive results are reported that use much higher resolutions.For example, del Álamo et al. (2004) and Hoyas and Jiménez (2006) reported computations with respectively $768 \times 769 \times 768$ and $6144 \times 633 \times 4608$ grid points. But they also work with much larger values of Re than in previous studies. Hence convergence studies at a fixed Re are still difficult and most of the time omitted in the reported results. The choice of numerical resolution is normally justified by the comparison of the grid spacing with the Kolmogorov length, which scales like $\mathrm{Re}^{-\frac{3}{4}}$. For example, in (Hoyas and Jiménez, 2006), the condition that $\Delta y = 1.5\eta$, where Δy is the wall-normal grid spacing, is considered sufficient to ensure that the flow is well resolved.

From this discussion, we have the feeling that the main factor limiting the ability of CFD to accurately simulate high Reynolds number wall-bounded flows is computer power. Due to the rapid development of parallel computing techniques, we expect that much larger Reynolds numbers will be achievable in the next few decades. To extract useful information from these upcoming simulations, it will be necessary to develop original analysis techniques. The wavelet analysis of a turbulent boundary layer reported in Sec. V.4 is a first step in this direction.

I.4.3.2 Plasma numerics

To put the results of Chapter 3 in perspective, we briefly review the computational methods used in the context of hot plasmas. Indeed the applications of plasma simulation are numerous. The most prominent ones include laser-plasma acceleration and controlled fusion.

Historically, the first methods were developed at Stanford by the group of Oscar Buneman, and gave birth to a whole school of plasma simulation using particles (Birdsall and Langdon, 1985; Hockney and Eastwood, 1988) that has now migrated to a number of adjacent fields of research. We refer to their approach as Lagrangian, because its main ingredient consists in following discrete marker particles described by so-called characteristic equations, whose trajectories are such that the particle distribution function is constant along them. Nowadays two other types of numerical methods for the Vlasov equations have appeared, namely semi-Lagrangian and Eulerian. In the semi-Lagrangian discretization, the particle distribution function is assumed to be sufficiently smooth and is represented by its values over a grid, but its time evolution is approximated by integrating the characteristic equations backwards in time starting from the grid-points, and using the conservation property. In Eulerian solvers, the Vlasov equation is interpreted as a hyperbolic conservation law which can be discretized using for example a finite volume scheme. Due to the high dimensionality of the problem (6 dimensions in general, 3 in space and 3 in velocity), the full discretization over an Eulerian

grid can be very costly and even prohibitive in some cases. The advantage of the Lagrangian schemes is that a grid is only required to represent the electromagnetic fields which depend only on the space variables. We come back to this point in more detail in Chapter 3.

To describe the core plasma of a magnetic fusion device, the equations currently used are called the gyrokinetic equations. They are an approximation of the Vlasov equations, obtained by "gyro-averaging", i.e. averaging out the fast rotation of the particles perpendicular to the magnetic field lines (Brizard and Hahm, 2007). Since the gyrokinetic equations are mathematically analogous to the Vlasov-Maxwell equations, they can be discretized by very similar numerical methods, i.e. Lagrangian, semi-Lagrangian, and Eulerian ones. A recent review of the various state-of-the-art methods and of the results obtained up to now can be found in (Batchelor et al., 2007). The demands in terms of computational power, memory and data storage to simulate a large tokamak like ITER are currently very high. As an example we consider ITER-like simulations using the GYSELA code developed at the french atomic energy commission (CEA) in Cadarache, at Université de Nancy and at Université de Strasbourg (Grandgirard et al., 2006). An important GYSELA run performed in 2010, on the occasion of a challenge to test a newly installed computer, took about 30 days on 8192 CPUs, i.e. more than $6.7 \cdot 10^6$ hours of CPU time. The discretization of the particle distribution function required $272 \cdot 10^9$ grid points, and 27.2 Gb of memory per node. The data accumulated as output of this run occupies a total of 6 terabytes (10^{12} bytes) on hard drives, of which 4.7 consist simply in restart files. Some important effects, in particular the kinetic response of the electrons in the plasma, have been neglected in this simulation. Taking them into account will imply a further increase in computational requirements. This example shows that, as in fluid turbulence, the computational complexity is and will be part of the limiting factors in our ability to predict the behavior of magnetically confined plasmas. Further research on fundamental statistical properties of plasma flows could help reduce this complexity, and also facilitate the extraction of useful information out of the huge quantity of data resulting from such simulations.

Part II

Wavelet tools for flow analysis

> Il serait bien à désirer que les ingénieurs, et en général tous les hommes qui ont à mesurer et à calculer, prissent définitivement et sans retour le parti de rapporter au mètre seul toutes les mesures de dimension ; la diversité des unités linéaires, superficielles et cubiques qu'on emploie, souvent dans la même phrase, forme une bigarrure choquante, et il est surtout étonnant qu'on se serve encore du *pouce de fontainier*.
>
> Gaspard-Clair-François-Marie Riche de Prony
> *Mémoire sur le jaugeage des eaux courantes* (1802)

Summary

II.1 Mathematical theory

Wavelet theory has developed into a whole branch of harmonic analysis ripe with interesting concepts and theorems. In this section, we introduce only the mathematical definitions and results that are essential to the understanding of the following chapters. Well established textbooks on the subject include (Daubechies, 1992; Mallat, 1999). The following presentation focuses on discrete wavelet families, and the topic of the continuous wavelet transform is left aside. Most of the material is classical, with a few exceptions, namely:

- the presentation of the adaptive wavelet transform (Sec. II.1.2.3), which is original but closely related to previous approaches,
- the explicit generalization of the Kingsbury complex wavelet transform to arbitrary space dimension (Sec. II.1.1.3).

The way the wavelets are indexed has been given some particular attention, and although it may seem rather esoteric at first sight, we hope that its overall effect is positive.

II.1.1 Multiresolution analysis

II.1.1.1 Wavelets

The construction first takes place in the Hilbert space $L^2(\mathbb{R}^d)$, where the scalar product is denoted by $\langle \cdot \mid \cdot \rangle$ and the norm by $\| \cdot \|$. We start with two functions φ and ψ in $L^2(\mathbb{R})$ such that $\|\varphi\| = \|\psi\| = 1$, which satisfy for $t \in \mathbb{R}$ the two scale refinement equations:

$$\varphi(t) = \sqrt{2} \sum_{i=0}^{S-1} h_i \varphi(2t - i) \tag{II.1.1}$$

$$\psi(t) = \sqrt{2} \sum_{i=0}^{S-1} g_i \varphi(2t - i + S - 2), \tag{II.1.2}$$

where $(h_i)_{0 \le i \le S-1}$, $(g_i)_{0 \le i \le S-1}$ are finite sequences of real numbers linked by

$$g_i = (-1)^{i+1} h_{S-1-i}. \tag{II.1.3}$$

and such that the discrete Fourier transform of h, defined by

$$\hat{h}(\xi) = \sum_{n \in \mathbb{Z}} h_n e^{\imath n \xi},$$

satisfies the quadrature property

$$\forall \xi, |\hat{h}(\xi)|^2 + |\hat{h}(\xi + \pi)|^2 = 2. \tag{II.1.4}$$

φ and ψ are respectively called the scaling function and the wavelet, and h and g are the associated quadrature mirror filters. For an easier correspondence with the numerical implementation, we have adopted the convention that g and h are both indexed by positive integers, which is different from the convention in (Mallat, 1999), hence the shift by $(S-2)$

in (II.1.2). We assume that (II.1.1-II.1.2) have well behaved solutions, whose supports are contained in $[0, S-1]$, see (Daubechies and Lagarias, 1991, 1992).

An important property of ψ is that its low order moments vanish

$$\int_{\mathbb{R}} x^m \psi(x) = 0 \text{ for } 0 \leq m \leq M-1, \tag{II.1.5}$$

where M is at least 1 for the lowest order wavelet family, called the Haar family, and in practice $M > 1$ for other commonly used families that we shall encounter. Regularity is a less important but nevertheless interesting matter concerning wavelets (Daubechies and Lagarias, 1991, 1992). Compactly supported wavelets are not infinitely differentiable functions. In fact, as we shall see below, the most common ones have relatively low order smoothness.

II.1.1.2 Orthogonal wavelet bases

In this presentation we directly define the wavelet basis in d dimensions, where $d \in \mathbb{N}^*$. Simpler notations for the wavelet transform in 1 or 2 dimensions may be found below in the introductions to the more specialized parts of this thesis. We start by defining a wavelet basis on $\mathbb{R}^d$, which we will have to periodize afterwards. The wavelets will be indexed by a parameter

$$\lambda \in \Lambda_A = \mathbb{Z} \times \{0,1\}^d \times \mathbb{Z}^d, \tag{II.1.6}$$

where

- the first factor indicates the scale the wavelet, and is denoted by j^λ,
- the second factor indicates the directions of oscillation of the wavelet, and is denoted by $\boldsymbol{\mu}^\lambda \in \{0,1\}^d$,
- the third factor indicates the position of the wavelet, and is denoted by a vector with integer components $\mathbf{i}^\lambda \in \mathbb{Z}^d$.

We shall thus always have $\lambda = (j^\lambda, \boldsymbol{\mu}^\lambda, \mathbf{i}^\lambda)$. Now for $\mathbf{x} = (x_k)_{1 \leq k \leq d} \in \mathbb{R}^d$, let

$$\psi_\lambda(\mathbf{x}) := 2^{\frac{dj^\lambda}{2}} \prod_{k, \mu_k^\lambda = 1} \psi(2^{j^\lambda} x_k - i_k^\lambda) \prod_{k, \mu_k^\lambda = 0} \varphi(2^{j^\lambda} x_k - i_k^\lambda). \tag{II.1.7}$$

For example, if $d = 2$, $\boldsymbol{\mu}^\lambda$ can take the four values $(0,0)$, $(1,0)$, $(0,1)$ and $(1,1)$. For the first value, the scaling function is obtained, and for the three remaining values, the three wavelets oscillating respectively in the x_1 direction, in the x_2 direction, and in both directions are obtained. In general it is convenient to consider the vector $\boldsymbol{\mu}^\lambda \in \{0,1\}^d$ as a binary expansion of the more classical direction index μ, which can take 3 nonzero values in 2 dimensions, 7 in 3 dimensions, and $2^d - 1$ in d dimensions. Note that we have adopted the convention that the scale index j_λ increases when going from coarse to fine scales. Denoting

$$\Lambda_\varphi = \left\{ \lambda \,\middle|\, \boldsymbol{\mu}^\lambda = 0 \right\} \tag{II.1.8}$$

$$\Lambda_\psi = \left\{ \lambda \,\middle|\, \boldsymbol{\mu}^\lambda \neq 0 \right\}, \tag{II.1.9}$$

we say that for $\lambda \in \Lambda_\varphi$, the function ψ is a scaling function, and in that case we denote $\varphi_\lambda = \psi_\lambda$. For $\lambda \in \Lambda_\psi$, ψ is called a wavelet. In this manner a "unified notation" designs at the same time scaling functions and wavelets. For clarity and consistency with classical notations, we nevertheless use a different notation for φ_λ and, except otherwise noticed, when writing ψ_λ we implicitly assume that $\boldsymbol{\mu}^\lambda \neq 0$, i.e. that ψ_λ is a wavelet and not a scaling function. When $\boldsymbol{\mu}^\lambda = 0$ we shall sometimes use the abbreviated notation $(j^\lambda, \mathbf{i}^\lambda)$ for λ. The main difference between wavelets and scaling functions is that the former have zero mean, whereas the latter do not.

The main property of $(\psi_\lambda)_{\Lambda_A}$, as can be deduced from the properties of ψ and φ that we have seen in the previous paragraph, is that it is an orthogonal basis of $L^2(\mathbb{R}^d)$. Hence $L^2(\mathbb{R}^d)$ is spanned by dilates and translates of a few elementary functions. A set of points of the form $2^{-j}\mathbf{i}$, where $j \in \mathbb{Z}$ and $\mathbf{i} \in \mathbb{Z}^d$, is called a dyadic grid. It is therefore commonly said that wavelets "live on a dyadic grid". Now defining

$$\Lambda_{\varphi,j} = \left\{ \lambda \,\middle|\, j^\lambda = j, \boldsymbol{\mu}^\lambda = 0 \right\}$$

for $j \in \mathbb{Z}$, we denote by V_j the space spanned by $(\varphi_\lambda)_{\lambda \in \Lambda_{\varphi,j}}$. Because of (II.1.2), $V_j \subset V_{j+1}$ for all j, so that we may also define W_j, the orthogonal complement of V_j in V_{j+1}. For any $J \in \mathbb{Z}$, $L^2(\mathbb{R}^d)$ can thus be decomposed in the following way, called a multi-resolution analysis:

$$L^2(\mathbb{R}^d) = V_J \oplus \bigoplus_{j \geq J} W_j, \tag{II.1.10}$$

and by construction, each W_j is spanned by the $(\psi_\lambda)_{\lambda \in \Lambda_{j,\psi}}$, where

$$\Lambda_{j,\psi} = \left\{ \lambda \,\middle|\, j^\lambda = j, \boldsymbol{\mu}^\lambda \neq 0 \right\}.$$

For every function $f \in L^2(\mathbb{R})$, V_J contains a coarse scale approximation of f, while the W_j, $j \geq J$ contain the details necessary to reconstruct f entirely. This information is encoded in the scaling coefficients and wavelet coefficients of f:

$$\overline{f}_\lambda = \langle f \mid \varphi_\lambda \rangle \tag{II.1.11}$$
$$\widetilde{f}_\lambda = \langle f \mid \psi_\lambda \rangle, \tag{II.1.12}$$

and it can be used to reconstruct f using its wavelet decomposition:

$$f = \sum_{\lambda \in \Lambda_{\varphi,J}} \overline{f}_\lambda \varphi_\lambda + \sum_{j \geq J} \sum_{\lambda \in \Lambda_{\psi,j}} \widetilde{f}_\lambda \psi_\lambda. \tag{II.1.13}$$

As for the wavelets, the notations for the coefficients can be unified since (II.1.11) can be seen as the special case of (II.1.12) corresponding to $\boldsymbol{\mu}^\lambda = 0$. But we prefer to distinguish $\overline{f}_\lambda$ and $\widetilde{f}_\lambda$, for the same reasons as before and also because the notation $\overline{\cdot}$ is evocative of averaging, while the notation $\widetilde{\cdot}$ is evocative of oscillation.

For numerical applications, we are usually not interested in wavelets defined on $\mathbb{R}$, but instead on wavelets defined on a torus

$$\mathbb{T}[\mathbf{m}] = \bigotimes_{k=1}^{d} \frac{\mathbb{R}}{m_k \mathbb{Z}},$$

where the m_k are integer periods. Most of the time, it is sufficient to work with $\mathbb{T}^d = \mathbb{T}[1]^d$, but for some anisotropic problems it is important to cover the general case for $\mathbf{m}$. To obtain periodized wavelets, one effectively replaces each ψ_λ by

$$\psi_\lambda(\mathbf{x}) \rightarrow \sum_{\mathbf{i}\in\mathbb{Z}^d} \psi_\lambda(\mathbf{x} + \mathbf{m} \otimes \mathbf{i}). \tag{II.1.14}$$

From now on, we are concerned only with the periodized case, but we pursue with the same notations for simplicity. It is important to notice that the sum in (II.1.14) contains more than one term if and only if the size of the support of ψ_λ is larger than m_k for a certain k. Since the support of ψ has length $S-1$, the one of ψ_λ has length $2^{-j^\lambda}(S-1)$. Therefore, a sufficient condition such that the size of the support of ψ_λ is smaller than m_k is

$$j^\lambda \geq \log_2(S-1) - \log_2(m_k). \tag{II.1.15}$$

We shall refer to (II.1.15) as the non-overlap condition. When J in (II.1.10) is chosen such that $J \geq \log_2(S-1) - \log_2(m_k)$, we see that the wavelet coefficients of a function $f \in L^2(\mathbb{T}[\mathbf{m}])$ in the periodized basis are identical to the coefficients of the periodic continuation of f to $\mathbb{R}^d$. Another important point is that periodized wavelets satisfying the non-overlap condition have the same number of vanishing moments as the mother wavelet. In contrast, the moments of coarse scale periodized wavelet that fail to satisfy the non-overlap condition others than the mean do not have vanish in general. Note that the wavelet coefficients in the periodized basis are also periodic themselves, so that the position index $\mathbf{i}^\lambda$ can be restricted to a finite subset of $\mathbb{Z}^d$ depending on j^λ, and that the definition (II.1.6) of Λ_A becomes:

$$\Lambda_A = \left\{ (j, \boldsymbol{\mu}, \mathbf{i}) \,\middle|\, j \in \mathbb{N}, \boldsymbol{\mu} \in \{0,1\}^d, \mathbf{i} \in \left\{0, 1, \ldots, m_k 2^j - 1\right\}^d \right\}.$$

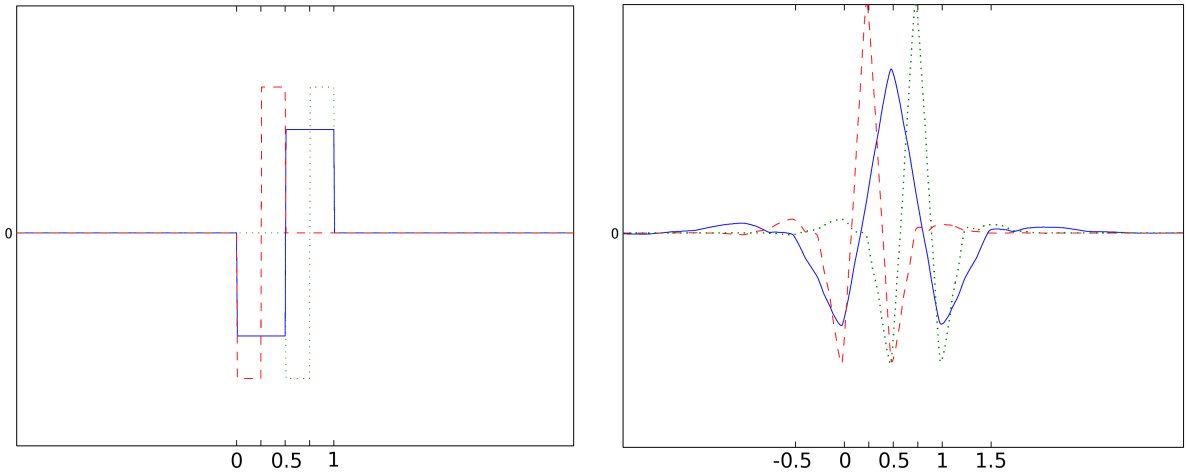
In the same manner, $\Lambda_{\varphi,j}$ and $\lambda_{\psi,j}$ are restricted so that they both become finite. In the following, position indices at scale j always have to be understood modulo $m_k 2^j$. To end this section two examples of wavelets in 1D are shown in Fig. II.1.1.

II.1.1.3 Dual tree complex wavelet transform

Sampling on a dyadic grid as in (II.1.7) is necessary to avoid redundancies in the representation of a signal, but it also introduces a lack of translation invariance that can be a severe drawback for some applications. Taking as example a simple set of 1D Heaviside signals (Fig. II.1.2), it appears that the position of the wavelet artificially imposed by the choice of the origin when constructing the basis entails a very different distribution of the information between all the scales for the shifted versions of the signal. This is a problem for many applications, among which one is the estimation of the variance of stochastic processes, as detailed in Sec. II.3.3.

To circumvent this issue, Kingsbury has been constructing since 1998 complex-valued wavelets families that are not bases but *tight frames* of $L^2(\mathbb{R}^d)$ (Kingsbury, 2000, 1999, 1998). As defined in (Daubechies et al., 1986), a tight frame $(\psi_\lambda)_{\lambda\in\Lambda}$ is characterized by the fact that for any $f \in L^2(\mathbb{R}^d)$:

$$\|f\|_2^2 = A \sum_{\lambda\in\Lambda} |\langle f \mid \psi_\lambda \rangle|^2 \tag{II.1.16}$$



(a) Haar wavelets (b) Coiflets of order 2

Figure II.1.1: For the most simple wavelet family, called the Haar family (a), and which was known well before the birth of wavelet theory proper, it is easily seen that the a wavelet (blue line), its contraction by a factor two (red dashed line), and its neighbor on the dyadic grid are orthogonal to each other. The Coiflet family of order 2 (b) enjoys the same properties; but its construction is much more delicate (see Sec.II.1.3.2).

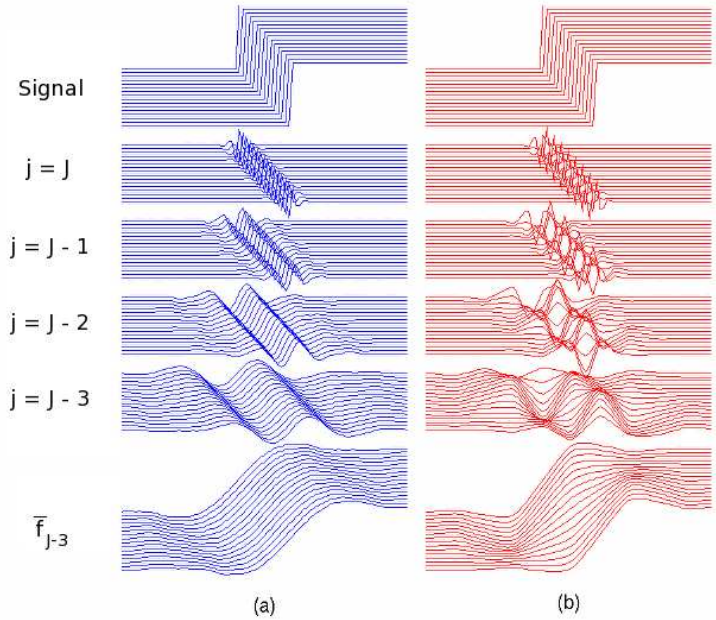


Figure II.1.2: Comparison between the reconstructions of sixteen shifts of a Heaviside function, from their wavelet coefficients at various scales, using either a frame of nearly translation invariant complex-valued wavelets (a), or an orthogonal wavelet basis (b). Figure reproduced by courtesy of N. Kingsbury.

where A is a constant. It can be shown that (II.1.16) implies a reconstruction formula that is nearly identical to that for an orthogonal family:

$$f = \frac{1}{A}\sum_{\lambda\in\Lambda}\langle\psi_\lambda \mid f\rangle\psi_\lambda \tag{II.1.17}$$

Since the elements of a tight frame are not linearly independent, the expansion of a function involves some redundancy. This redundancy was exploited by Kingsbury to reduce the problems associated to translations. We now briefly describe the construction in the d dimensional setting.

The basic ingredients are two mother wavelets ψ^R and ψ^I and their scaling functions φ^R and φ^I, associated to two pairs of quadrature mirror filters, which we denote respectively by (h^R, g^R) and (h^I, g^I) where the exponents R and I stand respectively for real and imaginary. The filters h^R and h^I are tailored in a specific way so that the complex-valued wavelet $\psi = \frac{1}{\sqrt{2}}(\psi^R + \iota\psi^I)$ and scaling function $\varphi = \frac{1}{\sqrt{2}}(\varphi^R + \iota\varphi^I)$ (where $\iota = \sqrt{-1}$) both have their Fourier transform nearly vanishing for negative wave-numbers:

$$\forall\xi < 0, \hat{\psi}(\xi) \simeq 0. \tag{II.1.18}$$

Since general functions have nonzero energy at negative wave-numbers, we see that we will need to use both ψ and ψ^* to represent them, where $\cdot^*$ denotes complex conjugation. To cope with this, we add a new degree of freedom to the index λ compared to orthogonal wavelets. Denoting by $\boldsymbol{\gamma}^\lambda \in \{0,1\}^d$ this new parameter, the definition of complex-valued wavelets goes like

$$\psi_\lambda(\mathbf{x}) := 2^{\frac{dj^\lambda}{2}} \prod_{\substack{k,\gamma_k^\lambda=0\\ \mu_k^\lambda=1}} \psi(2^{j^\lambda}x_k - i_k^\lambda) \prod_{\substack{k,\gamma_k^\lambda=0\\ \mu_k^\lambda=0}} \varphi(2^{j^\lambda}x_k - i_k^\lambda) \prod_{\substack{k,\gamma_k^\lambda=1\\ \mu_k^\lambda=1}} \psi^*(2^{j^\lambda}x_k - i_k^\lambda) \prod_{\substack{k,\gamma_k^\lambda=1\\ \mu_k^\lambda=0}} \varphi^*(2^{j^\lambda}x_k - i_k^\lambda), \tag{II.1.19}$$

From (II.1.19), we see that $\boldsymbol{\gamma}$ characterizes the part of Fourier space to which the support of the wavelet is approximately restricted due to (II.1.18), i. e. $\{\boldsymbol{\xi} \mid (\xi_k \geq 0 \text{ for } \gamma_k = 0) \text{ and } (\xi_k < 0 \text{ for } \gamma_k = 1)\}$. Thanks to this property, Kingsbury wavelets are more directionally selective than orthogonal wavelets: to each value of $(\boldsymbol{\mu}, \boldsymbol{\gamma})$ (with $\boldsymbol{\mu} \neq 0$) corresponds a different direction of oscillation, which makes a total of $(2^d - 1) \times 2^d$ directions. Note that since $\boldsymbol{\gamma}^\lambda \in \{0,1\}^d$, the family contains 2^d times more wavelets than an orthogonal wavelet family, which will imply an additional cost to compute all the coefficients.

The periodization and the definition of wavelet coefficients is analogous to what we have seen for orthogonal wavelets in the previous paragraph, and the reconstruction formula is given by (II.1.17) with $A = 2^d$, which can also be written

$$f = 2^{-d}\left(\sum_{\lambda\in\Lambda_{\varphi,J}} \overline{f}_\lambda\varphi_\lambda + \sum_{j\geq J}\sum_{\lambda\in\Lambda_{\psi,j}} \widetilde{f}_\lambda\psi_\lambda\right),$$

where the index sets $\Lambda_{\varphi,J}$ and $\Lambda_{\psi,j}$ have been modified to take into account $\boldsymbol{\gamma}$.

The transformation from f to its coefficients in the basis of complex-valued wavelets is called dual tree complex wavelet transform (DTCWT) (Kingsbury, 2001). In the following we use the shorthand DTCWT rather loosely to refer either to wavelets, to coefficients, or to algorithms related to the dual tree complex wavelet transform.

II.1.2 Fast wavelet transform

II.1.2.1 Non-adaptive transform

In this section we briefly recall the fast wavelet transform algorithm (see (Mallat, 1999) and references therein). The transform is first assumed to be real and orthogonal, and we explain the generalization to the DTCWT case afterwards. The 1D fast wavelet transform algorithm (Mallat, 1999) is obtained by injecting the refinement equations (II.1.1-II.1.2) into the expansion (II.1.13) of a function in the wavelet basis. The following relations between the coefficients at successive scales are thus derived:

$$\overline{f}_\lambda = \sum_{n_1=0}^{S-1} h_{n_1} \dots \sum_{n_d=0}^{S-1} h_{n_d} \overline{f}_{(j_\lambda+1, 2\mathbf{i}^\lambda+\mathbf{n})}, \tag{II.1.20}$$

$$\widetilde{f}_\lambda = \sum_{n_1=0}^{S-1} g_{n_1}^{\mu_1^\lambda} h_{n_1}^{1-\mu_1^\lambda} \dots \sum_{n_d=0}^{S-1} g_{n_d}^{\mu_d^\lambda} h_{n_d}^{1-\mu_d^\lambda} \overline{f}_{(j_\lambda+1, 2\mathbf{i}^\lambda+\mathbf{n}-(S-2)\boldsymbol{\mu}^\lambda)}, \tag{II.1.21}$$

so that the wavelet and scaling coefficients at scale j^λ can be computed from the scaling coefficients at scale $j^\lambda+1$ by applying (II.1.20-II.1.21) with the position $\mathbf{i}^\lambda$ varying successively along each direction. The (μ_k^λ) appearing as exponents in (II.1.21) simply mean that in each direction, either the filter g or the filter h has to be applied, depending on the direction being considered. A great advantage of the fast wavelet transform is that only one dimensional filters have to be applied. One can also remark that, in the unified notation, (II.1.20) can be seen as the special case of (II.1.21) corresponding to $\boldsymbol{\mu}^\lambda = 0$.

The shift of the index by the quantity $(S-2)\boldsymbol{\mu}^\lambda$ in (II.1.21) needs some clarification. This comes from the way the filter g is defined by (II.1.3), with only positive indices. Thanks to this shift, (II.1.21) holds for the standard definition of the wavelet coefficients, as used for example by (Mallat, 1999), despite the different definition of g. As we shall see below this has important practical consequences on the numerical implementation.

Now assuming that in the beginning the scaling coefficients at a certain scale J are known, the wavelet coefficients at scales j such that $0 \leq j \leq J$ can be computed by using (II.1.20-II.1.21) recursively. See Fig. II.1.3 for a schematic representation of the wavelet and scaling coefficients known at each step of this process. This whole process is called the fast wavelet transform algorithm. We come back to its practical implementation in Sec. II.2.2.

The inverse process is defined by inverting (II.1.20-II.1.21) thanks to the perfect reconstruction property of the filters. Using the unified notation, a compact formula for the reconstruction can be written:

$$\overline{f}_\lambda = \sum_{\substack{\boldsymbol{\mu}\in\{0,1\}^d \\ \mathbf{n}\in\{0,\dots,S-1\}^d \\ 2|\mathbf{i}-\mathbf{n}}} \prod_{k=1}^{d} h_{n_k}^{1-\mu_k} g_{n_k}^{\mu_k} \widetilde{f}_{(j^\lambda-1, \boldsymbol{\mu}, (\mathbf{i}-\mathbf{n}+(S-2)\boldsymbol{\mu})/2)}. \tag{II.1.22}$$

However, in practice, the computation is done by applying successively one dimensional filters, which is better understood when considering the following equivalent formulation:

$$\overline{f}_\lambda = \sum_{\substack{n_d=0 \\ 2|i_d-n_d}}^{S-1} \sum_{\mu_d=0}^{1} h_{n_d}^{1-\mu_d} g_{n_d}^{\mu_d} \dots \sum_{\substack{n_d=0 \\ 2|i_0-n_0}}^{S-1} \sum_{\mu_1=0}^{1} h_{n_1}^{1-\mu_1} g_{n_1}^{\mu_1} \widetilde{f}_{(j^\lambda-1, \boldsymbol{\mu}, (\mathbf{i}-\mathbf{n}+(S-2)\boldsymbol{\mu})/2)}. \tag{II.1.23}$$

For the DTCWT, the idea is to reduce the computation to a series of real wavelet transforms over bases constructed from the real and imaginary parts of the complex-valued wavelet, and then to use the fast wavelet transform as above. In this process the notations will be more and more tedious, and to attempt to simplify them we first introduce

$$\psi^{0,0} = \varphi^R, \quad \psi^{1,0} = \varphi^I, \quad \psi^{0,1} = \psi^R, \quad \psi^{1,1} = \psi^I.$$

We then rewrite (II.1.19) as follows:

$$\psi_\lambda(\mathbf{x}) = 2^{\frac{dj^\lambda - 1}{2}} \prod_{k=1}^{d} \left(\psi^{0,\mu_k}(2^{j^\lambda} x_k - i_k^\lambda) + (-1)^{\gamma_k^\lambda} \iota \psi^{1,\mu_k^\lambda}(2^{j^\lambda} x_k - i_k^\lambda) \right) \tag{II.1.24}$$

and expand this new expression:

$$\psi_\lambda(\mathbf{x}) = 2^{\frac{dj^\lambda - 1}{2}} \sum_{\mathbf{r} \in \{0,1\}^d} \iota^{|\mathbf{r}|} (-1)^{\boldsymbol{\gamma}^\lambda \cdot \mathbf{r}} \prod_{k=1}^{d} \psi^{r_k,\mu_k^\lambda}(2^{j^\lambda} x_k - i_k^\lambda). \tag{II.1.25}$$

Now using (II.1.25), the wavelet coefficients $\widetilde{f}_\lambda = \langle f \mid \psi_\lambda \rangle$ can be expanded as a linear combination of scalar products of f with functions of the form

$$\psi_\lambda^{\mathbf{r}}(\mathbf{x}) = \prod_{k=1}^{d} \psi^{r_k,\mu_k^\lambda}(2^{j^\lambda} x_k - i_k^\lambda). \tag{II.1.26}$$

Such functions can be seen as d-dimensional anisotropic wavelets, where in each direction k, either ψ^R or ψ^I is used, depending on the value of r_k. Formulas (II.1.20-II.1.21) and (II.1.23) can be generalized to compute the scalar products $\langle f \mid \psi_\lambda^{\mathbf{r}} \rangle$. We refrain from writing the generalized formulas explicitly, but they can be explained as follows. When computing the sum corresponding to direction k, the filter to use is respectively h^R if $(r_k, \mu_k) = (0,0)$, h^I if $(r_k, \mu_k) = (1,0)$, g^R if $(r_k, \mu_k) = (0,1)$ and g^I if $(r_k, \mu_k) = (1,1)$. Note that to ensure a better translation invariance at the discrete level, Kingsbury (2001) has proposed to use different filters for the first level of the wavelet transform. That means that if (II.1.20-II.1.21) is applied with $j^\lambda + 1 = J$, where J is the finest scale, a pair of biorthogonal filters are used instead of h^R and g^R, and the same filters shifted by one sample are used instead of h^I and g^I.

II.1.2.2 Scaling function coefficients and function values

In many applications, it is necessary to work in one way or another with the values of a function sampled on a Cartesian grid. To introduce wavelet series in these applications, it is therefore necessary to establish a link between the values of the function and its scaling coefficients. From the scaling coefficients, the wavelet coefficients can then be computed using the algorithm presented in the previous section.

Let us start from the definition:

$$\overline{f}_\lambda = \int_{\mathbb{T}[\mathbf{m}]} f(\mathbf{x}) \varphi_\lambda(\mathbf{x}) \mathrm{d}\mathbf{x} = 2^{\frac{dj^\lambda}{2}} \int_{\mathbb{T}[\mathbf{m}]} f(\mathbf{x}) \prod_{k=1}^{d} \varphi(2^{dj^\lambda} x_k - i_k^\lambda) \mathrm{d}\mathbf{x}$$

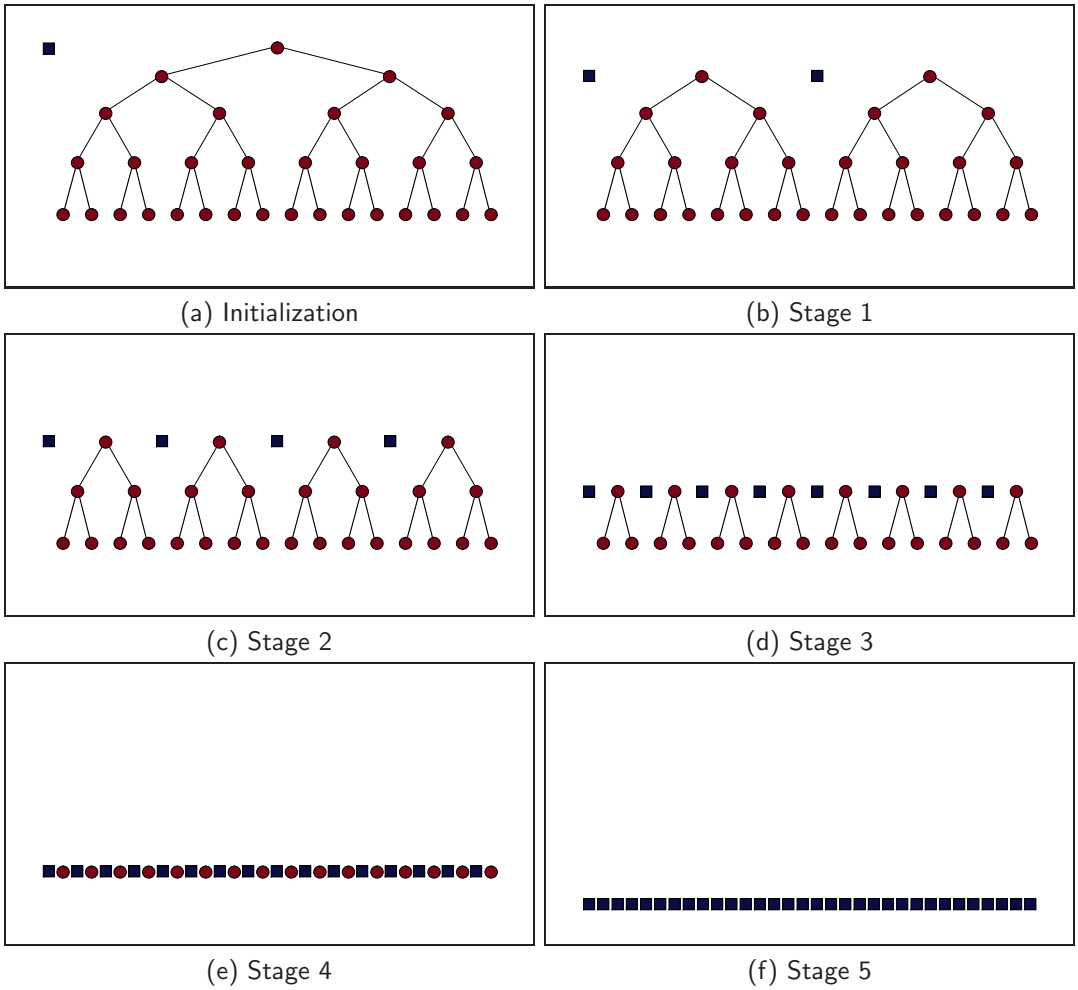


Figure II.1.3: Schematic representation of the available coefficients throughout successive stages of a non-adaptive 1D fast wavelet transform. Red circles: wavelet coefficients. Blue squares: scaling coefficients. The scale varies from $j = 0$ for the first row to $j = 4$ for the fifth row, and the position i increases from left to right.

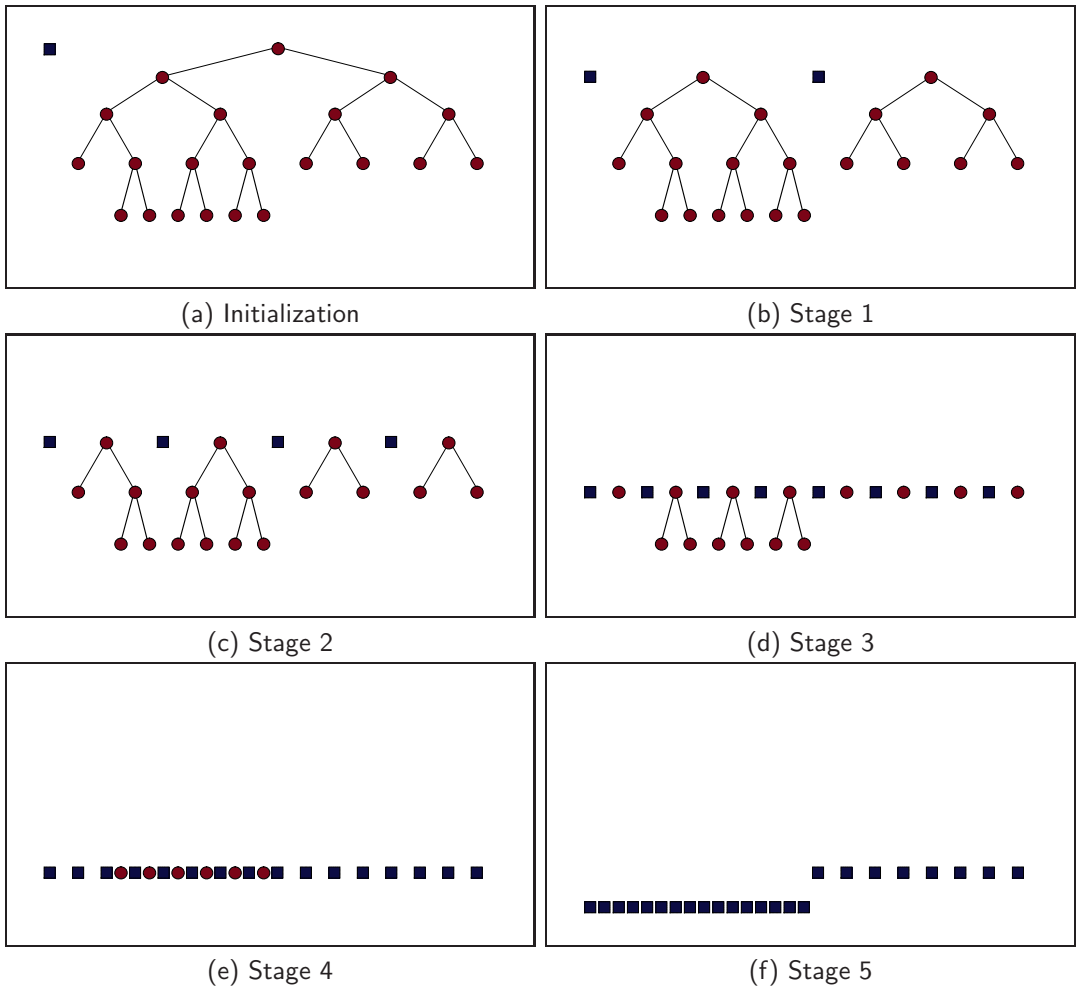


Figure II.1.4: Schematic representation of the available coefficients throughout successive stages of an adaptive 1D fast wavelet transform. Red circles: wavelet coefficients. Blue squares: scaling coefficients. In this example, the filter length is $S = 4$. Note the two additional scaling coefficients necessary to the left in (f), which correspond to the zone of influence of the wavelet coefficients present at stage 4 (see text and Eq. II.1.31).

Assuming that j^λ is sufficiently large so that the non overlap condition (II.1.15) holds, the integral can be extended to $\mathbb{R}^d$, and we can make the change of variable $\mathbf{y} = 2^{dj^\lambda}\mathbf{x} - \mathbf{i}^\lambda$:

$$\overline{f}_\lambda = 2^{-\frac{dj^\lambda}{2}} \int_{[0,S-1]^d} f\left(2^{-dj^\lambda}(\mathbf{y}+\mathbf{i}^\lambda)\right) \prod_{k=1}^{d} \varphi(y_k), \tag{II.1.27}$$

which can be rewritten

$$\overline{f}_\lambda = 2^{-\frac{dj^\lambda}{2}} f\left(2^{-dj^\lambda}\mathbf{i}^\lambda\right) + 2^{-\frac{dj^\lambda}{2}} \int_{[0,S-1]^d} \left(f\left(2^{-dj^\lambda}(\mathbf{y}+\mathbf{i}^\lambda)\right) - f\left(2^{-dj^\lambda}\mathbf{i}^\lambda\right)\right) \prod_{k=1}^{d} \varphi(y_k),$$

where we have used the property $\int_{[0,S-1]^d} \prod_{k=1}^{d} \varphi(y_k) = 1$. Now assuming that f is Lipschitz, we obtain the following approximation for $\overline{f}_\lambda$:

$$\overline{f}_\lambda = 2^{-\frac{dj^\lambda}{2}} \left(f\left(2^{-dj^\lambda}\mathbf{i}^\lambda\right) + O\left(2^{-dj^\lambda}\right)\right). \tag{II.1.28}$$

Note that the precision can be better if the scaling function φ has vanishing moments, as occurs for the wavelets known as "Coiflets" (Daubechies, 1992) (see Sec. II.1.3.2).

For some applications, a higher order approximation of $\overline{f}_\lambda$ is needed, and sometimes (II.1.28) cannot even be used because the values of f are not known on a dyadic grid but on a more complex grid or at arbitrary points. For all these cases, there exist an alternative approach, proposed by Sweldens and Piessens (1994) and developed by the quantum chemistry community for applications to ab initio density functional calculations (Johnson et al., 1999; Neelov and Goedecker, 2006). Suppose that values f_g of f are known at certain nodes $g \in \mathcal{G}$. The idea is to replace $f(2^{-dj^\lambda}(\mathbf{y}+\mathbf{i}^\lambda))$ in (II.1.27) by its approximation using a local Lagrange interpolation of the (f_g). The degree of the interpolating polynomial is chosen to be $S-1$. By reordering (II.1.27), one obtains an approximation of $\overline{f}_\lambda$ which depends on the coefficients of the Lagrange polynomial and on the moments of φ:

$$\mathcal{M}_k = \int_0^{S-1} x^k \varphi(x) \mathrm{d}x,$$

which can be determined using the recurrence relation (Johnson et al., 1999)

$$\mathcal{M}_k = \left(1 - 2^{-k}\right)^{-1} \sum_{p=0}^{k-1} \begin{pmatrix} k \\ p \end{pmatrix} \mathcal{M}_p \mathcal{M}'_{k-p} \ ,$$

where m'_p is a discrete moment of h:

$$\mathcal{M}'_p = \frac{1}{\sqrt{2}} \sum_{k=0}^{S-1} k^p h_k \ .$$

The reverse operation of going from the scaling coefficients to the grid point values can be done using the expansion

$$f(\mathbf{x}) \simeq \sum_{\lambda \in \Lambda_{\varphi,J}} \overline{f}_\lambda \varphi_\lambda(\mathbf{x}),$$

but if f is a smooth function this is not a very accurate expansion, except for J very large, because φ is normally not smooth. Depending on the situation, it may thus be more profitable to use another approach (Neelov and Goedecker, 2006) which we omit here.

II.1.2.3 Adaptive transform

As in the non-adaptive case, the goal of an adaptive wavelet transform is, starting from a certain set of scaling coefficients $\overline{f}_\lambda$, to compute a corresponding set of wavelet coefficients $\widetilde{f}_\lambda$. But in contrast to the non-adaptive case, the input scaling coefficients are not all located at the same scale. In the formulation that we have chosen, it is more natural to start by defining the inverse wavelet transform. Consider a finite set of indices $\Lambda \subset \Lambda_\psi$, and denote W_Λ the subspace of $L^2(\mathbb{T}[\mathbf{m}]^d)$ spanned by $(\psi_\lambda)_{\lambda\in\Lambda}$. A function $f \in W_\Lambda$ can be decomposed as

$$f = \overline{f} + \sum_{\lambda\in\Lambda} \widetilde{f}_\lambda \psi_\lambda, \tag{II.1.29}$$

where $\overline{f}$ is the mean value of f on $\mathbb{T}[\mathbf{m}]$. Now the aim of the inverse wavelet transform is to compute scaling function coefficients $(\overline{f}_\lambda)$ for λ in a certain set $\overline{\Lambda}$, such that they contain all the information needed to reconstruct f. The main difficulty compared to the non-adaptive case is that the corresponding scaling functions will in general not be pairwise orthogonal, and even not linearly independent, since they may live at different scales.

Necessary properties on $\overline{\Lambda}$. To construct $\overline{\Lambda}$, we proceed one scale at a time. Denote $J_0 = \min\left\{ j^\lambda \middle|\, \lambda \in \Lambda\right\}$ the coarsest scale present in $\Lambda_0 = \Lambda$. We would like to find a set $\Lambda_1 \subset \Lambda_A$ enjoying the following properties:

(i) $J_{p+1} := \min\left\{ j^\lambda \middle|\, \lambda \in \Lambda_{p+1}, \boldsymbol{\mu}^\lambda \neq 0\right\} = J_p + 1$,

(ii) all the coefficients of Λ_p can be computed from the coefficients of Λ_{p+1},

(iii) all the coefficients of Λ_{p+1} can be computed from the coefficients of Λ_p.

where for the moment $p = 0$. Note that for Λ_{p+1} we use the unified notation, that is, Λ_{p+1} may contain both wavelet and scaling function indices. Condition (i) means that Λ_{p+1} should not contain wavelet coefficients at scales coarser than J_p+1. Hence passing from Λ_p to Λ_{p+1} consists in replacing the wavelet coefficients at scale J_p by scaling coefficients at scale J_p+1, while leaving the indices at scales J_p+1 or finer unchanged (see Fig. II.1.3). To understand what to do concerning scale J_p, we start by considering condition (ii). From formula (II.1.21), it appears that a wavelet coefficient indexed by $\lambda = (J_p, \boldsymbol{\mu}^\lambda, \mathbf{i}^\lambda) \in \Lambda_p$ depends on the scaling coefficient indexed by $\kappa = (J_p + 1, 0, \mathbf{i}^\kappa)$ if and only if for all directions k,

$$2i_k^\lambda - (S-2) \leq i_k^\kappa \leq 2i_k^\lambda + 1. \tag{II.1.30}$$

We have thus shown that

$$\Lambda_{p+1} \supset \left\{ (J_p+1, 0, \mathbf{i}) \middle| \exists \lambda \in \Lambda, j^\lambda = J_0, \forall k, 2i_k^\lambda - (S-2) \leq i_k^\kappa \leq 2i_k^\lambda + 1 \right\} \tag{II.1.31}$$

where, as usual, everything is implicitly periodized with a period 2^{j^κ}. On the other hand, to satisfy condition (iii), we need to be able to compute the coefficients at scale J_p in Λ_{p+1} from those in Λ_p. Using (II.1.23), we deduce the following condition:

$$\Lambda_p \supset \left\{ (j-1, 0, (\mathbf{i}-\mathbf{n})/2) \,|(j,0,i) \in \Lambda_{p+1} \forall k, 0 \leq n_k \leq S-1 \text{ and } 2 \mid i_k - n_k \right\}. \tag{II.1.32}$$

Since we had assumed that Λ_0 contains only wavelet coefficients, (II.1.32) cannot be satisfied for $p = 0$ unless $J_0 = 0$, which we impose from now on. In that case (II.1.32) becomes unnecessary since the only scaling coefficient at scale 0 is already taken into account as the mean value appearing in (II.1.29). The situation becomes more interesting when the process is iterated to find Λ_2, Λ_3, etc., satisfying (i), (ii), (iii) respectively for $p = 1$, $p = 2$, etc. Imposing (II.1.31) will be sufficient to apply Eq. (II.1.21) so as to reconstruct all the wavelet coefficients in Λ_{p-1} from the coefficients of Λ_p. However, this will not be enough to compute the scaling coefficients of Λ_{p-1}, because the indices appearing in Eq. (II.1.20) are not the same, so that it seems that we may be missing some coefficients!

Edge matrices. To understand the situation better, it is useful to come back to the 1D case, and to let $S = 4$. Then Eqs. (II.1.20-II.1.21) can be written explicitly as follows:

$$\overline{f}_{(j,i)} = h_0\overline{f}_{(j+1,2i)} + h_1\overline{f}_{(j+1,2i+1)} + h_2\overline{f}_{(j+1,2i+2)} + h_3\overline{f}_{(j+1,2i+3)} \quad \text{(II.1.33)}$$

$$\widetilde{f}_{(j,i)} = g_0\overline{f}_{(j+1,2i-2)} + g_1\overline{f}_{(j+1,2i-1)} + g_2\overline{f}_{(j+1,2i)} + g_3\overline{f}_{(j+1,2i+1)}, \quad \text{(II.1.34)}$$

where the notation $(j, i) = \lambda$ for the indices has been used for clarity. By condition (II.1.31) on Λ_{p+1}, all scaling coefficients appearing in (II.1.34) are known. But if λ lies on the frontier of Λ_p, i.e. $(j, i+1) \notin \Lambda_p$, as we assume from now on, $\overline{f}_{(j+1,2i+2)}$ and $\overline{f}_{(j+1,2i+3)}$ are unknown. Fortunately, there are enough constraints on f to determine them exactly without employing any ad-hoc interpolation method. To see this, recall that $\widetilde{f}_\lambda = 0$ for $\lambda \notin \Lambda_p$. In particular,

$$\widetilde{f}_{(j,i+1)} = g_0\overline{f}_{(j+1,2i)} + g_1\overline{f}_{(j+1,2i+1)} + g_2\overline{f}_{(j+1,2i+2)} + g_3\overline{f}_{(j+1,2i+3)} = 0. \quad \text{(II.1.35)}$$

On the other hand, by (II.1.22) we have

$$\overline{f}_{(j+1,2i+2)} = h_0\overline{f}_{(j,i+1)} + h_2\overline{f}_{(j,i)} + g_0\widetilde{f}_{(j,i+2)} + g_2\widetilde{f}_{(j,i+1)} \quad \text{(II.1.36)}$$

$$\overline{f}_{(j+1,2i+3)} = h_1\overline{f}_{(j,i+1)} + h_3\overline{f}_{(j,i)} + g_1\widetilde{f}_{(j,i+2)} + g_3\widetilde{f}_{(j,i+1)}, \quad \text{(II.1.37)}$$

where $\widetilde{f}_{(j,i+1)} = 0$ according to (II.1.35). To proceed further, we make an assumption on the set Λ_p. Going back to an undetermined value of S, but still in the 1D setting, we impose that

$$((j,i) \in \Lambda_p \text{ and } (j,i+1)) \notin \Lambda_p \implies \forall k \in \left\{1, \ldots, \frac{S}{2}\right\}, (j, i+k) \notin \Lambda_p, \quad \text{(II.1.38)}$$

which roughly means that Λ_p cannot contain any hole of length smaller than $\frac{S}{2}$. Now using (II.1.38), we get that $\widetilde{f}_{(j,i+2)} = 0$, and (II.1.36-II.1.37) thus becomes

$$\overline{f}_{(j+1,2i+2)} = h_0\overline{f}_{(j,i+1)} + h_2\overline{f}_{(j,i)} \quad \text{(II.1.39)}$$

$$\overline{f}_{(j+1,2i+3)} = h_1\overline{f}_{(j,i+1)} + h_3\overline{f}_{(j,i)}, \quad \text{(II.1.40)}$$

which can be injected back into (II.1.33) to get:

$$\overline{f}_{(j,i)} = h_0\overline{f}_{(j+1,2i)} + h_1\overline{f}_{(j+1,2i+1)} + (h_2h_0 + h_3h_1)\overline{f}_{(j,i+1)} + (h_2^2 + h_3^2)\overline{f}_{(j,i)}.$$

Following (Beylkin, 1992), we now show that the interesting cancellation $h_2h_0 + h_3h_1 = 0$ holds as long as h satisfies (II.1.4). Indeed, coming back to the general case for S, (II.1.4) can be rewritten more explicitly as

$$\forall \xi, \sum_{p,q} \left(1 + (-1)^{p+q}\right) h_p h_q e^{\iota(p-q)\xi} = 2,$$

which, by regrouping equal powers of $e^{\iota\xi}$, is seen to imply:

$$\forall k \in \{1, (S-2)/2\}, \sum_{p=2k}^{S-1} h_p h_{p-2k} = 0. \tag{II.1.41}$$

Using this result, we finally get for our $S = 4$ example that

$$(1 - h_2^2 - h_3^2)\overline{f}_{(j,i)} = h_0 \overline{f}_{(j+1,2i)} + h_1 \overline{f}_{(j+1,2i+1)}, \tag{II.1.42}$$

which can be solved for $\overline{f}_{(j,i)}$. It turns out that (II.1.42) admits a generalization to $S > 4$, where the missing coefficients $\overline{f}_{(j,i)}, \ldots, \overline{f}_{(j,i+S/2-2)}$ are determined by a linear system of $S/2 - 2$ equations. The solution to this system can be written in the form

$$\begin{pmatrix} \overline{f}_{(j,i)} \\ \vdots \\ \overline{f}_{(j,i+S/2-2)} \end{pmatrix} = M_E \begin{pmatrix} \overline{f}_{(j+1,2i)} \\ \vdots \\ \overline{f}_{(j+1,2i+S-3)} \end{pmatrix}, \tag{II.1.43}$$

where we call M_E the "edge matrix".

Sufficient definition of $\overline{\Lambda}$. Now provided that condition (II.1.38) is satisfied at each step, we are able to construct a sequence (Λ_p) such that $\Lambda_0 = \Lambda$, and verifying (i), (ii) and (iii), simply by letting:

$$\begin{aligned} \Lambda_{p+1} &= \{\lambda \in \Lambda_p \mid j^\lambda \geq p+1\} \\ &\bigcup \left\{(p+1, 0, \mathbf{i}) \,\middle|\, \exists \lambda \in \Lambda_p, j^\lambda = p, \forall k, 2i_k^\lambda - (S-2) \leq i_k^\kappa \leq 2i_k^\lambda + 1\right\}. \end{aligned}$$

We stop the sequence at the index $p_{\max}$ such that $\Lambda_{p_{\max}}$ does not contain any wavelet coefficient, and we finally define

$$\overline{\Lambda} = \Lambda_{p_{\max}}. \tag{II.1.44}$$

The definitions in this paragraph were admittedly rather cryptic, and to help the understanding some explicit examples will be provided in Sec. II.2.3. A schematic view of the sets Λ_p is provided in Fig. II.1.4 for a simple 1D case with $S = 4$.

II.1.3 Wavelet families

In this section, we recall the main properties of the wavelet families that are commonly used in computational fluid dynamics. We also summarize for convenience a number of algorithms used to compute the coefficients of the associated filters. Note that the literature provides pre-computed values for the coefficients, but they sometimes suffer from a rather low precision.

II.1.3.1 Daubechies minimal support wavelets

These famous wavelet families were constructed in (Daubechies, 1988). Their main property is that the length of their support is minimal among all orthogonal wavelet families having a

given number of vanishing moments. If M denotes the number of vanishing moments, then the length of the filters is $S = 2M$, and the length of the support is thus $S - 1 = 2M - 1$.

To compute the filter coefficients, the procedure is as follows. Let P_M be the polynomial of degree $M - 1$ defined by

$$P_M(X) = \sum_{k=0}^{M-1} \left(\begin{array}{c} M-1+k \\ k \end{array} \right) X^k,$$

In (Daubechies, 1988) it is shown that the discrete Fourier transform of a finite length filter corresponding to an orthogonal wavelet basis can necessarily be factorized under the form:

$$\hat{h}(\xi) = 2^{-M} \left(1 + e^{\iota\xi}\right)^M Q\left(e^{\iota\xi}\right),$$

where Q is a polynomial such that

$$|Q\left(e^{\iota\xi}\right)|^2 = P_M\left(\sin^2\frac{\xi}{2}\right) + \left(\sin\frac{\xi}{2}\right)^{2M} R\left(\frac{1}{2}\cos(\xi)\right), \qquad \text{(II.1.45)}$$

and R is an odd polynomial.

To build Daubechies minimal support wavelets, one first chooses $R = 0$, and then one has to solve (II.1.45) for Q. The procedure to do this is explained in (Daubechies, 1988), and we recall it in a simplified manner. Denoting by $(a_k)_{1 \le k \le M-1}$ the complex roots of P_M, compute for each k the two roots r_k^+ and r_k^- of the equation

$$X^2 - 2(a_k - 2) + 1 = 0,$$

and let

$$r_k = \begin{cases} r_k^+ \text{ if } |r_k^+| \le 1 \\ r_k^- \text{ otherwise.} \end{cases} \qquad \text{(II.1.46)}$$

The resulting Q is given by

$$Q(X) = C \prod_{k=1}^{M-1} (X - r_k),$$

where the normalization constant C can be determined a posteriori using the constraint that $\sum_{i=0}^{S-1} h_i = \sqrt{2}$. This procedure is called "minimal phase spectral factorization", and is classical in filter design. It is possible to replace (II.1.46) by another choice of roots, in which case the spectral factorization is not minimal phase anymore, but still yields minimal support length Daubechies wavelets. For example, in (Daubechies, 1993), it was proposed to globally optimize the choice between the r_k^+ and the r_k^- so that the resulting wavelets are as symmetric as possible. These are known as "Symlets".

To implement the algorithm in practice, one needs a high precision polynomial root finder, which exists for example in C++ in the Nacre library. The advantage of using C++ and generic programming is that the algorithm can be made as precise as required by changing the type of the floating point numbers. For example, the minimal phase results are given in Table II.1.3.2 for $M = 2$, $M = 3$, $M = 4$, and $M = 5$, with 18 accurate digits. The accuracy can be checked a posteriori by testing the perfect reconstruction property.

II.1.3.2 Coiflets

Coiflets families were constructed by Daubechies in (Daubechies, 1993). For a given number of vanishing moments M, the corresponding Coiflet has a support of length $3M$. The characteristic property of Coiflets is that, in addition to vanishing moments for the wavelet (Eq. II.1.5), they benefit from vanishing moments for the scaling function, that is

$$\int_{\mathbb{R}} x^m \varphi(x) = 0 \text{ for } 1 \leq m \leq M-1. \tag{II.1.47}$$

This property endows the simple approximation of scaling function coefficients (II.1.28) in the case of Coiflets with a higher precision than for general wavelet families.

The construction of the Coiflets uses techniques that are similar to the above. The main idea is to relax the constraint $R = 0$, in order to be able to impose the additional condition (II.1.47). The filter values are provided in (Daubechies, 1993).

An interesting section at the very end of (Daubechies, 1993) defines specific low order Coiflets, which we call R-Coiflets to distinguish them from the previous ones. The R-Coiflets are specifically tuned to enjoy a better regularity than the standard Coiflets for given filter length and number of vanishing moments, hence the prefix "R". We recall here the algorithm for the computation of the corresponding filter coefficients for $M = 2$ and $M = 4$, that is respectively for the R-Coiflet 6 and R-Coiflet 12 families. In general for Coiflets, the following Ansatz for the filter in the Fourier domain is introduced:

$$\hat{h}(\xi) = \left(\cos^2 \frac{\xi}{2}\right)^K \left(P_K\left(\sin^2 \frac{\xi}{2}\right) + \left(\sin^2 \frac{\xi}{2}\right)^K f(\xi) \right), \tag{II.1.48}$$

where $M = 2K$, and f remains unknown. For R-Coiflet 6, one uses the Ansatz $f(\xi) = ae^{\iota\xi} + be^{2\iota\xi}$, for which (II.1.48) implies that $a = (s-1)/2, b = (-s+3)/2, \text{ where } s = \pm\sqrt{15}$. The R-Coiflet 6 corresponds to $s = +\sqrt{15}$ (and not $s = -\sqrt{15}$ as is mistakenly stated in (Daubechies, 1993)).

For R-Coiflet 12, the Ansatz is $f(\xi) = ae^{-\iota\xi} + b + ce^{\iota\xi} + de^{2\iota\xi}$. By developing (II.1.48) in powers of $e^{\iota\xi}$ and identifying term by term, we obtain the following system of equations:

$$\begin{cases} (a+b+c+d)^2 = 16d + 8(a+c) \\ ac + bd + 6ad = 0 \\ a + c + 4d + 2 = 0 \\ b = 3d + 5 \end{cases},$$

which can be pivoted to get the system

$$\begin{cases} 32d^2 + 88d + 45 = 0 \\ 2c^2 + (8d+4)c + (26d^2 + 78d + 45) = 0 \\ a = -2 - 4d - c \\ b = 3d + 5 \end{cases}$$

admitting 4 quadruplets (a, b, c, d) as solutions. One of these solutions yields a filter which agrees with the coefficients given in (Daubechies, 1993). The filter coefficients for the R-Coiflet 6 and R-Coiflet 12 wavelet families are provided in Table II.1.3.2.

R-Coiflet 6	**Daubechies 4**	**Daubechies 10**
0.568910863637797393	0.482962913144534143	0.160102397974192914
0.403360960493705453	0.836516303737807906	0.603829269797189670
-0.077161555495773499	0.224143868042013381	0.724308528437772927
0.607491641385684143	-0.129409522551260381	0.138428145901320731
0.215357473044523631		-0.242294887066382032
-0.303745820692842071	**Daubechies 6**	-0.032244869584638374
	0.332670552950082616	0.077571493840045713
R-Coiflet 12	0.806891509311092576	-0.006241490212798274
0.036797113136414754	0.459877502118491570	-0.012580751999081999
-0.006700453459679574	-0.135011020010254589	0.003335725285473771
-0.193465866829234017	-0.085441273882026661	
0.015361209805458600	0.035226291885709536	
0.626863205073584621		
0.712666476561508862	**Daubechies 8**	
0.238059669115279298	0.230377813308896501	
-0.041841810360839877	0.714846570552915647	
0.000935901149918644	0.630880767929858908	
0.039061962673359208	-0.027983769416859854	
-0.002083240459415777	-0.187034811719093084	
-0.011440604033259695	0.030841381835560763	
	0.032883011666885199	
	-0.010597401785069032	

Table II.1.1: Coefficients of R-Coiflet 6, R-Coiflet 12 and some Daubechies minimal support/minimal phase filters with 18 digits of accuracy.

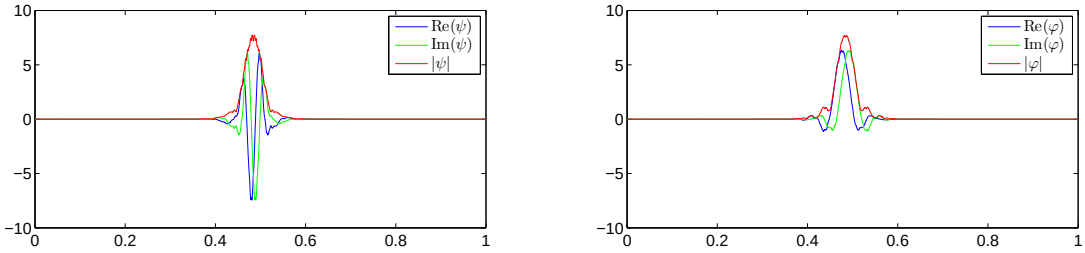


Figure II.1.5: Complex-valued wavelet from the Q-Shift B Kingslet family (left) and the corresponding scaling function (right).

II.1.3.3 Kingslets

We call Kingslets the Q-Shift wavelets described by Kingsbury in (Kingsbury, 2001). Here we briefly review these wavelet families, especially since we will use them for one application in Sec. IV.1. Kingslets belong to the class of DTCWT wavelets already introduced in Sec. II.1.1.3. Contrary to most other wavelet families, Kingslets filters are not defined by equations but are the result of an optimization procedure. By varying the parameters of the optimization procedure, various properties can be enforced to a good degree of approximation.

In the applications of Kingslets that we consider in Chapter 3, we shall only make use of the Kingslet named Q-Shift B, which is represented in Fig. II.1.5 next to the corresponding scaling function. It has two approximately vanishing moments (their amplitude is about 10^{-6}), and filters of length 14. Note that the extra length of the Kingslet filters allows them to enjoy good translation invariance properties, but their number of vanishing moments is quite low. In the future it may be possible to construct Kingslets with more vanishing moments.

II.1.3.4 Selesnick wavelets

Selesnick wavelets are alternative DTCWT wavelet families that were introduced in (Selesnick, 2002). If M is the number of vanishing moments (denoted by L in (Selesnick, 2002)), the length of the filter is $2(M+L)$, where L is an adjustable integer which controls how well the Fourier transform of the wavelet vanishes for negative wave-numbers.

The filter coefficients of minimal phase Selesnick wavelets are provided in Table II.1.2. They were computed with 15 digits precision using a C++ implementation of the Matlab code provided in (Selesnick, 2002), with some modifications to improve the precision of the result. Note that the filters differ from those in (Selesnick, 2002), which used mid-phase spectral factorizations instead of minimal phase.

II.1.4 Representation of differential operators

The representation of differential operators in wavelet spaces is essential for applications to the numerical approximation of partial differential equations. Important results on this topic

Selesnick 1/2	**Selesnick 2/4**	**Selesnick 3/3**	**Selesnick 4/2**
h^R	h^R	h^R	h^R
0.063722005204896	0.003713680476139	0.008846655842947	0.021102599277254
0.679872882462790	0.137657968389623	0.202541675343599	0.268378378150965
0.725610498781959	0.608838482922821	0.663564515821794	0.712965714026111
0.019527181672067	0.744293913959969	0.681689981462920	0.605543425035167
-0.082225722800308	0.099827234284518	0.029191650826211	-0.047563788371239
0.007706717051690	-0.211975875247260	-0.223975383050148	-0.214427465912031
h^I	0.002643711948873	0.018345881119283	0.034799514500651
0.001541343410338	0.037516078031551	0.048379737687138	0.052591871682046
-0.004114397277358	-0.008056349815213	-0.013509874338369	-0.016656062380741
-0.144918766341865	-0.000381526507967	-0.001500055319494	-0.004786091932662
0.392611152439421	0.000140021369405	0.000667951914677	0.002458804134508
0.850484204118075	-0.000003777439371	-0.000029174937471	-0.000193335836941
0.318610026024484	h^I	h^I	h^I
	-0.000000419715485	-0.000004167848210	-0.000038667167388
Selesnick 2/1	0.000004365516982	0.000028736130733	0.000182423487795
h^R	0.000429940683837	0.001512510330566	0.003409940503429
0.150856631227302	-0.004568431032085	-0.010767728009051	-0.017300957911382
0.681133683988679	0.004506876043497	0.008846447292283	0.006340085364341
0.700930246420804	0.054321878030781	0.066565500871343	0.086267604737072
-0.006070468508676	-0.115201566195227	-0.106398625212835	-0.094497819046072
-0.144680096461559	-0.160517482306511	-0.196613924769040	-0.231826057852929
0.032043565706544	0.469733549136782	0.376184343629676	0.294105321867595
h^I	0.784443326692122	0.785967606061924	0.764270772339713
0.01068118856884	0.347638401233140	0.426966272995064	0.497787919664638
-0.01974352930359	0.033423124285255	0.061926590900633	0.105512996386273
-0.14012240952992			
0.27428041680823			
0.83654800214762			
0.45256989368190			

Table II.1.2: Coefficients of some Selesnick filters with 15 digits of accuracy.

where derived in (Beylkin, 1992), and we briefly recall a small part of them that we will make use of.

We fix a positive integer J and work in the space V_J spanned by a family of scaling functions $(\varphi_\lambda)_{\lambda\in\Lambda_{\varphi,J}}$. A function $f \in V_J$ can thus be decomposed into

$$f = \sum_{\lambda\in\Lambda_{\varphi,J}} \overline{f}_\lambda \varphi_\lambda.$$

Now denoting by $D_{\mathbf{p}}$ a differential operator $\dfrac{\partial^{|\mathbf{p}|}}{\partial^{p_1}x_1\ldots\partial^{p_d}x_d}$ where $\mathbf{p}\in\mathbb{N}^d$ and $|\mathbf{p}| = \sum_{k=1}^d p_k$, we have:

$$D_{\mathbf{p}} f = \sum_{\lambda\in\Lambda_{\varphi,J}} \overline{f}_\lambda D_{\mathbf{p}}\varphi_\lambda.$$

The Galerkin representation of $D_{\mathbf{p}}$ in V_J, for which we use the shorthand notation $\overline{D}_{\mathbf{p}}$, is therefore given by

$$\overline{D}_{\mathbf{p}} f = \sum_{\lambda,\lambda'\in\Lambda_{\varphi,J}} \overline{f}_\lambda \left\langle D_{\mathbf{p}}\varphi_\lambda \middle|\, \varphi_{\lambda'}\right\rangle \varphi_{\lambda'}.$$

In (Beylkin, 1992), an effective way of computing the matrix elements $\left\langle D_{\mathbf{p}}\varphi_\lambda \middle|\, \varphi_{\lambda'}\right\rangle$ is derived. First, by a change of variables it is shown that

$$\left\langle D_{\mathbf{p}}\varphi_\lambda \middle|\, \varphi_{\lambda'}\right\rangle = 2^{|\mathbf{p}|J}\prod_{k=1}^d r^{(p_k)}_{i^{\lambda'}_k - i^\lambda_k},$$

where

$$r^{(p)}_i = \int_{\mathbb{T}} \varphi^{(p)}(y)\varphi(y-i)\mathrm{d}y.$$

Then, the $r^{(p)}_i$ are shown to satisfy the following linear system (Beylkin, 1992):

$$\begin{cases} r^{(n)}_i = 2^n\left(r^{2i} + \frac{1}{2}\sum_{k=1}^{S/2} a_{2k-1}(r^{(n)}_{2i-2k+1} + r^{(n)}_{2i+2k-1})\right) \\ \sum_i i^n r^{(n)}_i = (-1)^n n! \end{cases}, \tag{II.1.49}$$

where $(a_i)_{0\le i\le S-1}$ depends on the scaling function filter h via

$$a_i = \sum_{k=0}^{S-1-i} h_k h_{k+i}.$$

In (Beylkin, 1992), it is shown that (II.1.49) admits a unique solution, and this solution is given for some of the Daubechies minimal support wavelets. We provide here the solution for R-Coiflet 1 and R-Coiflet 2 wavelets, see Table II.1.3.

II.1.5 Denoising

In this section, we briefly recall the mathematical theory of denoising using wavelets. The presentation is meant to provide a heuristic basis for most of the developments and results

R-Coiflet 6, $r^{(1)}$	**R-Coiflet 6**, $r^{(2)}$	**R-Coiflet 12**, $r^{(1)}$	**R-Coiflet 12**, $r^{(2)}$
0	0.622642595976898	0	-3.658969347796646
0.619368727403235	-0.389477243531951	-0.867954504854107	2.291512188822066
1.403784906417689	0.179082676222473	0.285598759981490	-0.634322626940372
-2.427684222146337	-0.326853461357942	-0.095340803975880	0.234068159382585
0.839028531550099	0.225926730678971	0.025921268507550	-0.078417496426037
		-0.004471277482279	0.018448823282131
		0.000182709149420	-0.001655689924465
		0.000045997383079	-0.000116101105014
		0.000004103052443	-0.000032781272494
		-0.000000039014015	0.000000198414036
		0.000000000032848	-0.000000000334114

Table II.1.3: Coefficients of first and second order derivation filters for R-Coiflet 6 and R-Coiflet 12 wavelet families. The values of $r_i^{(n)}$ are given for $i \leq 0$ in increasing order. Those for $i < 0$ can be deduced by symmetry for even n, and by antisymmetry for odd n.

presented in the rest of this thesis. The topic is treated in more detail in the book by Mallat (1999).

We restrict ourselves from the beginning to the case of an additive noise, for which the observed signal can be written

$$X[n] = F[n] + W[n],\ n = 1..N \tag{II.1.50}$$

where W is a centered Gaussian process, and $(F[n])$ is a deterministic signal resulting from the uniform sampling of a function $F \in L^2(\mathbb{T})$, via $F[n] = F(\frac{n}{N})$. The problem of denoising consists in the estimation of F from the noisy measurements $(X[n])$. The efficiency of any proposed estimator $\mathcal{E}$ can be measured by the associated risk

$$R(\mathcal{E}, F) = \sum_{n=1}^{N} \mathbb{E}\left(|F[n] - \mathcal{E}(X)[n]|^2\right),$$

or equivalently the signal to noise ratio (SNR) measured in decibels (dB), and defined by

$$\mathrm{SNR}(\mathcal{E}, F) = 10 \log_{10}\left(\frac{\|F\|_2^2}{R(\mathcal{E}, F)}\right).$$

To model the a priori knowledge we have about the signal, we assume that it belongs to some subset of $L^2(\mathbb{T})$ denoted $\mathcal{T}$. Following the minimax approach, we then look for estimators $\mathcal{E}$ such that the maximal risk

$$R(\mathcal{E}, \mathcal{T}) = \sup_{F \in \mathcal{T}} R(\mathcal{E}, F) \tag{II.1.51}$$

is as small as possible (hence the name minimax). In the best of worlds, the risk of such an estimator will thus be

$$R(\mathcal{T}) = \inf_{\mathcal{E}} R(\mathcal{E}, \mathcal{T})$$

where the lower bound is taken over all possible estimators. Two things may influence the value of $R(\mathcal{T})$, namely

- the set $\mathcal{T}$ which is assumed to contain the signal,
- the properties of the noise W.

From now one we use the shorthand notation $\check{.}$ to denote statistical estimators, i.e. $\mathcal{E}(X) = \check{F}$.

Donoho and Jonhstone (1994) studied the efficiency of a wavelet thresholding approach to the minimax estimation problem in the case where W is uncorrelated (i.e. a white noise). They showed that this approach, which we now briefly recall, almost achieves the minimax risk for very diverse choices of $\mathcal{T}$. First, a 1D orthonormal wavelet basis is chosen, over which the wavelet coefficients $(\widetilde{X}_\lambda)$ of the noisy signal are computed. To do this, the approximation (II.1.28) is always used for the scaling function coefficients at the finest scale, since the noise cannot be interpolated. We therefore have to assume that $N = 2^J$, i.e., that the signal has been sampled on a dyadic grid. By linear transformation of (II.1.50), a similar equation is obtained in terms of wavelet coefficients:

$$\widetilde{X}_\lambda = \widetilde{F}_\lambda + \widetilde{W}_\lambda \tag{II.1.52}$$

Thanks to the properties of the Gaussian distribution and to the orthogonality of the wavelet basis, the wavelet coefficients $(\widetilde{W}_\lambda)$ of the noise are Gaussian and independent. The advantage of (II.1.50) compared to (II.1.50) is that for the types of signals that arise commonly in many applications, the energy of F is concentrated on a very small fraction of its wavelet coefficients, a property which is called sparsity. In contrast, the energy of the noise is evenly distributed between all of its wavelet coefficients. Consequently, the wavelet coefficients containing the majority of the signal's energy are only weakly affected by the noise, and conversely, the remaining coefficients are strongly affected by the noise but they do not contribute much to the signal. This heuristic observation leads to the following candidate for estimating the signal:

$$\check{\widetilde{F}}_\lambda = \begin{cases} \widetilde{Y}_\lambda \text{ if } |\widetilde{Y}_\lambda| > q\sigma \\ 0 \text{ otherwise} \end{cases}, \tag{II.1.53}$$

where σ is the standard deviation of the noise, and q is a chosen constant. Donoho and Jonhstone (1994) showed that, with the choice $q = \sqrt{2 \log N}$, the method was asymptotically minimax when $N \to \infty$ up to a logarithmic correction factor. Eq. (II.1.53) admits an obvious generalization to the case where $\sigma := \sigma_\lambda$ depends on the index λ. It has been proposed by many authors to apply this generalized thresholding approach to extend wavelet denoising to situations where the noise is correlated (Wang, 1996; Silverman, 1999; Johnstone and Silverman, 1997). Indeed, for some simple correlations models, the behavior of σ_λ can be computed. Explicit examples are given in Sec. II.3.4 to demonstrate the feasibility of the approach.

In practical denoising applications, the variance of the noise is often unknown and then has to be estimated from X. To do this in the case of a white noise, Donoho and Jonhstone (1994) have proposed to compute the median M_D of the $|\widetilde{Y}|_\lambda$ where λ is restricted to the finest scale. Indeed, for sufficiently sparse signals, the finest scale is likely to be mostly affected by the noise. By neglecting the influence of the signal, one can show that σ is related to M_D by $\sigma \simeq 1.4826 M_D$. In the case of a correlated noise, i.e. when σ_λ depends on scale, the same procedure can be applied scale by scale. For white noises, an alternative approach was introduced by Azzalini et al. (2004), and we have extended it to the case of

correlated noises. The reader is referred to Sec. V.3.5.1 for a detailed presentation of this method, called scale-wise CVE, which we do not reproduce here to avoid an unnecessary redundancy.

II.2 Implementation

In this section we complete the presentation of the FWT algorithms we have implemented, by insisting on some important points that fall more into the category of computer science than mathematics per se. C++ implementations of the algorithms are available under the GNU General Public License at the following URL:

`https://sourceforge.net/project/kicksey-winsey`.

II.2.1 Review of some existing implementations

The implementation of the fast wavelet transform algorithm immediately followed its mathematical derivation. Indeed, when working with moderately sized data, the algorithm is straightforward. Many toolboxes exist in Matlab for the 1D and 2D cases, among which the best known is Wavelab (Mallat, 1999). However, for the applications we are considering, this kind of implementation is insufficient, since it cannot take advantage of modern parallel architectures.

The parallelization of the fast wavelet transform algorithm was first demonstrated in architecture dependent frameworks (Goirand et al., 1994; Holmström, 1995), but since then much of the effort has been spent in developing efficient algorithms that run within generic distributed memory parallel environments such as the message passing interface (MPI). Early approaches used to mimic the parallel FFT algorithm, which involves a global transposition step that requires exchange of data between all processes, and thus strongly degrades efficiency. The parallelization strategy that we employ here was first proposed and studied by Nielsen and Hegland (2000) for both the 1D wavelet transform and the 2D tensor-product wavelet transform. The key ingredient to obtain good scalability is some sort of domain decomposition approach, since it preserves the quasi-locality of the wavelet transform (Yang and Misra, 1998; Gonzales et al., 2001). It was extended to 2D multiresolution analyses in (Chaver et al., 2002), who reported parallelization efficiency of up to 0.9 using a maximum of 32 processes and images of size up to 2048×2048. Recently, the wavelet transform was also ported to general purpose graphics processing units (GPGPU) (Tenllado et al., 2008).

Our main objective in the development of the implementation presented below was to be able to treat extremely large datasets, especially in 3D or more. By extremely large, we mean that the dataset does not fit inside the memory of a standard currently available personal computer, which is likely to be the case if it is larger than, say, $16Gb$.

II.2.2 General structure of our approach

II.2.2.1 Notations

To describe the implementation of the wavelet transform, it is convenient to define some new vocabulary and notations. We shall call "level" of a wavelet or scaling coefficient the integer $l = J - j$, where j is the scale of the coefficient. The wavelet transform is "adaptive" when the input scaling coefficients are not all at the same scale (see Sec. II.1.2.3). The fast wavelet transform algorithm consists in several successive "stages". After stage k, the level of all remaining scaling coefficients in the representation of the function must be greater than k. The multiple real wavelet transforms that have to be computed when doing the DTCWT of a signal (see Sec. II.1.1.3) will be called "trees".

In 1D we adopt the following shorthand notations to design the coefficients themselves:

$$\overline{ji} := \overline{f}_{j,i}, \tag{II.2.1}$$

$$\widetilde{ji} := \widetilde{f}_{j,i+(S-2)}. \tag{II.2.2}$$

Note the shift by $(S-2)$ of the wavelet coefficients, which will be understood later.

To store a d-dimensional array A in memory, one has to define a relationship between the d-dimensional index $i_1, \ldots, i_d$ of the elements and their memory position p. All the indices start from zero. Denoting by $(N_1, \ldots, N_d)$ the size of the array, we adopt the following convention:

$$p = i_1 + N_1 i_2 + \ldots + (N_1 \ldots N_{d-1}) i_d.$$

The corresponding element is then denoted $A(i_1, i_2, \ldots, i_d)$. The first direction is called the "major" direction, and the last one is called the "minor" direction.

II.2.2.2 Memory layout

Several memory layouts have been proposed for storing wavelet coefficients. In 1D, the most common choice is to group together coefficients of the same scale, as follows:

Input	$\overline{30}$	$\overline{31}$	$\overline{32}$	$\overline{33}$	$\overline{34}$	$\overline{35}$	$\overline{36}$	$\overline{37}$
Stage 1	$\overline{20}$	$\overline{21}$	$\overline{22}$	$\overline{23}$	$\widetilde{20}$	$\widetilde{21}$	$\widetilde{22}$	$\widetilde{23}$
Stage 2	$\overline{10}$	$\overline{11}$	$\widetilde{10}$	$\widetilde{11}$	$\widetilde{20}$	$\widetilde{21}$	$\widetilde{22}$	$\widetilde{23}$
Stage 3	$\overline{00}$	$\widetilde{00}$	$\widetilde{10}$	$\widetilde{11}$	$\widetilde{20}$	$\widetilde{21}$	$\widetilde{22}$	$\widetilde{23}$

however this approach has several disadvantages:

- the transform cannot be done in place
- the space locality of the wavelet coefficients is not reflected in their memory ordering, since coefficients that are very close in space but at different scales can be stored very far in memory.

To solve these two issues we have adopted the following *interlaced* storage order:

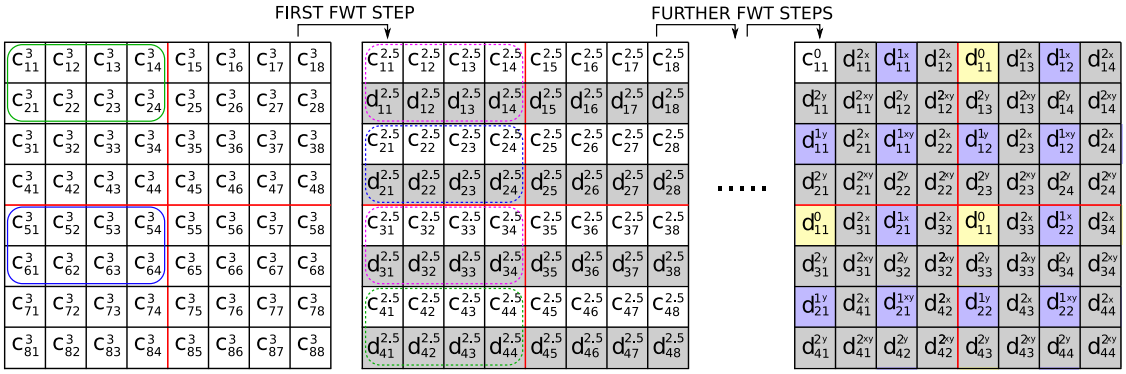


Figure II.2.1: Schematic view of the ordering of the wavelet coefficients before and after transformation. Red lines: possible domain decomposition for distributed memory parallelization involving 4 processes. Round edged boxes: communication between processes to compute the first step (see text). Here, the notations c and d design respectively the scaling and wavelet coefficients, the exponents indicate the scale and the two subscripts correspond to the positions i_1 and i_2.

Stage 0	$\overline{30}$	$\overline{31}$	$\overline{32}$	$\overline{33}$	$\overline{34}$	$\overline{35}$	$\overline{36}$	$\overline{37}$
Stage 1	$\overline{20}$	$\widetilde{20}$	$\overline{21}$	$\widetilde{21}$	$\overline{22}$	$\widetilde{22}$	$\overline{23}$	$\widetilde{23}$
Stage 2	$\overline{10}$	$\widetilde{20}$	$\widetilde{10}$	$\widetilde{21}$	$\overline{11}$	$\widetilde{22}$	$\widetilde{11}$	$\widetilde{23}$
Stage 3	$\overline{00}$	$\widetilde{20}$	$\widetilde{10}$	$\widetilde{21}$	$\widetilde{00}$	$\widetilde{22}$	$\widetilde{11}$	$\widetilde{23}$

The disadvantage of the new layout is that strided memory access is required to perform the filter at each stage except the first. We do not expect this to cost too much computationally since most of the time is spent on the first few stages of the transform, especially in higher dimensions.

The generalization of the layout to higher dimensions is obtained by storing the input scaling coefficients in a block n-dimensional array and by applying each stage of the 1D wavelet transform along each direction before going to the next stage (see Fig. II.2.1 for a 2D example).

In the DTCWT case, the coefficients corresponding to the 2^d possible values of the additional parameter $\boldsymbol{\gamma}$ are stored contiguously in memory, and the interlaced layout is utilized to arrange the resulting groups of coefficients.

II.2.2.3 In-place algorithm

Let us first consider the 1D case. The input scaling coefficients are stored in memory according to the layout described above, and the array is denoted $(A_i)_{0\leq i\leq N-1}$. The size of A can necessarily be put under the form $N = m2^J$, where m is an integer. To compute level l of the FWT, Algorithm 1 below follows immediately from the relationships between the scaling and wavelet coefficients at neighboring scales, (II.1.20-II.1.21). It is necessary to take care of periodization, because the computed coefficients at the end of the array depend on the input coefficients that were initially located at the beginning. These coefficients have to be stored beforehand in a buffer, called B. The amount of memory which is required in addition to the input array is therefore S, i.e. the contents of the buffer B, plus the temporary

Algorithm 1 Level l for 1D in-place FWT of array $(A_i)_{0\leq i\leq N-1}$

$s \leftarrow 2^l$
for $i = 0$ to $S-3$ **do**
  $B(i) \leftarrow A(si)$
end for
for $i = 0$ to $N - sS$ step $2s$ **do**
  $a \leftarrow \sum_{k=0}^{S_g-1} g[k]A(i+sk)$
  $b \leftarrow \sum_{k=0}^{S_h-1} h[k]A(i+sk)$
  $A(i) \leftarrow a$
  $A(i+s) \leftarrow b$
end for
for $i = N/s - S + 2$ to $N/s - 1$ step 2 **do**
  $a \leftarrow \sum_{k=0}^{\frac{N}{s}-i-1} g[k]A(s(i+k)) + \sum_{k=\frac{N}{s}-i}^{S_g-1} g[k]B(k-\frac{N}{s}+i)$
  $b \leftarrow \sum_{k=0}^{\frac{N}{s}-i-1} h[k]A(s(i+k)) + \sum_{k=\frac{N}{s}-i}^{S_h-1} h[k]B(k-\frac{N}{s}+i)$
  $A(si) \leftarrow a$
  $A(s(i+1)) \leftarrow b$
end for

variables a and b. The number of multiplications is $S2^{-l}N$. Note that all indices are to be taken modulo N, and recall that S is the length of the wavelet filter. The correctness of this algorithm relies on the shift by $(S-2)$ in (II.2.2). Without this shift, it would not be possible to derive such a simple in-place algorithm. To compute the full FWT, Algorithm 1 is repeated for $l = 0, \ldots, J-1$. For the inverse wavelet transform, which we omit here, the procedure is roughly the same except that the coefficients must be traversed backwards to preserve the in-place property.

Algorithm 2 d-dimensional in-place FWT of $(A(\mathbf{i}))$

for $l = 0$ to $J-1$ **do**
  $s \leftarrow 2^l$
  for $k = 1$ to d **do**
    for $\mathbf{i'} \in \bigotimes_{p\neq k}\{0, \ldots, N_p/s - 1\}$ **do**
      apply Algorithm 1 to $(A(si'_1, \ldots, si'_{k-1}, i, si'_k, \ldots, si'_{d-1}))_{1\leq i\leq N}$
    end for
  end for
end for

In the d-dimensional case, each level of the wavelet transform is decomposed into d stages acting along each of the dimensions, as can be seen from (II.1.21), and similarly for the inverse wavelet transform (II.1.23). Using this property, the generalization of Algorithm 1 to d dimensions is straightforward, see Algorithm 2. A denotes the array containing the input coefficients, and we take J such that for all k, $N_k = m_k 2^J$, where m_k is an integer. The loop on i' in Algorithm 2 can be seen as a loop on 1D cuts through the array A in the direction k. The complexity of Algorithm 2 is :

$$Q = \sum_{l=0}^{J-1} dSN_1N_2\ldots N_k 2^{-dl} \simeq 2dSN_1N_2\ldots N_k$$

For illustrative purposes, we detail the procedure in the case of a 2D array of size 8×8. The memory locations of the scaling coefficients at the finest scale are shown in Fig. II.2.1 (left). The red lines should be ignored for now. The FWT algorithm starts by computing a single level of the filter bank on the columns of the matrix, using (II.1.20) and (II.1.21). The results of the computation are stored in-place, alternating between scaling coefficient and wavelet coefficient. Due to the periodic wrapping of the column, the leading $S-2$ coefficients at scale J are needed to compute the trailing $S-2$ coefficients at scale $J-1$. These coefficients should therefore not be overwritten right away, but held in memory until they are not needed anymore. This can be accomplished thanks to a small buffer of size $S-2$. No additional memory is required. The results of this very first step are shown in Fig. II.2.1 (middle). The exponents 2.5 indicate that we stand halfway through the first level of the transform. Once all the columns have been filtered comes the turn of the lines. They are processed in exactly the same way, except that the stride in memory is now much larger, which makes this step usually more expensive due to cache issues. After all array dimensions have been filtered, the input data has been replaced by scaling and wavelet coefficients at scale $J-1$. The scaling coefficients sit in the memory locations with even-numbered row and column indices. The remaining memory locations are filled with the wavelet coefficients at scale $J-1$, which do not need to be touched by any further steps of the algorithm. The next level in the filter bank can hence be computed exactly in the same way as the first, as long as only the even-numbered rows and columns of the original memory space are considered. Continuing this procedure until only one row and one column remain yields the fully wavelet transformed data, arranged as shown in Fig. II.2.1 (right).

The generalization to the DTCWT is relatively straightforward, based on what we have outlined at the end of Sec. II.1.2.1. The FWT algorithm is first applied in place to compute the coefficients with respect to the anisotropic wavelets (II.1.26), and the linear combination (II.1.25) is then utilized to deduce the coefficients in the DTCWT basis. Using this approach the total cost of the DTCWT would be 2^d times the one of a real FWT, plus the cost required to compute the final linear combinations. In fact a trick, proposed by N. Kingsbury, can be implemented to reduce the overhead factor to $d^{-1}(2^d-1)$ (e.g., 1.5 instead of 4 in 2D and 2.33 instead of 8 in 3D).

II.2.3 Adaptive wavelet transform

The implementation of the adaptive wavelet transform is much more challenging than the non-adaptive one, due to the structure of the index sets Λ and $\overline{\Lambda}$ (see Sec. II.1.2.3). Hence we consider here only the 1D case. It follows from the definition (II.1.44) that $\overline{\Lambda}$ has at least as many coefficients as Λ. Therefore, the memory requirements for the adaptive wavelet transform are determined by $\overline{\Lambda}$. The coefficients are stored in a contiguous block of memory, ordered according to the position in space of their corresponding scaling function, that is $2^{-j}i$. To describe the set $\overline{\Lambda}$ itself, we also store the memory locations where the scale of the coefficients changes.

For simplicity we describe the FWT and IWT algorithms in detail only in the simple example corresponding to Fig. II.1.4, and we restrict ourselves to a few remarks concerning the generalization to more complex structures. As usual we assume that $S=4$. The input scaling coefficients (corresponding to panel (f) in Fig. II.1.4) are ordered in memory as follows:

$$\overline{50}\ \overline{51}\ \overline{52}\ \overline{53}\ \overline{54}\ \overline{55}\ \overline{56}\ \overline{57}\ \overline{58}\ \overline{59}\ \overline{510}\ \overline{511}\ \overline{512}\ \overline{513}\ \overline{514}\ \overline{515}\ \overline{48}\ \overline{49}\ \overline{410}\ \overline{411}\ \overline{412}\ \overline{413}\ \overline{414}\ \overline{415}$$

For the first stage, only the coefficients at scale 5 need to be considered, and those at coarser scales are left untouched. To compute $\overline{4i}$, $\widetilde{4i}$ for $i \leq 6$, (II.1.20-II.1.21) can be directly applied, as in Algorithm 1:

$$\begin{aligned} \overline{40} &= h_0\overline{50} + h_1\overline{51} + h_2\overline{52} + h_3\overline{53} \\ \widetilde{40} &= g_0\overline{50} + g_1\overline{51} + g_2\overline{52} + g_3\overline{53}, \text{ etc.} \end{aligned}$$

yielding:

$$\overline{40}\ \widetilde{40}\ \overline{41}\ \widetilde{41}\ \overline{42}\ \widetilde{42}\ \overline{43}\ \widetilde{43}\ \overline{44}\ \widetilde{44}\ \overline{45}\ \widetilde{45}\ \overline{46}\ \widetilde{46}\ \overline{514}\ \overline{515}\ \overline{48}\ \overline{49}\ \overline{410}\ \overline{411}\ \overline{412}\ \overline{413}\ \overline{414}\ \overline{415}$$

To compute $\overline{47}$, we use the edge matrix procedure outlined in Sec. II.1.2.3. On the other hand, by looking at the red circles in Fig. II.1.4 (panel (e)), we see that $\widetilde{47} = 0$. Stage 1 is thus completed:

$$\overline{40}\ \widetilde{40}\ \overline{41}\ \widetilde{41}\ \overline{42}\ \widetilde{42}\ \overline{43}\ \widetilde{43}\ \overline{44}\ \widetilde{44}\ \overline{45}\ \widetilde{45}\ \overline{46}\ \widetilde{46}\ \overline{47}\ 0\ \overline{48}\ \overline{49}\ \overline{410}\ \overline{411}\ \overline{412}\ \overline{413}\ \overline{414}\ \overline{415}$$

Note the wasted memory location indicated by the presence of the zero right after $\overline{47}$, and due to the fact that $\overline{\Lambda}$ is bigger than Λ. For subsequent stages, we may simply apply the non-adaptive algorithm to the set of scaling coefficients at scales 4, 3, 2 and 1 successively, since they are all known when they are needed. The final memory layout is:

$$\overline{00}\ \widetilde{40}\ \widetilde{30}\ \widetilde{41}\ \widetilde{20}\ \widetilde{42}\ \widetilde{31}\ \widetilde{43}\ \widetilde{10}\ \widetilde{44}\ \widetilde{32}\ \widetilde{45}\ \widetilde{21}\ \widetilde{46}\ \widetilde{33}\ \widetilde{47}\ \widetilde{00}\ \widetilde{34}\ \widetilde{22}\ \widetilde{35}\ \widetilde{11}\ \widetilde{36}\ \widetilde{23}\ \widetilde{37}$$

For the inverse transform, the adaptivity influences only the last stage as well, and even there it poses no specific difficulty, since we have everything we need to directly apply (II.1.23). If the sets Λ and $\overline{\Lambda}$ have a more complex structure, the same simple rules can be applied. Condition (II.1.38) ensures that the edge matrix is sufficient to resolve any indeterminacy that may arise during the process. Note that the more Λ is complex, the more the memory traversal pattern becomes complicated, especially during the intermediate stages of the transform, when the wavelet coefficients stored from previous stages have to be jumped over. Extensive studies will be necessary to determine the impact on the performance of the algorithm, as compared to other approaches using pointers.

II.2.4 Parallelization

Parallelization is considered here only for the non-adaptive wavelet transform. We adopt a hybrid MPI/OpenMP scheme, where the domain is, from the beginning, decomposed and distributed between processes communicating through MPI, and where the computations of each process are in turn shared between several threads using the OpenMP standard.

II.2.4.1 Shared memory OpenMP parallelization

The OpenMP parallelization layer is activated only for wavelet transforms involving two dimensions or more. According to Algorithm 2, to compute one stage of a forward or

inverse wavelet transform, a d-dimensional array has to be traversed successively along all the directions, and therefore a total of d times. When the array is traversed along direction k, the computations are reduced to a series of single-stage one dimensional wavelet transforms. The OpenMP parallelization consists in splitting the loop on i' into pieces, each piece being then attributed to a thread. The splitting is done along direction $k+1$ if $k < d$ and along direction 1 if $k = d$. We expect the speedup due to this parallelization layer to scale linearly with the number of threads, since there is no data dependency at all.

II.2.4.2 MPI parallelization

Algorithm 3 MPI-parallel 1D in-place FWT of $(A(i))$

for $l = 0$ to $J-1$ **do**
  $s \leftarrow 2^l$
  for $k = 1$ to d **do**
    Send data with `MPI_ISend` / `MPI_IRecv`
    for $\mathbf{i'} \in \bigotimes_{p \neq k} \{0, \ldots, N_p/s - 1\} \cap \{i_p^-, \ldots, i_p^+ - 1\}$ **do**
      A' points to $(A(si'_1, \ldots, si'_{k-1}, i, si'_k, \ldots, si'_{d-1}))_{i_k^- \leq i < i_k^+}$
      for $i = i_k^-$ to $i_k^+ - 1 - sS$ step $2s$ **do**
        $a \leftarrow \sum_{k=0}^{S_g - 1} g[k] A'(i + sk)$
        $b \leftarrow \sum_{k=0}^{S_h - 1} h[k] A'(i + sk)$
        $A'(i) \leftarrow a$
        $A'(i + s) \leftarrow b$
      end for
    end for
    Wait for data with `MPI_Wait`
    B is such that $B(\mathbf{i}) = A(si_1, \ldots, si_k - i_k^+, \ldots)$ for $0 \leq i_k < S - 2$
    for $\mathbf{i'} \in \bigotimes_{p \neq k} \{0, \ldots, N_p/s - 1\} \cap \{i_p^-, \ldots, i_p^+\}$ **do**
      A' points to $(A(si'_1, \ldots, si'_{k-1}, i, si'_k, \ldots, si'_{d-1}))_{i_k^- \leq i \leq i_k^+}$
      B' points to $(B(i'_1, \ldots, i'_{k-1}, i, i'_k, \ldots, i'_{d-1}))_{0 \leq i \leq 2S-5}$
      $q \leftarrow 0$
      for $i = i_k^- - sS + 2$ to $i_k^+ - 1$ step 2 **do**
        $a \leftarrow \sum_{k=0}^{i_k^+ - i - 1} g[k] A'(i + sk) + \sum_{k=i_k^+ - i}^{S_g - 1} g[k] B'(k - N/s + q)$
        $b \leftarrow \sum_{k=0}^{i_k^+ - i - 1} h[k] A'(i + sk) + \sum_{k=i_k^+ - i}^{S_h - 1} h[k] B'(k - N/s + q)$
        $A'(si) \leftarrow a$
        $A'(s(i+1)) \leftarrow b$
        $q \leftarrow q + 1$
      end for
    end for
  end for
end for

When the MPI layer is activated, we assume that the input data has already been distributed among processors, according to a certain domain decomposition pattern. Each piece of the global array A is assumed to be of the form $(A(i_1, \ldots, i_d))$ where each i_k varies in a certain interval $\{i_k^-, \ldots, i_k^+ - 1\}$. The intervals should have even length, but they are otherwise arbitrary.

The loops on l and k in Algorithm 2 are not affected by the parallelization, and the loop in i' is simply restricted to the local array, $\left\{i_k^-, \ldots, i_k^+ - 1\right\}$. The main difficulty is to parallelize Algorithm 1. The idea is to make use of the buffer B to communicate between processors. However, inserting the MPI calls directly within Algorithm 1 could be very inefficient because it would imply the exchange of many very small messages of size $(S-2)$. To avoid this issue, the idea is to merge the messages into bigger pieces, as summarized by Algorithm 3. We do not enter into the details of the exact contents of the messages that must be transferred between processors, but they are such that the condition indicated at line 15 is satisfied. Under that condition each process has all the input it needs to compute the requested wavelet and scaling coefficients. A major advantage of this communication scheme is that it allows to overlap communication and computation, since the data can be handed over to MPI before the main loop (line 4) and received after it has completed (line 14). As a counterpart, the necessary buffers become large when computing the coarse levels of the transform, which limits the parallel efficiency (See Sec. II.3.1).

For clarity we now come back to our 2D example (Fig. II.2.1) and describe its parallelization using MPI. We assume that the initial data is split between 4 processes, as indicated by the red lines in Fig. II.2.1 (left), that $S = 4$, and that the sizes of the subdomains belonging to each process are equal, and are powers of 2.. Processes are numbered in row-major order. To compute the coefficients in the dotted blue box (Fig. II.2.1) , all the coefficients in the solid blue box need to be sent from process 2 to process 1. Because of periodicity, all the coefficients in the solid green box need to be sent from process 1 to process 2. The same kind of communication occurs between processes 3 and 4 during this step. The coefficients in the magenta boxes can be computed without waiting for the communications to be completed. For large array sizes, they will correspond to the majority of coefficients, ensuring a good overlap of computation and communication.

II.3 Verification and benchmarking

In this section, we report the results of a few test cases that required particular attention.

II.3.1 Parallel efficiency

The consumed CPU time for a forward wavelet transform using the algorithm described in Sec. II.2.4, with Daubechies filters of length 4, was measured on an IBM Regatta Power6 machine. The results are shown in Fig. II.3.1. One observes near perfect speedup up to 32 processes. The sudden degradation of efficiency when going from 32 to 64 processes is mostly due to hardware limitations, since each node on the machine contains 32 Power6 cores. Simulations with a higher number of cores require communication through a different network, which has a longer latency.

We have also performed benchmarks on a BlueGene/P machine, and the results are reported in Fig. II.3.2. This machine is made of basic cubes, each one containing 1024 nodes. Each node is made of 4 CPUs sharing $2Gb$ of memory. The machine can be used in three different modes, virtual node (VN) where each process is attributed to one CPU, dual mode where each process is attributed to two CPUs, and SMP mode, where each process is attributed to four CPUs. For the tests reported here, only the VN and SMP modes are

considered, and no OpenMP parallelization was used. The algorithm shows good scaling results for 3D arrays using up to 1024 processes. In 2D, for an array of size 16384×16384, we observe near perfect scaling up to 1024 processors in SMP mode, and an efficiency of 0.5 up to 1024 processors in VN mode.

The results of this section demonstrate the good scaling properties of the FWT algorithms we have presented, on two quite different machines, and for large datasets.

II.3.2 Representation of translation operators

We now make a brief digression in order to show the limitations of translation invariance in complex-valued wavelet families. Moreover, we will derive a formula that will be useful for the next section. Here, φ can be any square integrable, complex-valued, 1-periodic function on $\mathbb{R}$, and not necessarily a scaling function. Our goal is to calculate the matrix of a translation operator with respect to a family of evenly spaced translates of φ. We look for conditions under which the subspace spanned by these integer translates is stable, or approximately stable, under all translations.

II.3.2.1 Expansion over a family of translates

For $\xi \in \mathbb{T}$ and $f \in L^2(\mathbb{T})$, let Tf be defined by:

$$\forall x \in \mathbb{R}, (T_\xi f)(x) = f(x - \xi). \tag{II.3.1}$$

To study T it is convenient to introduce the Fourier modes $(c_k)_{k\in\mathbb{Z}}$, defined for $x \in \mathbb{T}$ by $c_k(x) = e^{2\imath\pi kx}$, and which make up an orthogonal basis of $L^2(\mathbb{T})$. Denoting by $(\widehat{f_k})_{k\in\mathbb{Z}}$ the Fourier coefficients of a function f, T is thus characterized by

$$\widehat{Tf}_k = e^{2\imath\pi k\xi}\widehat{f}_k. \tag{II.3.2}$$

Now let M be any positive integer, define for $0 \leq m < M$ $\varphi_m = T_{m/M}\varphi$, and consider the mapping F from $L^2(\mathbb{T})$ to $\mathbb{C}^M$ which is characterized by $(Ff)[m] = \langle f \mid \varphi_m\rangle$, where the elements of a sequence $a \in \mathbb{C}^M$ are denoted $a[m]$, $m = 0, \ldots, M-1$. Using Parseval's identity and (II.3.2) we get

$$(Ff)[m] = \langle \widehat{f} \mid \widehat{\varphi_m}\rangle = e^{-2\imath\pi k\frac{m}{M}}\langle \widehat{f} \mid \widehat{\varphi}\rangle, \tag{II.3.3}$$

where we have denoted by $\langle \cdot \mid \cdot \rangle$ the canonical scalar product on $l^2(\mathbb{Z})$. Now denoting by $\epsilon_\alpha[m] = \frac{1}{M}e^{2\imath\pi\alpha\frac{m}{M}}$ for $\alpha \in \mathbb{Z}$ the discrete Fourier basis in $\mathbb{C}^M$, we obtain a matrix representation of the operator $F : L^2(\mathbb{T}) \to \mathbb{C}^M$:

$$\langle Fe_k \mid \epsilon_\alpha\rangle = \langle \widehat{e_k} \mid \widehat{\varphi}\rangle \sum_{m=0}^{M-1} \frac{1}{M} e^{2\imath\pi(\alpha-k)\frac{m}{M}}.$$

Denoting momentarily by $S_{k\alpha}$ the Poisson sum on the right hand side, we have the classical result that $S_{k\alpha} = 1$ if $k - \alpha$ is a multiple of M, and $S_{k\alpha} = 0$ otherwise. Therewith:

$$\langle Fe_k \mid \epsilon_\alpha\rangle = S_{k\alpha}\overline{\widehat{\varphi}_k} \tag{II.3.4}$$

and moreover, letting $F^* : \mathbb{C}^M \to L^2(\mathbb{T})$ be the adjoint of F, we have

$$\begin{aligned}\langle F^*Fe_k \mid e_{k\prime}\rangle &= \langle Fe_k \mid Fe_{k\prime}\rangle = \sum_{\alpha=-M/2+1}^{M/2} \langle Fe_k \mid \epsilon_\alpha\rangle\langle\epsilon_\alpha \mid Fe_{k\prime}\rangle \qquad \text{(II.3.5)}\\ &= \sum_{\alpha=-M/2+1}^{M/2} S_{k\alpha}\overline{\widehat{\varphi}_k}S_{k\prime\alpha}\widehat{\varphi}_{k\prime} \qquad \text{(II.3.6)}\\ &= \overline{\widehat{\varphi}_k}\widehat{\varphi}_{k\prime} \text{ if } M \mid k-k' \text{ and } 0 \text{ otherwise.} \qquad \text{(II.3.7)}\end{aligned}$$

II.3.2.2 Translation operators and aliasing terms

We are looking for a representation of T_ξ in $\mathbb{C}^M$, that is, a mapping $T_\xi^+ : \mathbb{C}^M \to \mathbb{C}^M$ so that the following diagram commutes:

$$\begin{array}{ccc} E & \xleftarrow{F^*} & V \\ \downarrow{\scriptstyle T_\xi} & & \downarrow{\scriptstyle T_\xi^+} \\ E & \xrightarrow{F} & V \end{array}$$

An obvious solution to this problem would be given by $T_\xi^+ = FT_\xi F^*$. Let us first examine the outcome of this choice.

$$\begin{aligned}\langle FT_\xi F^*\epsilon_\beta \mid \epsilon_\alpha\rangle &= \langle T_\xi F^*\epsilon_\beta \mid F^*\epsilon_\alpha\rangle = \sum_{k\in\mathbb{Z}}\langle T_\xi F^*\epsilon_\beta \mid e_k\rangle\langle e_k \mid F^*\epsilon_\alpha\rangle\\ &= \sum_{k\in\mathbb{Z}} e^{2\imath\pi k\xi}\langle\epsilon_\beta \mid Fe_k\rangle\langle Fe_k \mid \epsilon_\alpha\rangle = \sum_{k\in\mathbb{Z}} e^{2\imath\pi k\xi}S_{k\beta}\widehat{\varphi}_k S_{k\alpha}\overline{\widehat{\varphi}_k}\\ &= \delta_{\alpha\beta}\sum_{p\in\mathbb{Z}} e^{2\imath\pi(\alpha+Mp)\xi}|\widehat{\varphi}[\alpha+Mp]|^2,\end{aligned}$$

where $\delta_{\alpha\beta} = 1$ if $\alpha = \beta$ and 0 otherwise. If the (φ_m) are pairwise orthogonal, then for all α (Mallat, 1999, p.222):

$$\sum_{p\in\mathbb{Z}} |\widehat{\varphi}[\alpha+Mp]|^2 = 1 \qquad \text{(II.3.8)}$$

so that $FT_0F^* = FF^* = Id_V$. More generally, the condition $T_0^+ = Id_V$ is natural to impose. However if φ is a complex Kingsbury scaling function, this will not be the case. Instead, $\varphi = \varphi^R + i\varphi^I$ where $(\varphi_m^R)_{0\le m\le M-1}$ and $(\varphi_m^I)_{0\le m\le M-1}$) are orthogonal families, but in general $\langle\varphi_m^R \mid \varphi_{m\prime}^I\rangle \neq 0$. This is the main difference with the Fourier case, where the $\sin$ and $\cos$ families are completely orthogonal to each other. To overcome this difficulty, we need to consider another lifting of T_ξ. Let us denote by σ the complex conjugation with respect to the canonical basis in V. One has:

$$\begin{aligned}\langle\sigma FT_\xi F^*\sigma\epsilon_\beta \mid \epsilon_\alpha\rangle &= \overline{\langle FT_\xi F^*\epsilon_{-\beta} \mid \epsilon_{-\alpha}\rangle} = \overline{\delta_{\alpha\beta}\sum_{p\in\mathbb{Z}} e^{2\imath\pi(-\alpha+Mp)\xi}|\widehat{\varphi}[-\alpha+Mp]|^2}\\ &= \delta_{\alpha\beta}\sum_{p\in\mathbb{Z}} e^{2\imath\pi(\alpha+Mp)\xi}|\widehat{\varphi}[-\alpha-Mp]|^2\end{aligned}$$

Defining now $T_\xi^+ = \frac{1}{2}(FT_\xi F^* + \sigma FT_{-\xi}F^*\sigma)$, one has:

$$\langle T_\xi^+ \epsilon_\beta \mid \epsilon_\alpha \rangle = \delta_{\alpha\beta}\frac{1}{2}\sum_{p\in\mathbb{Z}} e^{+2\iota\pi(\alpha+Mp)\xi}|\widehat{\varphi}[+\alpha+Mp]|^2 + e^{-2\iota\pi(\alpha+Mp)\xi}|\widehat{\varphi}[-\alpha-Mp]|^2. \quad \text{(II.3.9)}$$

Now putting $\xi = 0$:

$$\begin{aligned}
\langle T_0^+ \epsilon_\beta \mid \epsilon_\alpha \rangle &= \delta_{\alpha\beta}\frac{1}{2}\sum_{p\in\mathbb{Z}} |\widehat{\varphi}[+\alpha+Mp]|^2 + |\widehat{\varphi}[-\alpha-Mp]|^2 \\
&= \delta_{\alpha\beta}\frac{1}{2}\sum_{p\in\mathbb{Z}} \left|\widehat{\varphi^R}[+\alpha+Mp]\right|^2 + \left|\widehat{\varphi^I}[+\alpha+Mp]\right|^2 \\
&+ 2\mathrm{Im}\left(\widehat{\varphi^R}[+\alpha+Mp]\overline{\widehat{\varphi^I}[+\alpha+Mp]}\right) + \left|\widehat{\varphi^R}[-\alpha-Mp]\right|^2 + \left|\widehat{\varphi^I}[-\alpha-Mp]\right|^2 \\
&+ 2\mathrm{Im}\left(\widehat{\varphi^R}[-\alpha-Mp]\overline{\widehat{\varphi^I}[-\alpha-Mp]}\right) \\
&= \delta_{\alpha\beta}\sum_{p\in\mathbb{Z}} \left|\widehat{\varphi^R}[+\alpha+Mp]\right|^2 + \left|\widehat{\varphi^I}[+\alpha+Mp]\right|^2.
\end{aligned}$$

If $2\varphi^R$ and $2\varphi^I$ independently satisfy (II.3.8), as in the case of a DTCWT pair, then we finally have $T_0^+ = Id_V$.

More generally, the operator T_ξ^+ defined by (II.3.9) tells us how the coefficients of a function over the family of translates of φ is transformed under the action of the translation operator T_ξ. For an arbitrary choice of φ, this process will entail some loss of information, but for certain clever choices of φ, this loss of information can be made minimal. This is the meaning of "translation invariance" in our context. The goal of the construction of complex-valued wavelets is precisely to maximize translation invariance, by tailoring $\widehat{\varphi}$ so that only one term dominates all the others in the sum (II.3.9). In the next paragraph, we give actual examples of the application of T_ξ^+ to some functions, in order to assess the translation invariance properties of these wavelets.

II.3.2.3 Numerical examples

In this section, we use the above expression (II.3.9) to translate functions numerically. On Fig.II.3.3 we plot the phase and modulus of a translation filter computed for the Q-Shift D wavelet with $\xi = 1/8th$ of a sample. We see in particular that the phase is not linear but seems to have some sort of cosine shape. The modulus is not constant equal to one, which it would have to be if the transform was perfectly translation invariant. On Fig.II.3.4, the phase of the Fourier coefficient of a discrete Dirac delta function is plotted as a function of wave-number, for increasing translation parameters. Perfect translation at all wave-numbers would correspond to a linear increase of the phase. On Fig.II.3.5, we show the functions obtained by repeatedly applying a translation operator with a very small parameter to a Gaussian function. The dispersion after each translation is very small, but it accumulates and thus slowly deteriorates the result.

II.3.3 Power spectrum estimation

II.3.3.1 Statement of the problem

Given one realization of a zero-mean discrete stationary Gaussian process $(X_n)_{0\leq n\leq N-1}$ $(N = 2^J)$, one would like to find out some information about its power spectrum $\hat{c}(k)$, where the covariance function c is defined by

$$c(p) = \mathbb{E}(X_n X_{n+p}).$$

A popular solution when the process is known to posses some self-similarity properties is to use the discrete wavelet coefficients to perform this estimation. In particular the power spectrum is related to the variance of the wavelet coefficients :

$$\widetilde{X}_{jb} = \langle X \mid \psi_{jb} \rangle.$$

For example, in (Kaplan and Kuo, 1996; Abry et al., 1993; Abry and Veitch, 1998; Tewfik and Kim, 1992) orthonormal wavelet decompositions are used to estimate the Hurst exponent of fractional Brownian motions. In particular it is shown that if the mother wavelet has enough vanishing moments, the variation of $\sigma^2_{(j)}$ with j is closely related to the spectral behavior of the process. The wavelet coefficients at different scales and/or positions are almost uncorrelated, and bounds for the residual correlation are given. Then an estimator for the Hurst exponent H is constructed, and the variance of this estimator is studied. However the results provided are asymptotic and assume complete uncorrelation between the wavelet coefficients which is never exactly the case. As one can expect intuitively, more weight has to be given to finer scales because they benefit from a larger number of orthogonal wavelets. In the following, we propose to give some quantitative results concerning the estimation of $\sigma^2_{(j)}$ with known wavelet families. In particular, we are interested in the comparison between orthogonal wavelet families and DTCWT complex-valued wavelets.

In the following we will drop the index j and consider the more general problem of estimating

$$\sigma^2 = \mathbb{E}\left(|\langle X \mid \psi \rangle|^2\right), \tag{II.3.10}$$

where ψ is any given function in $L^2(\mathbb{T})$, normalized to $\|\psi\|_2 = 1$. By the Parseval identity, σ^2 can be related to the power spectrum of the process:

$$\sigma^2 = \frac{1}{N} \sum_{k\in\mathbb{Z}} \widehat{c}_k \left|\widehat{\psi}_k\right|^2,$$

and for $\xi \in \mathbb{R}$, we also have thanks to the stationarity of X:

$$\sigma^2 = \mathbb{E}(|\langle X \mid T_\xi \psi \rangle|^2) \tag{II.3.11}$$

where T_ξ was introduced in the previous section. In the following we study estimators for σ^2 of the form

$$\Sigma^2 = \frac{1}{2^j} \sum_{i=0}^{2^j-1} |\langle X \mid T_{2^{-j}i} \psi \rangle|^2. \tag{II.3.12}$$

II.3.3.2 Variance of Σ^2

The estimator Σ^2 is always unbiased, because

$$\mathbb{E}\left(\Sigma^2\right) = \frac{1}{2^j}\sum_{i=0}^{2^j-1}\mathbb{E}\left(|\langle X \mid T_{2^{-j}i}\psi\rangle|^2\right) = \sigma^2, \tag{II.3.13}$$

where we have made use of (II.3.11). The statistical uncertainty associated to Σ^2 can be quantified by computing its variance, denoted $\mathrm{Var}\left(\Sigma^2\right)$. Using Parseval again, we can rewrite Σ^2 as:

$$\Sigma^2 = \frac{1}{2^j}\sum_{i=0}^{2^j-1}\left\langle \widehat{X} \mid \widehat{T_{2^{-j}i}\psi}\right\rangle \overline{\left\langle \widehat{X} \mid \widehat{T_{2^{-j}i}\psi}\right\rangle}, \tag{II.3.14}$$

or in a more condensed form:

$$\Sigma^2 = \frac{1}{2^j}\widehat{X}^\dagger F^\dagger F \widehat{X} \tag{II.3.15}$$

where $\cdot^\dagger$ denotes the Hermitian conjugate of a matrix, and F is the matrix defined by $(F\widehat{X})_i = \left\langle \widehat{X} \mid \widehat{T_{2^{-j}i}\psi}\right\rangle$, that we have already encountered in the previous section. As the process is stationary and periodic, we known that the discrete Fourier transform achieves a Karhunen-Loeve decomposition, which means that the components of the vector $\widehat{X}$ are independent random variables. A classical theorem about the variance of a quadratic form evaluated for a jointly Gaussian vector hence gives us:

$$\mathrm{Var}\left(\Sigma^2\right) = \frac{2}{2^{2j}}\mathrm{Tr}\left(\left(F^\dagger F \mathrm{Cov}\left(\widehat{X}\right)\right)^2\right)$$

where the diagonal matrix $\mathrm{Cov}\left(\widehat{X}\right)$ is the covariance of $\widehat{X}$, or in other words the power spectrum of the process, $(\hat{c}[k])$, that we have already encountered. Using the result (II.3.5) from the previous section, we obtain:

$$\mathrm{Var}\left(\Sigma^2\right) = 2\sum_{k\in\mathbb{Z}}|\widehat{c}_k|^2\left|\widehat{\psi}_k\right|^2\sum_{\gamma\in\mathbb{Z}}\left|\widehat{\psi}_{k+2^j\gamma}\right|^2 \tag{II.3.16}$$

which is the essential result of this section. The terms corresponding to $\gamma \neq 0$ in this expression are called aliasing term, and they should be made as small as possible if we want the variance of Σ^2 to be small.

II.3.3.3 Behavior of the variance

There are cases where the aliasing terms in (II.3.16) are exactly or almost equal to zero, so that considering more and more translated version of our function doesn't bring any improvement to the variance estimator. The most simple one arises when we take a Fourier mode:

$$\psi(x) = \exp(2\iota\pi k_0 x),$$

for which the aliasing products $\left|\widehat{\psi}_k\right|^2\left|\widehat{\psi}_{k+2^j\gamma}\right|^2$ are zero whatever the values of k_0 and γ and the variance of the estimator is always simply given by

$$\mathrm{Var}(\Sigma^2) = 2\left|\widehat{c}_{k_0}\right|^2.$$

	Haar	Coiflet 12	Coiflet 18	Daub. 4	Daub. 8	Daub. 12	QShift-B
Σ	0.123	0.125	0.125	0.125	0.125	0.125	0.105
r	3.00	1.57	1.43	2.03	1.60	1.45	1.003

Table II.3.1: Deterioration of the variance estimator due to lack of translation invariance, for white noise and different wavelets.

The complex-valued wavelets corresponding to the dual tree complex wavelet transform (DTCWT) (Kingsbury, 2001) possess the key shift-invariance property that makes the aliasing products almost equal to zero (Kingsbury, 2000, 1999). We shall see numerical examples of that phenomenon below. For that purpose, let r_j be defined by:

$$r = \frac{\sum_{k\in\mathbb{Z}} |\widehat{c}_k|^2 \left|\widehat{\psi}_k\right|^2 \sum_{\gamma=0}^{2^j-1} \left|\widehat{\psi}_{k+2^j\gamma}\right|^2}{\sum_{k\in\mathbb{Z}} |\widehat{c}_k|^2 \left|\widehat{\psi}_k\right|^4},$$

which quantifies the increase of variance due to the presence of aliasing terms.

II.3.3.4 Numerical examples

We now report results on the evaluation of Σ and r for some popular wavelets. They are summarized in table II.3.1. The estimation is made at scale $j = 7$ (which contains 128 wavelet coefficients), for a white noise ($c_k = 1$ for all k). We see that smaller values of r are attained by higher order wavelets. This is explained by the property that higher order wavelets are more localized in Fourier space, and thus less prone to aliasing effects. On the other hand, the corresponding wavelets have a longer support, and the wavelet coefficients are thus affected by more fluctuations, which explains why the value of Σ is close to 0.125 for all wavelet families. The only exception is the QShift-B wavelet, for which the standard deviation is reduced to 0.105. From these results, we conclude that Kingslets allow for a better estimation of the power spectrum of random processes, thanks to their translation invariance properties.

II.3.4 Denoising correlated noises

One of the important research subjects that we have explored is the denoising of turbulent vorticity fields under the assumptions that the noise is Gaussian and additive, but with as little restriction as possible on the correlation of the noise. To understand the inherent limitations implied by the presence of correlations, we have made some tests with academic signals which are reported here. We circumvent for a moment the problem of estimating the variance of the noise by using synthetic noise with a known variance. Even the constant q in (II.1.53) is determined from the data itself in order to minimize the risk. In this way we obtain a lower bound on the risk which is achievable in practice, and we concentrate on the effect due to the correlation of the noise, independently of the chosen algorithm. This procedure is called "oracle thresholding", see (Mallat, 1999).

We focus on two types of correlated noise that are idealized versions of those often encountered in practice, namely fractional Gaussian noise and exponentially correlated Gaussian

noise. The correlation function of a fractional Gaussian noise is given by

$$\mathbb{E}(X_H[0]X_H[n]) = \frac{\sigma^2}{2}(|n+1|^{2H} + |n-1|^{2H} - 2|n|^{2H}),$$

where H is called the Hurst exponent. Some examples of realizations are presented in Fig. II.3.6, along with their time integral, which are called fractional Brownian motions. The covariance of an exponentially correlated Gaussian noise is given by

$$\mathbb{E}\left(X[0]X[n]\right) = \sigma^2 \exp\left(-\frac{|n|}{\tau_C}\right),$$

where τ_C is called the correlation time.

The results, presented in Fig. II.3.8 show that in each case, the performance of denoising is strongly limited by the correlation in the noise. The limitation is observed when the Hurst exponent tends to 1, and when the correlation time exceeds a few percents of the total length of the signal. This can be heuristically understood by noticing that the more a noise is correlated, the more it will affect the wavelet coefficients at coarse scales. Since there are few wavelet coefficients at coarse scale, there is simply no room left for the signal to emerge out of the background noise. Anyway, as we have seen in the previous section, it would be very difficult to estimate precisely the variance of the noise at coarse scales even if it was separated from the signal, because the small number of available wavelet coefficients implies a large uncertainty on the estimation. Based on these simple tests, we have decided to systematically restrict the denoising to scales finer than some minimal scale j_0, whose value will be specified later on.

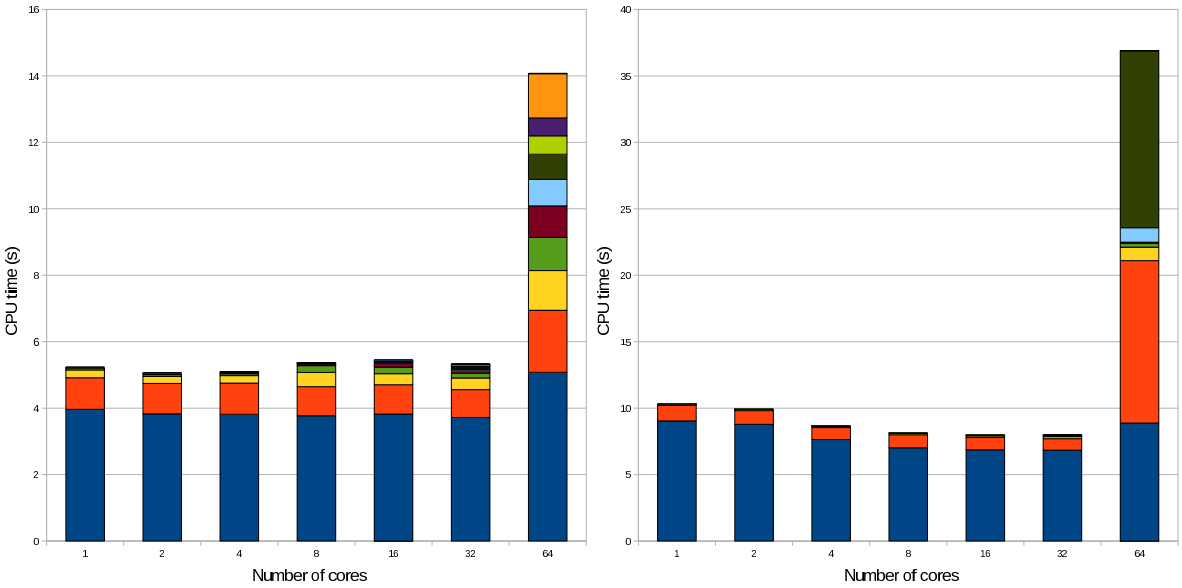


Figure II.3.1: Consumed CPU time for increasing number of MPI processes on the IBM Regatta Power6. Left: 4096×4096 2D array. Right: $256 \times 256 \times 256$ 3D array. The different colors indicate successive scales in the wavelet transform, starting from the finest scale at the bottom.

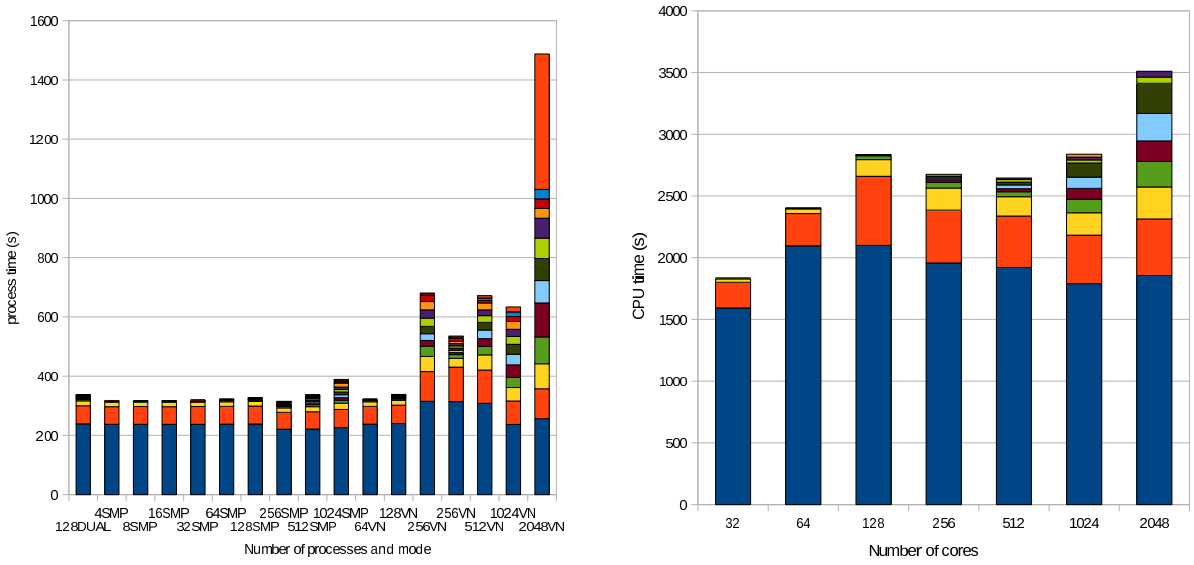


Figure II.3.2: FWT performance on IBM BlueGene machine. Left: CPU time consumed to compute one FWT of a 2D array of size 16384×16384, as a function of the number of processors and parallel mode (see text). Right: CPU time consumed to compute on FWT of a 3D array of size $1024 \times 1024 \times 1024$ using the SMP parallel mode. In both figures the various colors distinguish the contributions of the successive scales of the FWT, from the finest scale in blue at the bottom to the coarsest scale at the top.

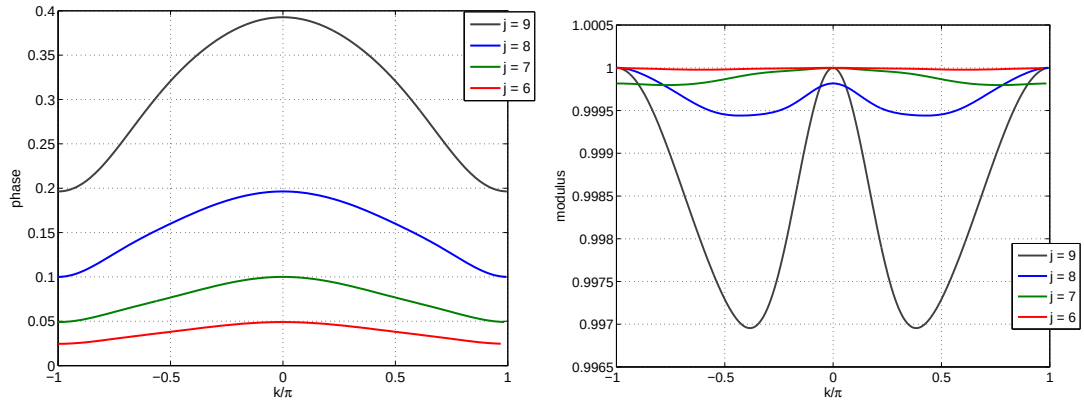


Figure II.3.3: Phase (left) and modulus (right) of the translating filter (II.3.9) for the qshift d wavelet at several scales. The translation parameter ξ equals one eighth of a sample.

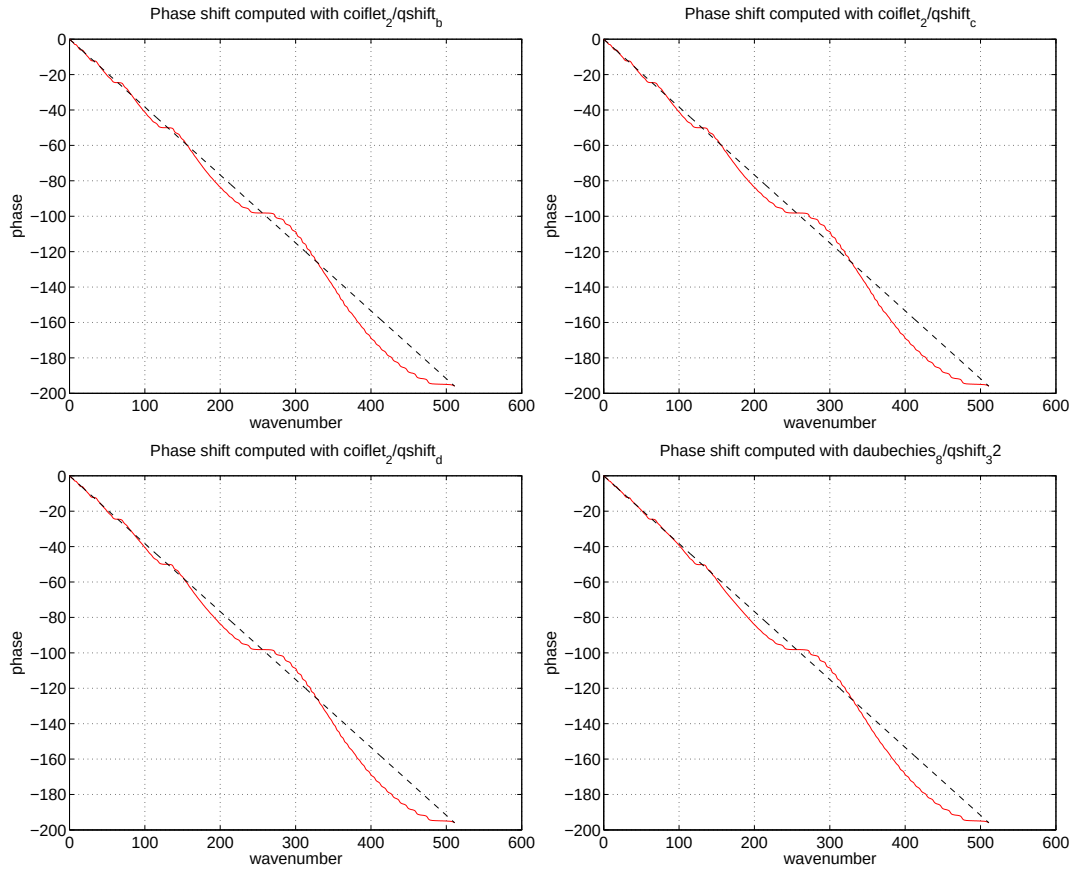


Figure II.3.4: The phase of the Fourier transform of a Dirac delta function at the origin, after translation. Top,left: QShift B, Top,right: QShift C, Bottom,left: QShift D, Bottom,right: QShift 32

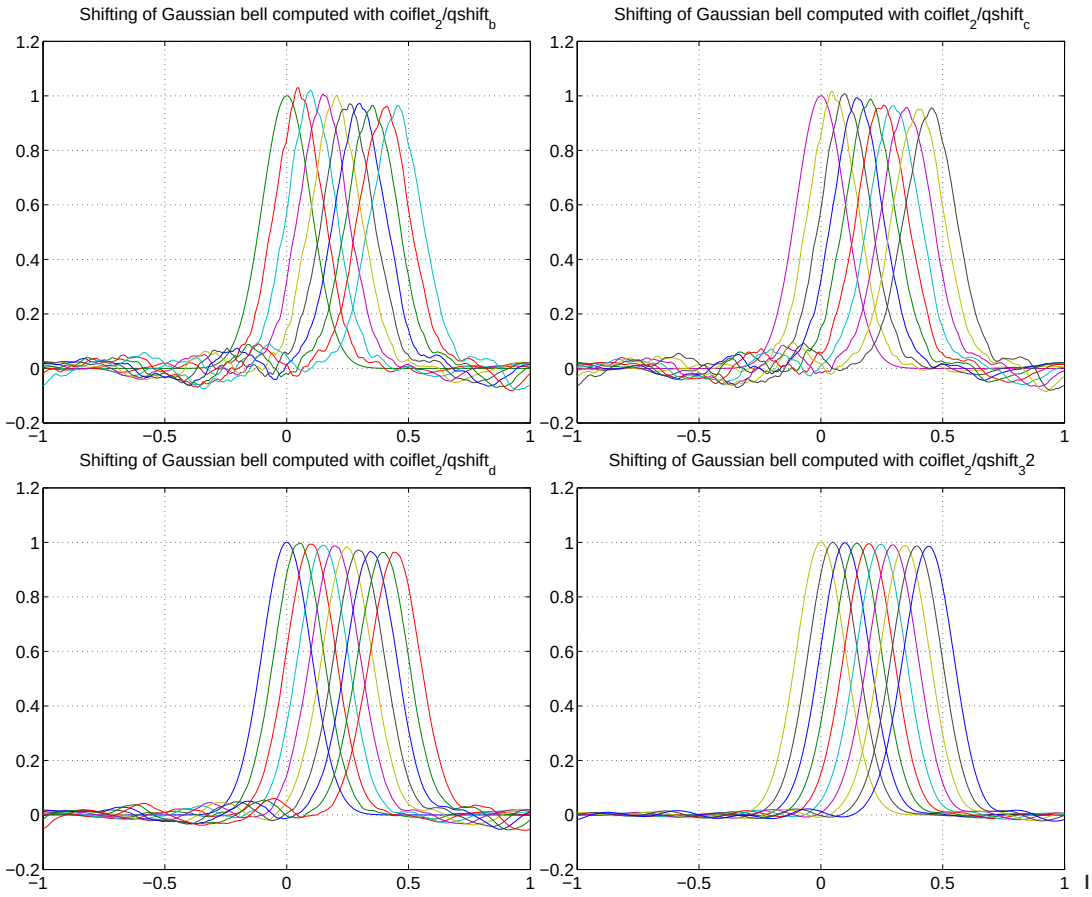


Figure II.3.5: The shifts of a Gaussian by increasing amounts, computed using (II.3.9) Top,left: QShift B, Top,right: QShift C, Bottom,left: QShift D, Bottom,right: QShift 32,

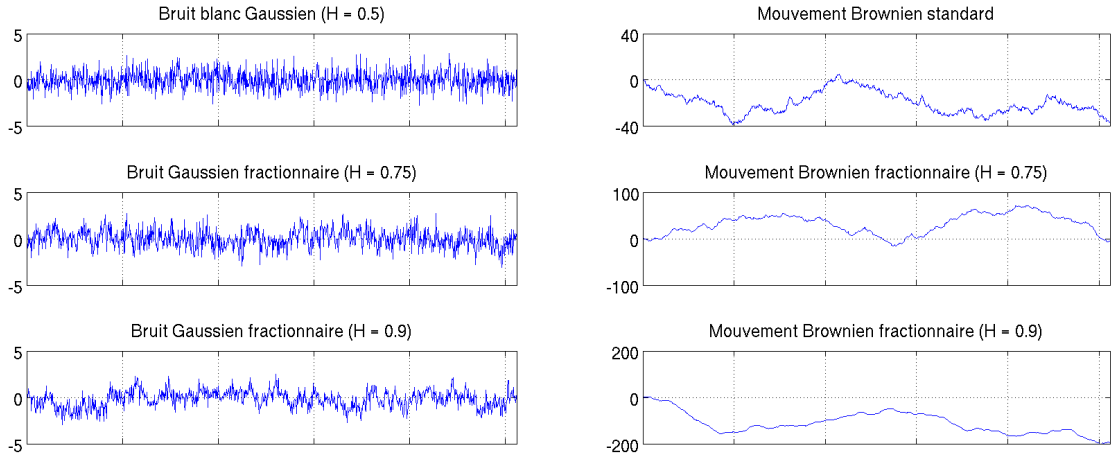


Figure II.3.6: Examples of realizations of fractional Gaussian noises (left) and the associated fractional Brownian motions (right) for several values of the Hurst exponent H. The number of samples is $N = 1024$.

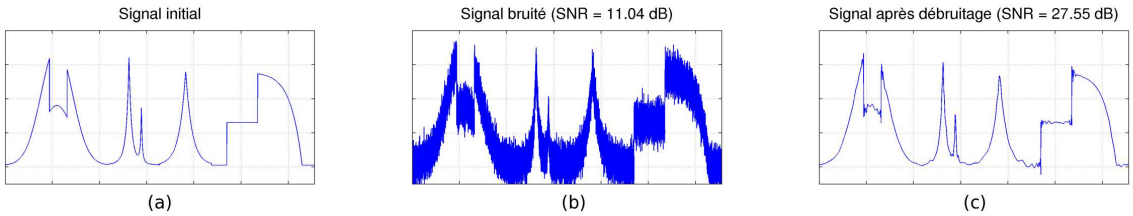


Figure II.3.7: Denoising of an academic signal corrupted by Gaussian white noise by thresholding of its wavelet coefficients. The procedure used to estimate the variance of the noise is the global CVE algorithm, which will be explained in Sec. V.3.5.1.

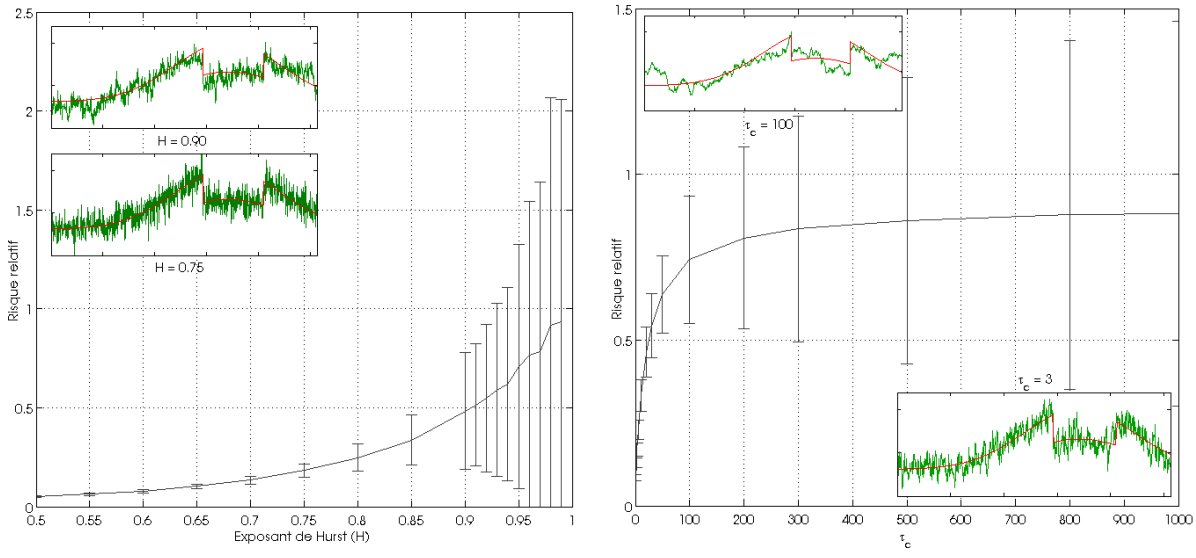


Figure II.3.8: Influence of noise correlation on denoising performance, for the academic signal in Fig. II.3.7. Left: optimal risk for fractional Gaussian noise as a function of the Hurst exponent. Right: optimal risk for exponentially correlated noise as a function of the correlation time.

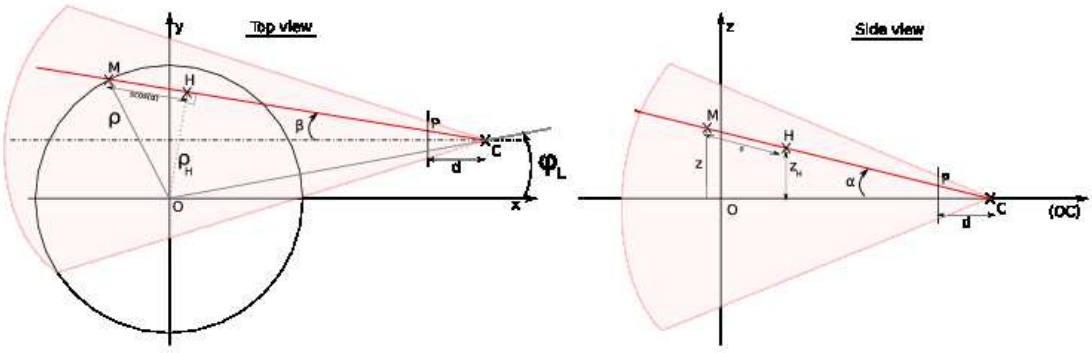


Figure II.4.1: Schematic representation of the camera setting with appropriate notations.

II.4 Application: edge plasma tomography

This section stems from a collaborative effort, with the participation of Nicolas Fedorczak and Pascale Monier-Garbet from CEA-IRFM, Frédéric Brochard and Gérard Bonhomme from LPMI-CNRS-IJL (Nancy Université), Kai Schneider and Marie Farge. It will form the basis of a future publication after some remaining questions are addressed.

II.4.1 Introduction

Turbulent fluctuations of plasma density in the edge of tokamaks are important for cross-field particle and heat transport, and for plasma-wall interaction. The plasma is at a relatively low temperature in these regions, allowing it under certain conditions to emit visible light, which can be observed by a fast camera.

Since the photons received by each pixel on the camera screen have been emitted all along the corresponding line of sight, the received intensity cannot be directly related to the volumic emissivity of the plasma. Nonetheless, as we will show, it is theoretically possible to invert the transformation if one assumes that the light emission is constant along field lines. This is a reasonable hypothesis if the toroidal angle does not vary too much within the field of view covered by the camera.

A similar problem has been encountered in space imaging, or in X-ray sensing of mechanical devices, but in these cases the simplifying assumption was made, that the object was lying at infinity. In that case the transformation is called the Abel transform. We first introduce a generalized definition of the Abel transform, and propose a way to solve the associated inverse problem by wavelet-vaguelette decomposition. We then show a few academic examples of inversion, validate the procedure by applying it to independently generated images, and finally, apply it to Tore Supra movies.

II.4.2 Reconstruction method

II.4.2.1 Notations

We start with magnetic coordinates (Ψ, θ, φ), where Ψ is a flux coordinate, φ is the toroidal angle, and θ is an appropriate poloidal coordinate such that the magnetic field lines have the equation:

$$\begin{cases} \Psi = \psi_0 \\ \varphi - q(\Psi)\theta = \text{constant} \end{cases} \tag{II.4.1}$$

where q is the safety factor. We assume that (Ψ, θ, φ) can be related to cylindrical coordinates (ρ, z, φ), where ρ is the distance to the axis of symmetry, and z is the vertical coordinate with respect to the horizontal midplane. In general, such a relation can be obtained using a magnetic reconstruction code, but for the purpose of this study, we limit ourselves to magnetic surfaces having a circular cross section, and we allow for a Shafranov shift. Taking for ψ the radius r of the flux surface, we obtain the analytic expressions:

$$\rho = R\left(1 + \gamma\left(1 - \frac{r^2}{a^2}\right)\right) + r\cos\left(\theta + \frac{\varphi}{q(r)}\right) \tag{II.4.2}$$

$$z = r\sin(\theta) \tag{II.4.3}$$

where γ characterizes the amplitude of the Shafranov shift. The axially symmetric case corresponds to $q = +\infty$.

The camera is modeled by a vanishingly small diaphragm located in cylindrical coordinates at $C(\rho_C, z_C, \varphi_C)$. The vertical and horizontal tilts of the screen are denoted respectively by α_C and β_C, and the distance between the screen and the diaphragm is d. The rays going through C can be parametrized by two angles : α is the angle with the plane $\varphi = 0$, and β is the angle with the plane $z = 0$. We call H the point on a ray which has the smallest ρ, and ρ_H is then the distance of closest approach of the ray to the z-axis. See Fig. II.4.1 for a schematic view of the setup.

Elementary geometry allows us to express α and β as functions of ρ_H and z_H as follows:

$$\sin(\alpha - \varphi_C) = \frac{\rho_H}{\rho_C} \tag{II.4.4}$$

$$\tan(\beta) = \frac{z_H - z_C}{\sqrt{\rho_C^2 - \rho_H^2}} \tag{II.4.5}$$

and, denoting by s the arc length along a ray, with $s = 0$ at H, we obtain a parametric representation of the ray in cylindrical coordinates:

$$\rho(s) = \sqrt{\rho_H^2 + s^2\cos^2(\beta)} \tag{II.4.6}$$

$$z(s) = z_H + s\sin(\beta) \tag{II.4.7}$$

$$\rho(s)\cos(\varphi(s) - \varphi_C) = \rho_C - s\cos(\alpha - \varphi_C) \tag{II.4.8}$$

II.4.2.2 Helical Abel transform

Let $S_0(\psi, \theta, \varphi)$ be the emissivity of the plasma at a point M defined by its field line coordinates (ψ, θ, φ). S_0 is expressed in $W.m^{-3}.Sr^{-1}$. We assume that the emission spectrum is

independent of position, so that the spectral response of the camera can be eliminated by a proper calibration. Moreover, we consider a transparent plasma, so that the flux density received by the camera screen around a point $P = (x, y)$ is given, up to a dimensional constant depending on d and the opening of the diaphragm, by the integral of the volume emissivity along the unique ray going through P and C:

$$I_0(x, y) = \int_{s_C}^{\infty} S_0(\psi, \theta, \varphi) \mathrm{d}s \tag{II.4.9}$$

where

$$x = d \tan(\alpha - \alpha_C) \tag{II.4.10}$$
$$y = d \tan(\beta - \beta_C) \tag{II.4.11}$$

and

$$s_C = -\frac{\sqrt{\rho_C^2 - \rho_H^2}}{\cos \alpha} \tag{II.4.12}$$

We denote by K the operator such that $I_0 = KS_0$. We are interested in the action of K on special kinds of emissivity fields, namely those that vary slowly along magnetic field lines as defined by Eq. II.4.1. Due to the fact that the safety factor q can be irrational, it would not be realistic to assume that S_0 is constant along field lines, since by continuity that would, in most cases, mean that S_0 is constant on magnetic surfaces. In the following, we shall assume that S_0 is constant on any connected portion of a field line visible in the camera field. This requirement will not exactly be met by real data, as we shall see, and this is one of the reasons why we need a very robust inversion algorithm. Now the restriction of K to these helically symmetric fields is a generalization of the classical Abel transform. To compute I_0, (II.4.2-II.4.3) need to be inverted and (II.4.9) can then be used. In our case the inversion can be done analytically, since r is defined as a root of a quadratic polynomial.

To discretize the operator K, we restrict (r, θ) to a rectangular domain $[r_{\min}, r_{\max}] \times [\theta_{\min}, \theta_{\max}]$ and use a regular Cartesian grid (r_i, θ_j) with $i = 1, \ldots, N_r$ and $j = 1, \ldots, N_\theta$. We assume that S vanishes for $r > r_{\max}$ and for $r < r_{\min}$. In the θ direction, we assume extend S periodically with the period $\theta_{\max} - \theta_{\min}$. Similarly, (x, y) is restricted to $[x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}]$ and discretized using by a grid (x_i, y_j) with $i = 1, \ldots, N_x$ and $j = 1, \ldots, N_y$. The integral in (II.4.9) is approximated by the method of rectangles using 1024 points between the bounds $s_{\min}$ and $s_{\max}$ defined by:

$$\cos(\alpha)^2 s_{\min}^2 = \rho_C^2 - \rho_H^2 \tag{II.4.13}$$
$$\cos(\alpha)^2 s_{\max}^2 = (R + a)^2 - \rho_H^2 \tag{II.4.14}$$
$$\tag{II.4.15}$$

The discretized operator K, for which we keep the same notation for simplicity, thus transforms matrices defined on the grid (r_i, θ_j) into matrices defined on the grid (x_i, y_j).

II.4.2.3 Wavelet-vaguelette decomposition

In practice, the measured I is corrupted by noise:

$$I = I_0 + W = KS_0 + W,$$

and even if K is invertible, a regularization method must be used to obtain a reasonable estimate for S_0. We propose to use the wavelet-vaguelette decomposition (Donoho, 1992) (WVD) method, which is now briefly recalled. Consider an orthogonal wavelet basis $(\psi_\lambda)_{\lambda\in\Lambda}$ on the domain $[r_{\min}, r_{\max}] \times [\theta_{\min}, \theta_{\max}]$ with periodic boundary conditions. The elements of the index set Λ are of the form $(j, \mathbf{i})$, where j indicates the scale of the wavelet, and $\mathbf{i}$ its position. The corresponding vaguelette families $(u_\lambda)_{\lambda\in\Lambda}$ and $(v_\lambda)_{\lambda\in\Lambda}$ are then defined by:

$$K\psi_\lambda = \kappa_\lambda v_\lambda \quad \text{(II.4.16)}$$

$$K^* u_\lambda = \kappa_\lambda \psi_\lambda, \quad \text{(II.4.17)}$$

where K^* is the adjoint of K, and the constants κ_λ are chosen in order to impose $\|u_\lambda\| = 1$ for all λ. Therewith v_λ is defined in the image plane (x, y), while u_λ is defined in the (r, θ) plane. From their definitions (II.4.16-II.4.17) it follows that the families (u_λ) and (v_μ) satisfy the biorthogonality relations:

$$\langle u_\lambda \mid v_\mu \rangle = \delta_{\lambda,\mu} \quad \text{(II.4.18)}$$

where $\delta_{\lambda,\mu}$ is the Kronecker symbol.

The WVD-reconstructed density S_R is then defined by:

$$S = \overline{S}_0 + \sum_{\lambda\in\Lambda} \mathbb{1}_{[\Theta,+\infty[}\left(\langle I \mid u_\lambda\rangle\, \kappa_\lambda^{-1}\right) \psi_\lambda \quad \text{(II.4.19)}$$

where $\overline{S}_0$ is the space average of S_0 and Θ_λ is a global threshold to be specified later. In practice, S_0 is unknown, so that our algorithm will actually only allow us to reconstruct $S - S_0$. This is however sufficient for the applications we are interested in.

The optimal threshold to use depends on the level of noise. To estimate it directly from the observed camera image, we propose to use an iterative algorithm(Azzalini et al., 2004).

To implement formula (II.4.19) in practice, the vaguelettes v_λ are first computed by applying K to the wavelets, and the u_λ are then obtained using (II.4.18). To invert the linear system (II.4.18), we use sparse Cholesky decomposition. Overall, the algorithm thus consists in the following steps:

(i) construct a sparse matrix representation of K,

(ii) compute the v_λ from (II.4.16),

(iii) compute the u_λ from (II.4.18) using sparse Cholesky decomposition,

(iv) compute the vaguelette coefficients $\langle I \mid u_\lambda\rangle$ using sparse matrix multiplication,

(v) reconstruct S from (II.4.19).

Note that the only expensive step in this algorithm is (iii), which needs to be done only once for a given geometric configuration of the camera. In the following we take as wavelet family the most regular Coiflets (Daubechies, 1993) with filters of length 12. Moreover, to be able to use the was wavelet transform algorithm, we impose that $N_r = N_\theta = 2^J$ for some integer J, but there is no restriction on N_x and N_y. Note that the restriction on N_r and N_θ could be diminished by using scaling function coefficients.

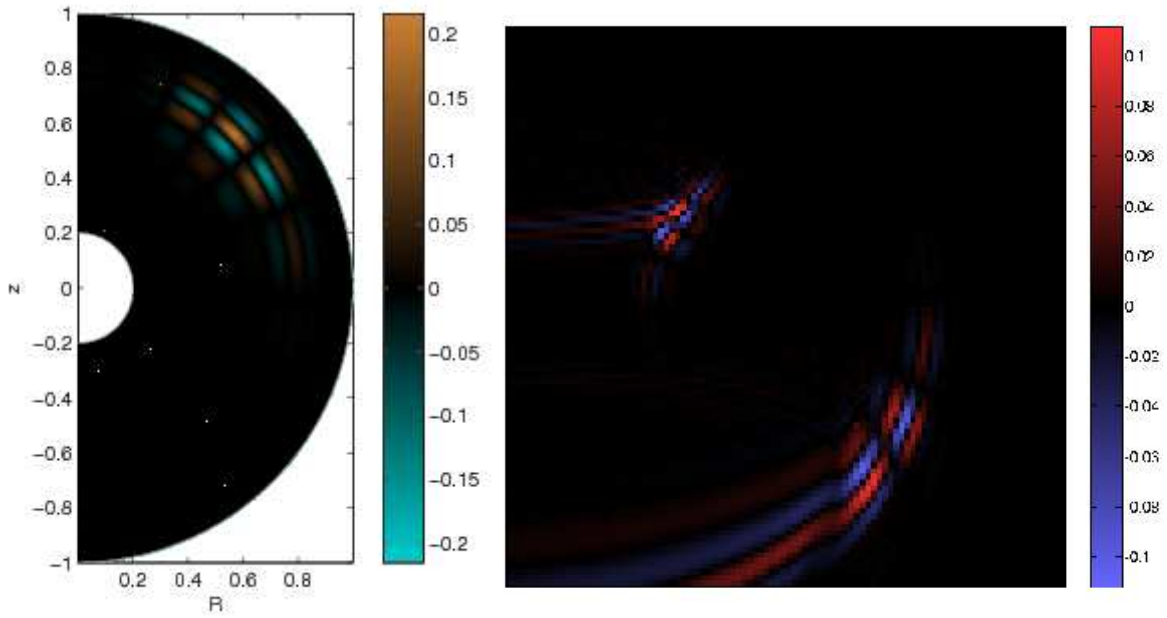


Figure II.4.2: Representation of a wavelet in the (r, θ) plane (left) and of the corresponding vaguelette v_λ in the image plane (right). The geometric parameters are indicated in Table. II.4.1.

II.4.3 Validation

II.4.3.1 First examples

In this section we report the results of some simple test cases, using the geometric parameters in the first column of Table II.4.1. For simplicity there is no Shafranov shift and no magnetic shear.

To give an idea of the type of functions we are working with, we first provide a representation of a wavelet and the corresponding vaguelette v_λ (Fig .II.4.2). As expressed by (II.4.16), the vaguelette v_λ is simply the image of the wavelet ψ_λ seen by the camera. The conjugated vaguelettes u_λ are not shown here because they do not benefit from such a simple interpretation. Moreover, since they are obtained from the v_λ by solving the inverse problem (II.4.18), they tend to be individually quite noisy, and only the whole family (u_λ) is meaningful.

To test WVD-reconstruction, we now apply it to a simple academic test case. We start with a homogeneous emissivity map (Fig. II.4.3, a) in a toric shell extending from $r = 0.47$ to $r = 0.73$, and we first apply the operator K to obtain a synthetic camera image (Fig. II.4.3, b). It is interesting to remark the critical curves that appear in the image plane due to the integration along lines of sight intersecting the toric shell. Then, we perturb the image with a Gaussian white noise having a standard deviation $\sigma = 0.125$ (Fig. II.4.3, c), and apply WVD-reconstruction to reconstruct the emissivity field from the noisy image. The result (Fig. II.4.3, d) preserves the main features of the input (Fig. II.4.3, a), although the consequences of the degradation, like spurious peaks and oscillations, are visible.

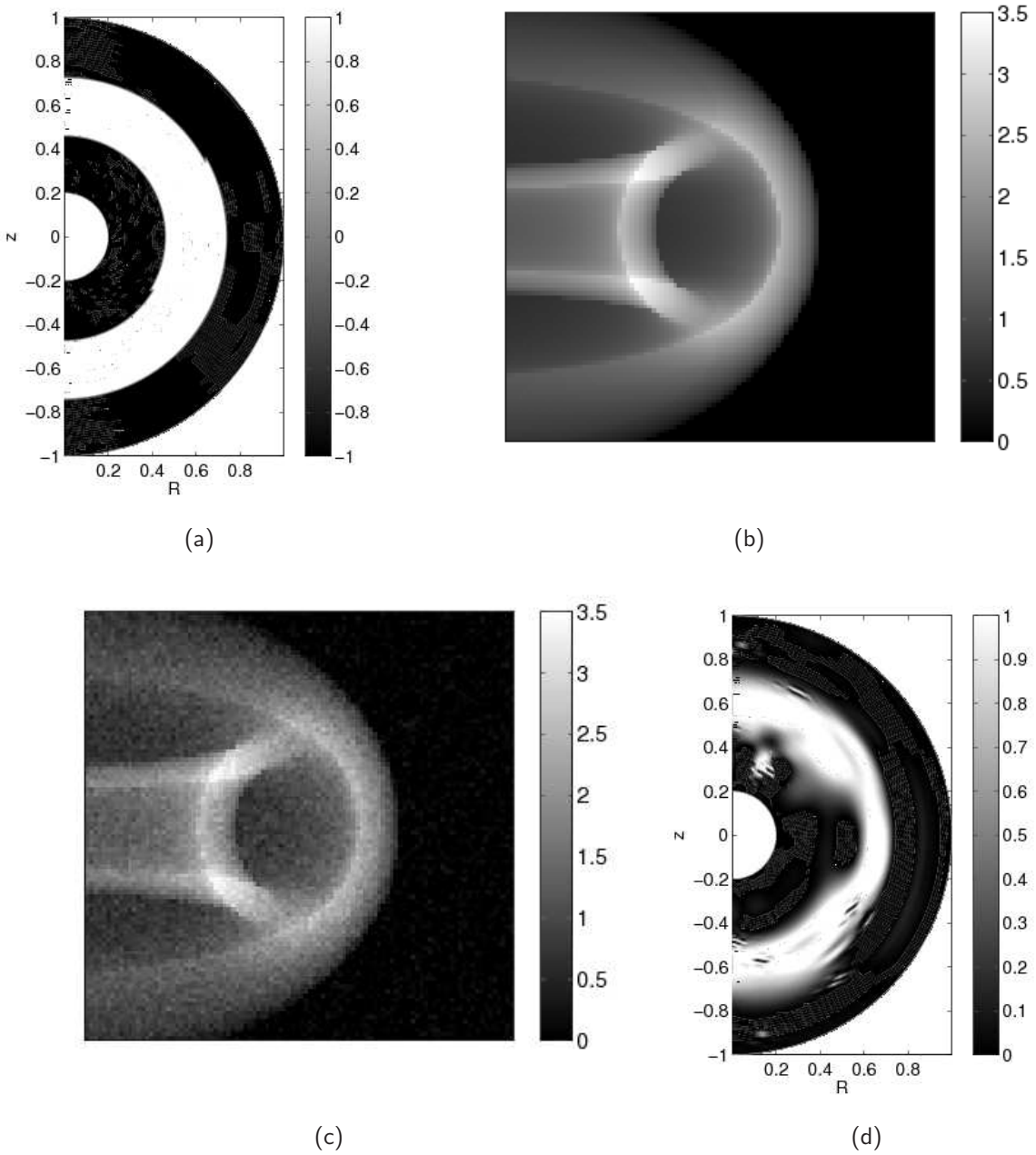


Figure II.4.3: Denoising test with uniform radiating shell from $r = 0.47$ to $r = 0.73$. The geometric parameters are indicated in Table. II.4.1. (a) Source emission intensity S_0. (b) Corresponding noise-less image I_0. (c) Noisy image I, with $\sigma = 0.125$. (d) WVD-inverted emission intensity S.

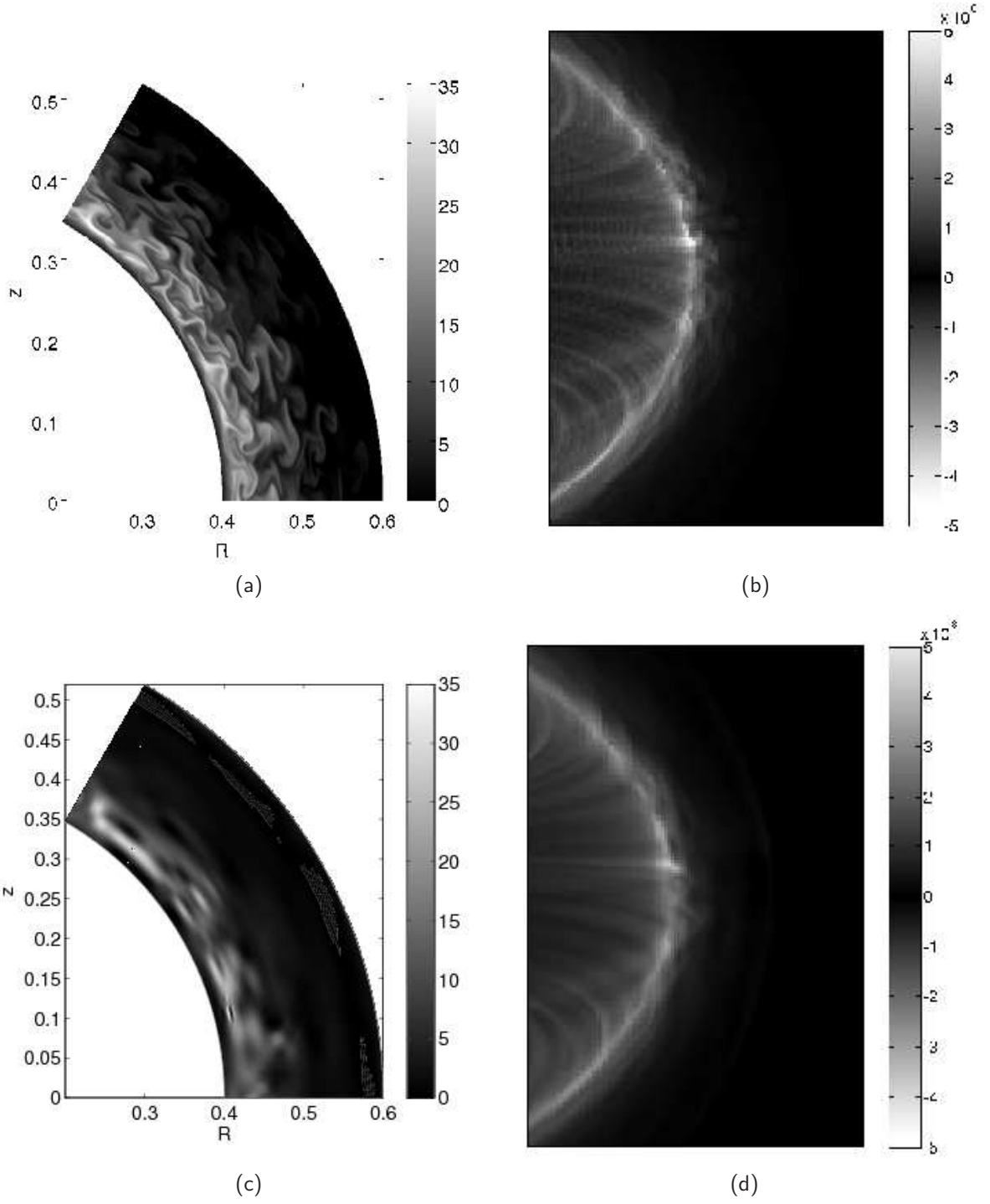
	Examples	TOKAM	TS 42967
R	2.30	2.30	2.22
γ	0	0	0.0650
q	∞	-3.00	-3.25
ρ_C (m)	4.00	3.20	3.53
z_C (m)	0	0	0.190
α_C	0	0	0
φ_L(rad)	-0.846	-0.846	-0.727
$r_{\min}$ (m)	0.200	0.400	0.28
$r_{\max}$ (m)	1.00	0.600	0.47
$\theta_{\min}$(rad)	$-\pi/2$	0	0
$\theta_{\max}$(rad)	$\pi/2$	$\pi/3$	$\pi/2$
$x_{\min}$ (d)	-0.750	0	0.111
$x_{\max}$ (d)	0.300	0.404	0.382
$y_{\min}$ (d)	-0.404	-0.404	-0.431
$y_{\max}$ (d)	0.404	0.404	0.266
N_x	100	101	110
N_y	100	150	283

Table II.4.1: Parameters of reported numerical experiments.

II.4.3.2 Test with Tokam data

To validate the method further, we apply it to camera images that are generated artificially but using a different method than the one presented above. In order to come closer to the experimental situations we are going to face, we take as input emissivity field the fourth power of the ion density in a computation of edge plasma fluctuations obtained with the Tokam 2D code (Ghendrih et al., 2005), shown in Fig. II.4.4 (a). The artificial camera image (Fig. II.4.4,b) was obtained by accumulating projections of successive poloidal cross-sections through S using Matlab, i.e. the operator K is discretized by splitting the integral (II.4.9) into a sum over φ instead of a sum over s as we normally do. This can be seen as a primitive way of evaluating the impact of all the small effects that we have not taken into account in our definition of K. The geometric parameters for this test case are shown in the second column of Table II.4.1.

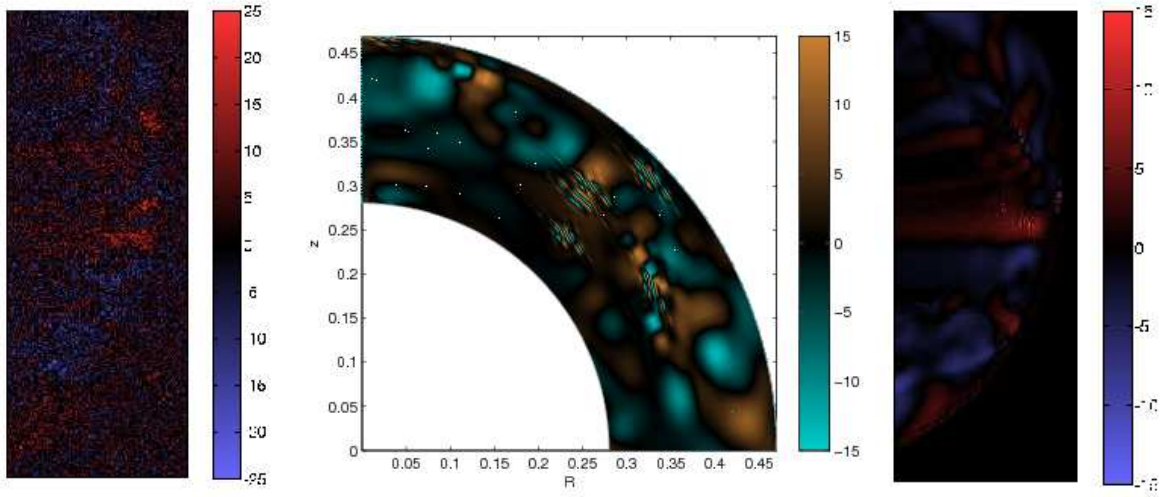
The emissivity map reconstructed by WVD is shown in Fig. II.4.4 (c). It looks as a smooth approximation of the input map. In particular the radial position of maximum average emissivity is well captured, as well as some of the main blobs, like the one in the top left corner. By applying the operator K again to this inverted emissivity map, we obtain an image (Fig. II.4.4, d) that is visually very similar to the one we start from (Fig. II.4.4, b), which supports the fact that not much information has been lost in the reconstruction process.



(a) (b)

(c) (d)

Figure II.4.4: Inversion test with artificial image generated from the TOKAM code. (a) Emissivity in the (r, θ) plane obtained from TOKAM run. (b) Artificial image obtained by stacking method. (c) WVD-inverted emissivity I. (d) Regenerated image KI using our method.



(a) (b) (c)

Figure II.4.5: Inversion test with experimental movie from Tore Supra shot 42967. (a) Movie frame. (b) WVD-reconstructed emissivity map S in (r, θ) plane. (c) Artificial movie frame regenerated by applying K to S.

II.4.4 Application to Tore Supra movies

In this section we present preliminary results obtained by applying the WVD-reconstruction method to an experimental movie frame acquired during the Tore Supra shot 42967. The plasma is in a detached phase, and therefore is strongly radiative in the neighborhood of a magnetic surface located at $r \simeq 0.4$. The camera is oriented so that its line of sight is approximately tangent to this magnetic surface. The geometric parameters are provided in the last column of Table. II.4.1. Note that these parameters were estimated from the movie itself using a key-point detection method based on some visible features of the vessel.

One frame (Fig. II.4.5, a) was picked out randomly within the acquired movie to test WVD-inversion. Note that the time average of the whole movie has been subtracted from the frame, so that only the fluctuations are visible. The inverted emissivity map is shown in Fig. II.4.5 (b). The strong radiative activity going on around $r = 0.4$ is well detected by the algorithm. Except some very intense and localized artifacts, the field is quite smooth. The movie reveals that the structures are preserved in time and propagate counter-clockwise. In Fig. II.4.5 (c), we show the artificial movie frame obtained by applying K to the inverted emissivity map. The main features that were visible by eye in the original movie frame (Fig. II.4.5, a) are strongly enhanced in the artificial one, while the noise has been reduced to a very low level.

II.4.5 Conclusion

We have proposed a new method for reconstructing the volumic light emissivity map of a tokamak plasma using a single camera. Our method relies on the hypothesis that the emissivity varies sufficiently slowly along magnetic field lines. We have demonstrated its feasibility using simple academic test cases, and validated its robustness by applying it to

independently generated artificial movies based on numerical computations of plasma edge fluctuations by the Tokam 2D code. Finally, we have presented a preliminary result showing that the method can be applied to an actual experimental movie. The comparison with other diagnostics, like Doppler reflectometry, is ongoing and we hope that it will validate the method.

The technical tool underlying our approach is the wavelet-vaguelette decomposition (Donoho, 1992) (WVD), which is an efficient way of solving ill-posed inverse problems in the presence of noise. Thanks to the localization of the wavelets, features such as blobs and fronts are preserved in the denoised emissivity map. We have seen that some artifacts persist in the denoised output, but there are good hopes that the method could be improved in the future, for example by choosing the threshold in a more refined way.

Part III

Particle-in-Wavelets approach for the Vlasov equation

> "Ah, this is marvelous!" said Lord Wen-hui. "Imagine skill reaching such heights!"
> Cook Ting laid down his knife and replied:
> "What I care about is the Way, which goes beyond skill. When I first began cutting up oxen, all I could see was the ox itself. After three years I no longer saw the whole ox. And now – now I go at it by spirit and don't look with my eyes. Perception and understanding have come to a stop and spirit moves where it wants. I go along with the natural makeup, strike in the big hollows, guide the knife through the big openings, and follow things as they are. So I never touch the smallest ligament or tendon, much less a main joint."
>
> ZHUANGZI, chapter 2

Summary

III.1 Wavelet-based density estimation

The contents of this section are reproduced from a published paper, see (Nguyen van yen et al., 2010b).

III.1.1 Introduction

Particle-based numerical methods are routinely used in plasma physics calculations (Birdsall and Langdon, 1985; Hockney and Eastwood, 1988). In many cases these methods are more efficient and simpler to implement than the corresponding Eulerian methods. However, particle methods face the well known statistical sampling limitation of attempting to simulate a physical system containing N particles using $N_p \ll N$ computational particles. Particle methods do not seek to reproduce the exact individual behavior of the particles, but rather to approximate statistical macroscopic quantities like density, current, and temperature. These quantities are determined from the particle distribution function. Therefore, a problem of relevance for the success of particle-based simulations is the reconstruction of the particle distribution function from discrete particle data.

The difference between the distribution function reconstructed from a simulation using N_p particles and the exact distribution function gives rise to a discretization error generically known as "particle noise" due to its random-like character. Understanding and reducing this error is a complex problem of importance in the validation and verification of particle codes, see for example Refs. (Nevins et al., 2005; Krommes, 2007; McMillan et al., 2008) and references therein for a discussion in the context of gyrokinetic calculations. One obvious way to reduce particle noise is by increasing the number of computational particles. However, the unfavorable scaling of the error with the number of particles, $\sim 1/\sqrt{N_p}$ (Krommes, 1993; Aydemir, 1994), puts a severe limitation on this straightforward approach. This has motivated the development of various noise reduction techniques including finite size particles (FSP) (Hockney, 1966; Langdon and Birdsall, 1970), Monte-Carlo methods (Aydemir, 1994), weight spreading (Brunner et al., 1999), Fourier-filtering (Jolliet et al., 2007), coarse-graining (Chen and Parker, 2007), Krook operators (McMillan et al., 2008), smooth interpolation (Cormier-Michel et al., 2008), low noise collision operators (Lewandowski, 2005), and Proper Orthogonal Decomposition (POD) methods (del Castillo-Negrete et al., 2008) among others.

In the present paper we propose a wavelet-based method for noise reduction in the reconstruction of particle distribution functions from particle simulation data. The method, known as Wavelet Based Density Estimation (WBDE), was originally introduced in Ref. (Donoho et al., 1996) in the context of statistics to estimate probability densities given a finite number of independent measurements. However, to our knowledge, this method has not been applied before to particle-based computations. WBDE, as used here, is based on a truncation of the wavelet representation of the Dirac delta function associated with each particle. The method yields almost optimal results for functions with unknown local smoothness without compromising computational efficiency, assuming that the particles' coordinates are statistically independent. As a first step in the application of the WBDE method to plasma particle simulations, we limit our attention to "passive denoising". That is the WBDE method is treated as a post-processing technique applied to independently generated particle data. The problem of "active denoising", e.g. the application of WBDE methods in the evaluation of self-consistent fields in particle in cell simulations, will not be addressed. This simplification

will allow us to assess the efficiency of the proposed noise reduction method in a simple setting. Another simplification pertains the dimensionality. Here, for the sake of simplicity, we limit attention to the reconstruction and denoising problem in two dimensions. However, the extension of the WBDE method to higher dimensions is in principle straightforward.

Collisions, or the absence of them, play an important role in plasma transport problems. Particle methods handle the collisional and non-collisional parts of the dynamics differently. Fokker-Planck-type collision operators are typically introduced in particle methods using Langevin-type stochastic differential equations. On the other hand, the non-collisional part of the dynamics is described using deterministic ordinary differential equations. Collisional dominated problems tend to wash out fine scale structures whereas collisionless problems typically develop fine scale filamentary structures in phase space. Therefore, it is important to test how the efficiency of denoising depends on the level of collisionality. Here we test the WBDE method in strongly collisional, weakly collisional and collisionless regimes. For the strongly collisional regime we consider particle data of force-free collisional relaxation involving energy and pinch-angle scattering. The weakly collisional regime is illustrated using guiding-center particle data of a magnetically confined plasma in toroidal geometry. The collisionless regime is studied using particle in cell (PIC) data corresponding to bump-on-tail and two streams instabilities in the Vlasov-Poisson system.

Beyond the role of collisions, the data sets that we are considering open the possibility of exploring the role of external and self-consistent fields in the reconstruction of the particle density. In the collisional relaxation problem no forces act on the particles, in the guiding-center problem particles interact with an external magnetic field, and in the Vlasov-Poisson problem particle interactions are incorporated through a self-consistent electrostatic mean field. One of the goals of this paper is to compare the WBDE method with the Proper Orthogonal Decomposition (POD) density reconstruction method proposed in Ref. (del Castillo-Negrete et al., 2008).

The rest of the paper is organized as follows. In Sect. II we review the main properties of kernel density estimation (KDE) and show its relationship with finite size particles (FSP). We then review basic notions on orthogonal wavelet and multiresolution analysis and outline a step by step algorithm for WBDE. For completeness, we also include in this section a brief description of the POD reconstruction method proposed in Ref. (del Castillo-Negrete et al., 2008). Section III discusses applications of the WBDE method and the comparison with the POD method. We start by post-processing a simulation of plasma relaxation by random collisions against a background thermostat. We then turn to a δf Monte-Carlo simulation in toroidal geometry, whose phase space has been reduced to two dimensions. Finally, we analyze the results of particle-in-cell (PIC) simulations of a 1D Vlasov-Poisson plasma. The conclusions are presented in Sec. IV.

III.1.2 Methods

This section presents the wavelet-based density estimation (WBDE) algorithm. We start by reviewing basic ideas on kernel density estimation (KDE) which is closely related to the use of finite size particles (FSP) in PIC simulations. Following this, we give a brief introduction to wavelet analysis and discuss the WBDE algorithm. For completeness, we also include a brief summary of the POD approach.

III.1.2.1 Kernel density estimation

Given a sequence of independent and identically distributed measurements, the nonparametric density estimation problem consists in finding the underlying probability density function (PDF), with no a priori assumptions on its functional form. Here we discuss general ideas on this difficult problem for which a variety of statistical methods have been developed. Further details can be found in the statistics literature, e.g. Ref. (Silverman, 1986).

Consider a number N_p of statistically independent particles with phase space coordinates $(\mathbf{X}_n)_{1\leq n\leq N_p}$ distributed in $\mathbb{R}^d$ according to a PDF f. This data can come from a PIC or a Monte-Carlo, full f or δf simulation. Formally, the sample PDF can be written as

$$f^\delta(\mathbf{x}) = \frac{1}{N_p}\sum_{n=1}^{N_p}\delta(\mathbf{x}-\mathbf{X}_n) \tag{III.1.1}$$

where δ is the Dirac distribution. Because of its lack of smoothness, Eq. (III.1.1) is far from the actual distribution f according to most reasonable definitions of the error. The dependence of f^δ on the statistical fluctuations in $(\mathbf{X}_n)$ can lead to an artificial increase of the collisionality, which could be problematic in the modeling of near collisionless plasmas of interest to controlled fusion. Beyond introducing dissipation, noise can lead to other problems including self-heating and momentum spread which, for example, is known to be an issue in laser-plasma interaction computations. Also, computations involving derivatives of f, like for example quasilinear fluxes in wave-particle interaction calculations, can be seriously compromised by poor reconstruction techniques.

The simplest method to introduce some smoothness in f^δ is to use a histogram. Consider a tiling of the phase space by a Cartesian grid with N_g^d cells. Let $\{B_\lambda\}_{\lambda\in\Lambda}$ denote the set of all cells with characteristic function χ_λ defined as $\chi_\lambda = 1$ if $x\in B_\lambda$ and $\chi_\lambda = 0$ otherwise. Then the histogram corresponding to the tiling is

$$f^H(\mathbf{x}) = \sum_{\lambda\in\Lambda}\left(\frac{1}{N_p}\sum_{n=1}^{N_p}\chi_\lambda(\mathbf{X}_n)\right)\chi_\lambda(\mathbf{x}) \tag{III.1.2}$$

which can also be viewed as the orthogonal projection of f^δ on the space spanned by the χ_λ. The main difference between f^δ and f^H is that the latter cannot vary at scales finer than the grid scale which is of order N_g^{-1}. By choosing N_g small enough, it is therefore possible to reduce the variance of f^H to very low levels, but the estimate then becomes more and more biased towards a piecewise continuous function, which is not smooth enough to be the true density. Histograms correspond to the nearest grid point (NGP) charge assignment scheme used in the early days of plasma physics computations (Hockney, 1966).

One of the most popular methods to achieve higher level of smoothness is kernel density estimation (KDE) (Parzen, 1962). Given $(\mathbf{X}_n)_{1\leq n\leq N_p}$, the kernel estimate of f is defined as

$$f^K(\mathbf{x}) = \frac{1}{N_p}\sum_{n=1}^{N_p}K(\mathbf{x}-\mathbf{X}_n)\,, \tag{III.1.3}$$

where the smoothing kernel K is a positive function, normalized such that $\int K = 1$. Equation (III.1.3) corresponds to the convolution of K with the Dirac delta measure corresponding

to each particle. A typical example is the Gaussian kernel

$$K_h(\mathbf{x}) = \frac{1}{(\sqrt{2\pi}h)^d} e^{-\frac{\|\mathbf{x}\|^2}{2h^2}} \tag{III.1.4}$$

where the so-called "bandwidth", or smoothing scale, h, is a free parameter. The optimal smoothing scale depends on how the error is measured. For example, in the one dimensional case, to minimize the mean L^2-error between the estimate and the true density, the smoothing volume h^d should scale like ${N_p}^{-\frac{1}{5}}$, and the resulting error scales like $N_p^{-\frac{2}{5}}$ (Silverman, 1986). As in the case of histograms, the choice of h relies on a trade-off between variance and bias. In the context of plasma physics simulations the kernel K corresponds to the charge assignment function (Hockney and Eastwood, 1988).

A significant effort has been devoted in the choice of the function K since it has a strong impact on computational efficiency and on the conservation of global quantities. Concerning h, it has been shown that it should not be much larger than the Debye length λ_D of the plasma to obtain a realistic and stable simulation (Birdsall and Langdon, 1985). Given a certain amount of computational resources, the general tendency has thus been to reduce h as far as possible in order to fit more Debye lengths inside the simulation domain, which means that the effort has been concentrated on reducing the bias term in the error. Since the force fields depend on f through integral equations, like the Poisson equation, that tend to reduce the high wavenumber noise, we do not expect the disastrous scaling $h^d \propto {N_p}^{-\frac{1}{5}}$, which would mean $N_p \propto \lambda_D^{5d}$ in d dimensions, to hold. Nevertheless, the problem remains that if we want to preserve high resolution features of f or of the electromagnetic fields, we need to reduce h, and therefore greatly increase the number of particles to prevent the simulation from drowning into noise. Bandwidth selection has long been recognized as the central issue in kernel density estimation (Chiu, 1991). We are not aware of a theoretical or numerical prediction of the optimal value of h taking into account the noise term. To bypass this difficulty, it is possible to use new statistical methods which do not force us to choose a global smoothing parameter. Instead, they adapt locally to the behavior of the density f based on the available data. Wavelet based-density estimation, which we will introduce in the next two sections, is one of these methods.

III.1.2.2 Bases of orthogonal wavelets

Wavelets are a standard mathematical tool to analyze and compute non stationary signals. Here we recall basic concepts and definitions. Further details can be found in Ref. (Farge, 1992) and references therein. The construction takes place in the Hilbert space $L^2(\mathbb{R})$ of square integrable functions. An orthonormal family $(\psi_{j,i}(x))_{j\in\mathbb{N},i\in\mathbb{Z}}$ is called a wavelet family when its members are dilations and translations of a fixed function ψ called the mother wavelet:

$$\psi_{j,i}(x) = 2^{j/2}\psi(2^j x - i) \tag{III.1.5}$$

where j indexes the scale of the wavelets and i their positions, and ψ satisfies $\int \psi = 0$. In the following we shall always assume that ψ has compact support of length S. The coefficients $\langle f \mid \psi_{j,i}\rangle = \int f\psi_{j,i}$ of a function f for this family are denoted by $(\tilde{f}_{j,i})$. These coefficients describe the fluctuations of f at scale 2^{-j} around position $\frac{i}{2^j}$. Large values of j correspond to fine scales, and small values to coarse scales. Some members of the commonly used Daubechies 6 wavelet family are shown in the left panel of Fig. 1.

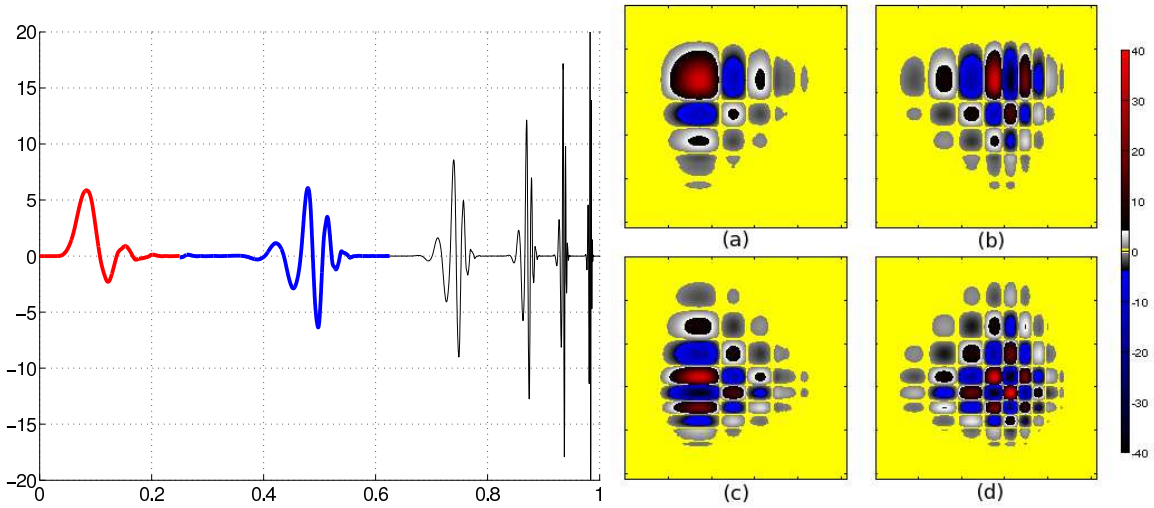


Figure III.1.1: Daubechies 6 wavelet family. Left, bold red: scaling function φ at scale $j = 5$. Left, bold blue: wavelet ψ at scale $j = 5$. Left, thin black, from left to right: wavelets at scales 6, 7, 8 and 9. Right : (a) 2D scaling function $\varphi(x_1)\varphi(x_2)$. (b) first 2D wavelet $\psi(x_1)\varphi(x_2)$. (c) second 2D wavelet $\varphi(x_1)\psi(x_2)$. (d) third 2D wavelet $\psi(x_1)\psi(x_2)$.

It can be shown that the orthogonal complement in $L^2(\mathbb{R})$ of the linear space spanned by the wavelets is itself orthogonally spanned by the translates of a function φ, called the scaling function. Defining

$$\varphi_{L,i} = 2^{\frac{L}{2}}\varphi(2^L x - i) \tag{III.1.6}$$

and the scaling coefficients $\bar{f}_{L,i} = \langle f \mid \varphi_{L,i}\rangle$, one thus has the reconstruction formula:

$$f = \sum_{i=-\infty}^{\infty} \bar{f}_{L,i}\varphi_{L,i} + \sum_{j=L}^{\infty}\sum_{i=-\infty}^{\infty} \tilde{f}_{j,i}\psi_{j,i} \tag{III.1.7}$$

The first sum on the right hand side of Eq. (III.1.7) is a smooth approximation of f at the coarse scale, 2^{-L}, and the second sum corresponds to the addition of details at successively finer scales.

If the wavelet ψ has M vanishing moments:

$$\int x^m \psi(x)dx = 0 \tag{III.1.8}$$

for $0 \leq m < M$, and if f is locally m times continuously differentiable around some point x_0, then a key property of the wavelet expansion is that the coefficients located near x_0 decay like $2^{-j(m+\frac{1}{2})}$ when $j \to \infty$ (Jaffard, 1991). Hence, localized singularities or sharp features in f affect only a finite number of wavelet coefficients within each scale. Another important consequence of (III.1.8) of special relevance to particle methods is that, for $0 \leq m < M$, the moments $\int x^m f(x)dx$ of the particle distribution function depend only on its scaling coefficients, and not on its wavelet coefficients.

If the scaling coefficients $\overline{f}_{J,i}$ at a certain scale J are known, all the wavelet coefficients at coarser scales ($j \leq J$) can be computed using the fast wavelet transform (FWT) algorithm (Mallat, 1999). We shall address the issue of computing the scaling coefficients themselves in section III.1.2.4.

The generalization to d dimensions involves tensor products of wavelets and scaling functions at the same scale. For example, given a wavelet basis on $\mathbb{R}$, a wavelet basis on $\mathbb{R}^2$ can be constructed in the following way:

$$\psi^1_{j,i_1,i_2}(x_1,x_2) = 2^j\psi(2^jx_1-i_1)\varphi(2^jx_2-i_2) \quad \text{(III.1.9)}$$
$$\psi^2_{j,i_1,i_2}(x_1,x_2) = 2^j\varphi(2^jx_1-i_1)\psi(2^jx_2-i_2) \quad \text{(III.1.10)}$$
$$\psi^3_{j,i_1,i_2}(x_1,x_2) = 2^j\psi(2^jx_1-i_1)\psi(2^jx_2-i_2)\,, \quad \text{(III.1.11)}$$

where we refer to the exponent $\mu = 1,2,3$ as the direction of the wavelets. This name is easily understood by looking at different wavelets shown in Fig. III.1.1 (right). The corresponding scaling functions are simply given by $2^j\varphi(2^jx_1-i_1)\varphi(2^jx_2-i_2)$. Wavelets on $\mathbb{R}^d$ are constructed exactly in the same way, but this time using 2^d-1 directions. To lighten the notation we write the d-dimensional analog of Eq. (III.1.7) as

$$f = \sum_{\lambda\in\Lambda_{\phi,L}} \overline{f}_\lambda\phi_\lambda + \sum_{\lambda\in\Lambda_{\psi,L}} \tilde{f}_\lambda\psi_\lambda \quad \text{(III.1.12)}$$

where $\lambda = (j,\mathbf{i},\mu)$ is a multi-index, with the integer j denoting the scale and the integer vector $\mathbf{i} = (i_1,i_2,\ldots)$ denoting the position of the wavelet.

The wavelet multiresolution reconstruction formula in Eq. (III.1.7) involves an infinite sum over the position index i. One way of dealing with this sum is to determine a priori the non-zero coefficients in Eq. (III.1.7), and work only with these coefficients, but still retaining the full wavelet basis on $\mathbb{R}^d$ as presented above. An alternative, which we have chosen because it is easier to implement, is to periodize the wavelet transform on a bounded domain (Mallat, 1999). Assuming that the coordinates have been rescaled so that all the particles lie in $[0,1]^d$, we replace the wavelets and scaling functions by their periodized counterparts:

$$\psi_{j,i}(x) \to \sum_{l=-\infty}^{\infty} \psi_{j,i}(x+l) \quad \text{(III.1.13)}$$
$$\varphi_{j,i}(x) \to \sum_{l=-\infty}^{\infty} \varphi_{j,i}(x+l)\,. \quad \text{(III.1.14)}$$

Throughout this paper we will consider only periodic wavelets. For the sake of completeness we mention a third alternative which is technically more complicated. It consists in constructing a wavelet basis on a bounded interval (Cohen et al., 1993). The advantage of this approach is that it does not introduce artificially large wavelet coefficients at the boundaries for functions f that are not periodic.

III.1.2.3 Wavelet based density estimation

The multiscale nature of wavelets allows them to adapt locally to the smoothness of the analyzed function (Mallat, 1999). This fundamental property has triggered their use in a variety of problems. One of their most fruitful applications has been the denoising of intermittent signals (Donoho and Jonhstone, 1994). The practical success of wavelet thresholding to reduce noise relies on the fact that the expansion of signals in a wavelet basis is typically sparse. Sparsity means that the interesting features of the signal are well summarized by

a small fraction of large wavelet coefficients. On the contrary, the variance of the noise is spread over all the coefficients appearing in Eq. (III.1.12). Although the few large coefficients are of course also affected by noise, curing the noise in the small coefficients is already a very good improvement. The original setting of this technique, hereafter referred to as global wavelet shrinkage, requires the noise to be additive, stationary, Gaussian and white. It found a first application in plasma physics in Ref. (Farge et al., 2006), where coherent bursts were extracted out of plasma density signals. Since Ref. (Donoho and Jonhstone, 1994), wavelet denoising has been extended to a number of more general situations, like non-Gaussian or correlated additive noise, or to denoise the spectra of locally stationary time series (von Sachs and Schneider, 1996). In particular, the same ideas were developed in Ref. (Vannucci and Vidakovic, 1998; Donoho et al., 1996) to propose a wavelet-based density estimation (WBDE) method based on independent observations. At this point we would like to stress that WBDE assumes nothing about the Gaussianity of the noise, nor on its stationarity. In fact, under the independence hypothesis – which is admittedly quite strong – the statistical properties of the noise are entirely determined by standard probability theory. We refer to Ref. (Vidakovic, 1999) for a review on the applications of wavelets in statistics. In Ref. (Gassama et al., 2007), global wavelet shrinkage was applied directly to the charge density of a 2D PIC code, in a case were the statistical fluctuations were quasi Gaussian and stationary. In particular, an iterative algorithm (Azzalini et al., 2004), which crucially relies on the stationnarity hypothesis, was used to determine the level of fluctuations. However,in the next section we will show an example where the noise is clearly non-stationary, and this procedure fails.

Let us now describe the WBDE method as we have generalized it to several dimensions. The first step is to expand the sample particle distribution function, f^δ, in Eq. (III.1.1) in a wavelet basis according to Eq. (III.1.12) with the wavelet coefficients

$$\overline{f}^\delta_\lambda = \langle f^\delta \mid \varphi_\lambda \rangle = \frac{1}{N_p} \sum_{n=1}^{N_p} \varphi_\lambda(X_n) \qquad \text{(III.1.15)}$$

$$\tilde{f}^\delta_\lambda = \langle f^\delta \mid \psi_\lambda \rangle = \frac{1}{N_p} \sum_{n=1}^{N_p} \psi_\lambda(X_n)\,. \qquad \text{(III.1.16)}$$

Since this reconstruction is exact, keeping all the wavelet coefficients does not improve the smoothness of f^δ. The simple and yet efficient remedy consists in keeping only a subset of the wavelet coefficients in Eq. (III.1.12). A straightforward prescription would be to discard all the wavelet coefficients at scales finer than a cut-off scale L. This approach corresponds to a generalization of the histogram method in Eq. (III.1.2) with $N_g = 2^L$. Because the characteristic functions χ_λ of the cells in a dyadic grid are the scaling functions associated with the Haar wavelet family, Eqs. (III.1.12) and (III.1.2) are in fact equivalent for this wavelet family. Accordingly, like in the histogram case, we would have to choose L quite low to obtain a stable estimate, at the risk of losing some sharp features of f. Better results can be obtained by keeping some wavelet coefficients down to a much finer scale $J > L$. However, to prevent that statistical fluctuations contaminate the estimate, only those coefficients whose modulus are above a certain threshold should be kept. We are thus naturally led to a nonlinear thresholding procedure. In the one dimensional case, values of J, L, and of the threshold within each scale that yield theoretically optimal results have been given in Ref. (Donoho et al., 1996). This reference discusses the precise smoothness requirements on f, which can accommodate well localized singularities, like shocks and

filamentary structures known to arise in collisionless plasma simulations. There remains the question of how to compute the wavelet coefficients $\tilde{f}_{j,i}$ based on the positions of the particles. Although more accurate methods based on (III.1.15) may be developed in the future, our present approximation relies on the computation of a histogram, which creates errors of order N_g^{-1}. The complete procedure is described in the following **Wavelet-based density estimation** algorithm:

(i) construct a histogram f^H of the particle data with $N_g = 2^{J_g}$ cells in each direction,

(ii) approximate the scaling coefficients at the finest scale J_g by :

$$\overline{f}_{J_g,\mathbf{i}} \simeq 2^{-J_g/2} f^H(2^{-J_g}\mathbf{i}) \qquad \text{(III.1.17)}$$

(iii) compute all the needed wavelet coefficients using the FWT algorithm,

(iv) keep all the coefficients for scales coarser than L, defined by $2^{dL} \sim N_p^{\frac{1}{1+2r_0}}$ where r_0 is the order of regularity of the wavelet (1 in our case),

(v) discard all the coefficients for scales strictly finer than J defined by $2^{dJ} \sim \frac{N_p}{\log_2 N_p}$,

(vi) for scales j in between L and J, keep only the wavelet coefficients $\tilde{f}_\lambda$ such that $|\tilde{f}_\lambda| \geq T_j = C\sqrt{\frac{j}{N_p}}$ where C is a constant that must in principle depend on the smoothness of f and on the wavelet family (Donoho et al., 1996).

The choice of parameters for the algorithm was justified rigourously in (Donoho et al., 1996). The dependence of the threshold on N_p and j can be intuitively understood as follows. Since the particles are assumed to be statistically independent, the standard deviation of each wavelet coefficient is proportional to $\frac{1}{\sqrt{N_p}}$. Because of the L^2 normalization of the wavelets (see (III.1.5)), the standard deviation is to a good approximation scale-independent. On the other hand, the central limit theorem implies that the fluctuations are almost Gaussian provided the number of particles is large enough. Therefore, to filter the fluctuations, the threshold should be larger than the typical value taken by a Gaussian random variable with standard deviation $\sigma \sim \frac{1}{\sqrt{N_p}}$. A standard choice to ensure this level of denoising is to take the threshold proportional to $\sigma\sqrt{2\ln(M)}$, where M is the number of samples (Mallat, 1999). Since there are 2^{dj} wavelet coefficients at scale j, we get a threshold proportional to $\sigma\sqrt{2\ln 2^{dj}} = C\sqrt{1/N_p}\sqrt{j}$. In the following, except otherwise indicated, we will assume the proportionality constant to be, $C = \frac{1}{2}$. For the wavelet bases we used orthonormal Daubechies wavelets with 6 vanishing moments and thus support of size $S = 12$ (Daubechies, 1992). In our case, $r_0 = 1$, which means that the wavelets have a first derivative but no second derivative, and the size of the wavelets at scale L for $d = 1$ is roughly $N_p^{-\frac{1}{3}}$. Since $N_p \gg 1$, it follows from the definition at stage 5 of the algorithm that the size of the wavelets at scale J is orders of magnitude smaller than $N_p^{-\frac{1}{3}}$. Using the adaptive properties of wavelets, we are thus able to detect fine scale structures of f without compromising the stability of the estimate. Note that the error at stage (ii) could be reduced by using Coiflets (Daubechies, 1993) instead of Daubechies wavelets, but the gain would be negligible

compared to the error made at stage (i). We will denote the WBDE estimate of f as f^W. In the one-dimensional case,

$$f^W = \sum_{i=1}^{2^L} \overline{f}_{L,i}\varphi_{L,i} + \sum_{j=L}^{J}\sum_{i=1}^{2^j} \tilde{f}_{j,i}\rho_j(\tilde{f}_{j,i})\psi_{j,i} \qquad \text{(III.1.18)}$$

where ρ_j is the thresholding function as defined by stage (vi) of the algorithm : $\rho_j(y) = 0$ if $|y| \leq T_j$ and $\rho_j(y) = 1$ otherwise.

Finally, let us propose two methods for applying WBDE to postprocess δf simulations. Recall that the Lagrangian equations involved in the δf schemes are identical to their full f counterparts. The only difficulty introduced by the δf method lies in the evaluation of phase space integrals of the form $\delta I = \int A \cdot (f - f_0)$, where A is a function on phase space and f_0 is a known reference distribution function. In these integrals, $f - f_0$ should be replaced by δf, which is in turn written as a product wf, where w is a "weighting" function. Numerically, w is known via its values at particles positions, $w(X_n)$, and the usual expression for δI is thus $\delta I = \sum_{n=1}^{N_p} A(X_n)w(X_n)$. We cannot apply WBDE directly to δf, since this function is not a density function. An elegant approach would be to first apply WBDE to the unweighted distribution f^δ to determine the set of statistically significant wavelet coefficients, and to include the weights only in the final reconstruction (III.1.18) of f^W. A simpler approach, which we will illustrate in section III.1.3.2, consists in renormalizing δf, so that $\int |\delta f| = 1$, and treating it like a density.

III.1.2.4 Further issues related to practical implementation

In this section we discuss how the WBDE method handles two issues of direct relevance to plasma simulations: conservation of moments and computational efficiency. As mentioned before, due to the vanishing moments of the wavelets in Eq. (III.1.8), the moments up to order M of the particle distribution distribution are solely determined by its scaling function coefficients. As a consequence, we expect the thresholding procedure to conserve these moments, in the sense that

$$\mathcal{M}^W_{m,k} = \int x_k^m f^W(\mathbf{x})\mathrm{d}\mathbf{x} \simeq \int x_k^m f^\delta(\mathbf{x})\mathrm{d}\mathbf{x} = \mathcal{M}^\delta_{m,k} \qquad \text{(III.1.19)}$$

for $0 \leq m \leq M-1$ and for all $i \in \{1, \ldots, d\}$. This conservation holds up to round-off error if the wavelet coefficients can be computed exactly. Due to the type of wavelets that we have used, we were not able to achieve this in the results presented here. There remains a small error related to stages 1 and 2 of the algorithm, namely the construction of f^H and the approximation of the scaling function coefficients by Eq. (III.1.17). They are both of order N_g^{-1}. We will present numerical examples of the moments of f^W in the next section.

Conservation of moments is closely related to a peculiarity of the denoised distribution function resulting from the WBDE algorithm: it is not necessarily everywhere positive. Indeed, wavelets are oscillating functions by definition, and removing wavelet coefficients therefore cannot preserve positivity in general. Further studies are needed to assess if this creates numerical instabilities when f^W is used in the computation of self-consistent fields. The same issue was discussed in Ref. (Denavit, 1972) where a kernel with two vanishing moments was used to linearly smooth the distribution function. The fact that this kernel is

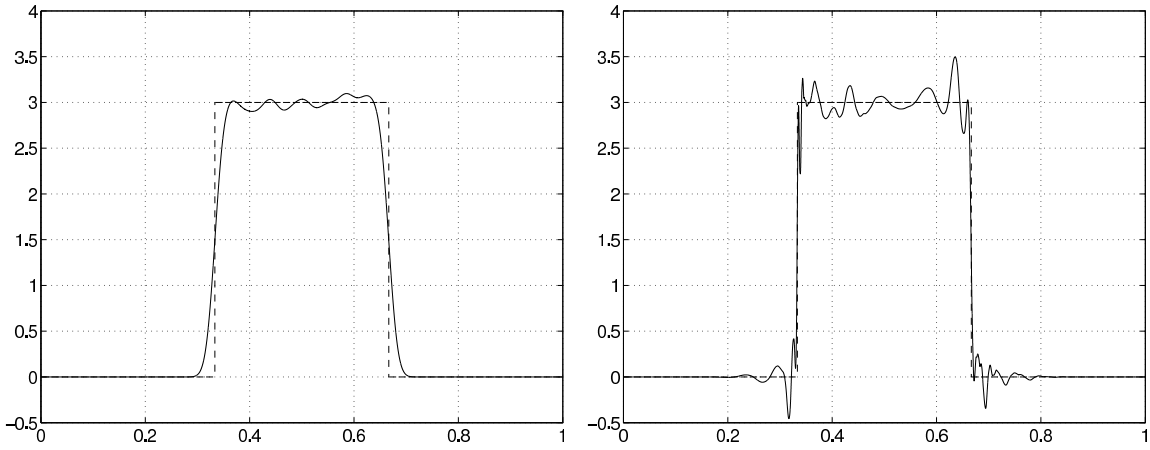


Figure III.1.2: Estimation of the density of a sample of size 2^{14} drawn uniformly in $[1/3, 2/3]$, using Gaussian kernels (left) or wavelets (right). The discontinuous analytical density is plotted with a dashed line in the two cases.

not everywhere positive was not considered harmful in this reference. We acknowledge that it may render the resampling of new particles from f^W, if it is needed in the future, more difficult. There are ways of forcing f^W to be positive, for example by applying the method to $\sqrt{f}$ and then taking the square of the resulting estimate, but this implies the loss of the moment conservation, and we have not pursued in this direction.

The number of multiplications required to perform forward and inverse wavelet transforms between scales J and L using the FWT algorithm in d dimensions is $4dS2^{d(J-L)}$, where S is the length of the wavelet filter (12 for the Daubechies filter that we are using). From the definitions of J and L it follows that $2^{d(J-L)} \sim N_p^{\frac{2}{3}} / \log_2 N_p$. Thus, for a simulation in d dimensions involving $N_p = 10^m$ particles, the number of operations scales as $\sim (4Sd \log_{10} 2/m) \times 10^{2m/3}$. For example, for a 2-dimensional simulation with 10^7 particles, $J = 10$ and $L = 4$. In this case, the computation of the wavelet coefficients require $\sim 1.5 \times 10^6$ multiplications which takes about 0.3 sec of CPU time on a desktop computer. For a more demanding 4-dimensional simulation with 10^8 particles, we find that $J = 6$ and $L = 3$, which increases the number of multiplications to 12.5×10^6. The cost of the binning stage is of order N_p, so that the total number of operations for computing f^W is proportional to N_p, but overall the number of multiplications per particle remains small. On the other hand, advancing the particles in time is likely to require several multiplications per particle. If one wishes to use a finer grid to ensure high accuracy conservation of moments, the storage requirement grows like N_g^d. Thanks to optimized in-place algorithms, the amount of additional memory needed during the computation does not exceed $3S$. Also, an important issue that needs to be kept in mind is that the FWT algorithm requires N_g to be an integer multiple of 2^{J-L}. For comparison purposes, let us recall that most algorithms to compute the POD in 2 dimensions have a complexity proportional to N_g^3, in addition to the order N_p cost of binning.

To conclude this subsection, Fig. III.1.2 presents an example of the reconstruction of a 1D discontinuous density that illustrates the difference between the KDE and WBDE methods. The probability density function is uniform on the interval $\left[\frac{1}{3}, \frac{2}{3}\right]$ and the estimates were computed on $[0, 1]$ to include the discontinuities. The sample size was 2^{14}, and the binning used $N_g = 2^{16}$ cells to compute the scaling function coefficients. For this 1D case the value

	$m=0$	$m=1$	$m=2$	$m=4$
f^K	$1.81 \cdot 10^{-5}$	$1.70 \cdot 10^{-5}$	$7.52 \cdot 10^{-4}$	$3.90 \cdot 10^{-3}$
f^W	$1.08 \cdot 10^{-11}$	$1.52 \cdot 10^{-5}$	$2.93 \cdot 10^{-5}$	$5.52 \cdot 10^{-5}$

Table III.1.1: Relative absolute difference between the moments of f^δ and those of f^K and f^W, for the distribution function corresponding to Fig. III.1.2.

$C = 2$ was used to determine the thresholds (step (vi) of the algorithm). The KDE estimate is computed using a Gaussian kernel with smoothing scale $h = 0.0138$ (Ihler, 2003). The relative mean squared errors associated with the KDE and WBDE estimates are respectively 19.6×10^{-3} and 6.97×10^{-3}. The error in the KDE estimate comes mostly from the smoothing of the discontinuities. The better performance of WBDE stems from the much sharper representation of these discontinuities. It is also observed that the WBDE estimate is not everywhere positive. The approximate conservation of moments is demonstrated on Table III.1.1. Note that the error on all these moments for f^W could be made arbitrary low by increasing N_g. The overshoots could also be mitigated by using nearly shift invariant wavelets (Kingsbury, 2001).

III.1.2.5 Proper Orthogonal Decomposition Method

For completeness, in this subsection we present a brief review of the POD density reconstruction method. For the sake of comparison with the WBDE method, we limit attention to the time independent case. Further details, including the reconstruction of time dependent densities using POD methods can be found in Ref. (del Castillo-Negrete et al., 2008).

The first step in the POD method is to construct the histogram f^H from the particle data. This density is represented by an $N_x \times N_y$ matrix $\hat{f}_{ij}$ containing the fraction of particles with coordinates (x, y) such that $X_i \leq x < X_{i+1}$ and $Y_i \leq y < Y_{i+1}$. In two dimensions, the POD method is based on the singular value decomposition of the histogram. According to the SVD theorem (Golub and van Loan, 1996), the matrix $\hat{f}$ can always be factorized as $\hat{f} = UWV^t$, where U and V are $N_x \times N_x$ and $N_y \times N_y$ orthogonal matrices, $UU^t = VV^t = I$, and W is a diagonal matrix, $W = \mathrm{diag}\,(w_1, w_2, \ldots w_N)$, such that $w_1 \geq w_2 \geq \ldots \geq w_N \geq 0$ with $N = \min(N_x, N_y)$.

In vector form, the decomposition can be expressed as

$$\hat{f}_{ij} = \sum_{k=1}^{N} w_k \, u_i^{(k)} \, v_j^{(k)} \,, \tag{III.1.20}$$

where the N_x-dimensional vectors, $u_i^{(k)}$, and the N_y-dimensional vectors, $v_j^{(k)}$, are the orthonormal POD modes and correspond to the columns of the matrices U and V respectively. Given the decomposition in Eq. (III.1.20), we define the rank-r approximation of $\hat{f}$ as

$$\hat{f}_{ij}^{(r)} = \sum_{k=1}^{r} w_k \, u_i^{(k)} \, v_j^{(k)} \,, \tag{III.1.21}$$

where $1 \leq r < N$, and define the corresponding rank-r reconstruction error as

$$e(r) = ||\hat{f} - \hat{f}^{(r)}||^2 = \sum_{i=r+1}^{N} w_i^2 \,, \tag{III.1.22}$$

where $||A|| = \sqrt{\sum_{ij} A_{ij}^2}$ is the Frobenius norm. Since $\hat{f}^{(r=N)} = \hat{f}$, we define $e(N) = 0$. The key property of the POD is that the approximation in Eq. (III.1.21) is optimal in the sense that

$$e(r) = \min \left\{ ||\hat{f} - g||^2 \ |\mathrm{rank}(g) = r \right\} \,. \tag{III.1.23}$$

That is, of all the possible rank-r Cartesian product approximations of $\hat{f}$, $\hat{f}^{(r)}$ is the closest to $\hat{f}$ in the Frobenius norm.

The SVD spectrum, $\{w_k\}$, of noise free coherent signals decays very rapidly after a few modes, but the spectrum of noise dominated signals is relatively flat and decays very slowly. When a coherent signal is contaminated with low level noise, the SVD spectrum exhibits an initial rapid decay followed by a weakly decaying spectrum known as the noisy plateau. In the POD method the denoised density is defined as the truncation $f^P = \hat{f}^{(r_c)}$, where r_c corresponds to the rank where the noisy plateau starts. In general it is difficult to provide a precise a priori estimate of r_c, and this is one of the potential limitations of the POD method. One possible quantitative criterion used in Ref. (del Castillo-Negrete et al., 2008) is to consider the relative decay of the spectrum, $\Delta(k) = (w_{k+1} - w_k)/(w_2 - w_1)$, for $k > 1$, and define r_c by the condition $\Delta(r_c) \leq \Delta_c$ where Δ_c is a predetermined threshold.

III.1.3 Applications

In this section, we apply the WBDE method to reconstruct and denoise the particle distribution function starting from discrete particle data. The data corresponds to three different groups of simulations: collisional thermalization with a background plasma, guiding center transport in toroidal geometry, and Vlasov-Poisson electrostatic instabilities. We will compare the WBDE and POD methods in all three cases. Note that the first two groups of simulations were already analyzed using POD methods in Ref. (del Castillo-Negrete et al., 2008). The third group, which is new here, allows the testing of the reconstruction algorithms in a collisionless system that incorporates the self-consistent evaluation of the forces acting on the particles, as opposed to the colliional, test particle problems analyzed before. When comparing the two methods, it is important to keep in mind that POD has one free parameter, namely the number r of singular vectors that are retained to reconstruct the denoised distribution function. In the cases studied here, we used a best guess for r based on the properties of the reconstruction. In Ref. (del Castillo-Negrete et al., 2008) the POD method was developed and applied to time independent and time dependent data sets. However, in the comparison with the WBDE method, we limit attention to 2-dimensional time independent data sets.

The accuracy of the reconstruction of the density at a fixed time t will be monitored using the mean square error

$$e = \sum_{i,j} |f^{est}(x_i, y_j; t) - f^{ref}(x_i, y_j; t)|^2 \,, \tag{III.1.24}$$

where (x_i, y_j) are the coordinates of the nodes of a prescribed $N_g \times N_g$ grid in phase space, and f^{est} denotes the estimated density computed from a sample with N_p particles. For the WBDE method $f^{est} = f^W$, and for the POD method $f^{est} = f^P$. In principle, the reference density, f^{ref}, in Eq. (III.1.24) should be the density function obtained from the exact solution of the corresponding continuum model, e.g. the Fokker-Planck or the Vlasov-Poisson system. However, when no explicit solution is available, we will set $f^{ref} = f^H$, where f^H is the histogram corresponding to a simulation with a maximum number of particles available, which in the cases reported here correspond to $N_p = 10^6$. We will also use the normalized error

$$e_0 = \frac{e}{\sum_{i,j} |f^{ref}(x_i, y_j; t)|^2} \,. \tag{III.1.25}$$

III.1.3.1 Collisional thermalization with a background plasma

This first example models the relaxation of a non equilibrium plasma by collisional damping and pitch angle scattering on a thermal background. The plasma is spatially homogeneous and is represented by an ensemble of N_p particles in a three-dimensional velocity space. Assuming a strong magnetic field, the dynamics can be reduced to two degrees of freedom: the magnitude of the particle velocity, v, and the particle pitch, $\lambda = \cos\theta$, where θ is the angle between the particle velocity and the magnetic field. In the continuum limit the particle distribution function is governed by the Fokker-Planck equation, which in the particle description corresponds to the stochastic differential equations

$$d\lambda = -\lambda \nu_D \, dt - \sqrt{\nu_D \, (1 - \lambda^2)} \, d\eta_\lambda \,, \tag{III.1.26}$$

$$dv = -\left[\alpha \, \nu_s \, v - \frac{1}{2v^2} \frac{d}{dv} \left(\nu_{||} v^4\right)\right] \, dt + \sqrt{v^2 \, \nu_{||}} \, d\eta_v \,, \tag{III.1.27}$$

describing the evolution of $v \in (0, \infty)$ and $\lambda \in [-1, 1]$ for each particle, where $d\eta_\lambda$ and $d\eta_v$ are independent Wiener stochastic processes and ν_D, ν_s and $\nu_{||}$ are functions of v. For further details on the model see Ref. (del Castillo-Negrete et al., 2008) and references therein.

We considered simulations with $N_p = 10^3$, 10^4, 10^5 and 10^6 particles. The initial conditions of the ensemble of particles were obtained by sampling a distribution of the form

$$f(v, \lambda, t = 0) = A v^2 \exp\left\{-\frac{1}{2}\left[\frac{(\lambda - \lambda_0)^2}{\sigma_\lambda^2} + \frac{(v - v_0)^2}{\sigma_v^2}\right]\right\} \,, \tag{III.1.28}$$

where a v^2 factor has been included in the definition of the initial condition so that the volume element is simply $\mathrm{d}v\mathrm{d}\mu$, A is a normalization constant, $\lambda_0 = 0.25$, $v_0 = 5$, $\sigma_\lambda = 0.25$ and $\sigma_v = 0.75$. This relatively simple problem is particularly well suited for the WBDE method because the simulated particles do not interact and therefore statistical correlations cannot build-up between them.

Before applying the WBDE method, we analyze the sparsity of the wavelet expansion of f^δ, and compare the number of modes kept and the reconstruction error for different thresholding rules. The plot in the upper left panel of Fig. III.1.3 shows the absolute values of the wavelet coefficients in decreasing order at different fixed times. The wavelet coefficients exhibit a clear rapid decay beyond the few significant modes corresponding to the gross shape of the Maxwellian distribution. A similar trend is observed in the coefficients of the POD expansion shown in the upper right panel of Fig. III.1.3. However, in the wavelet

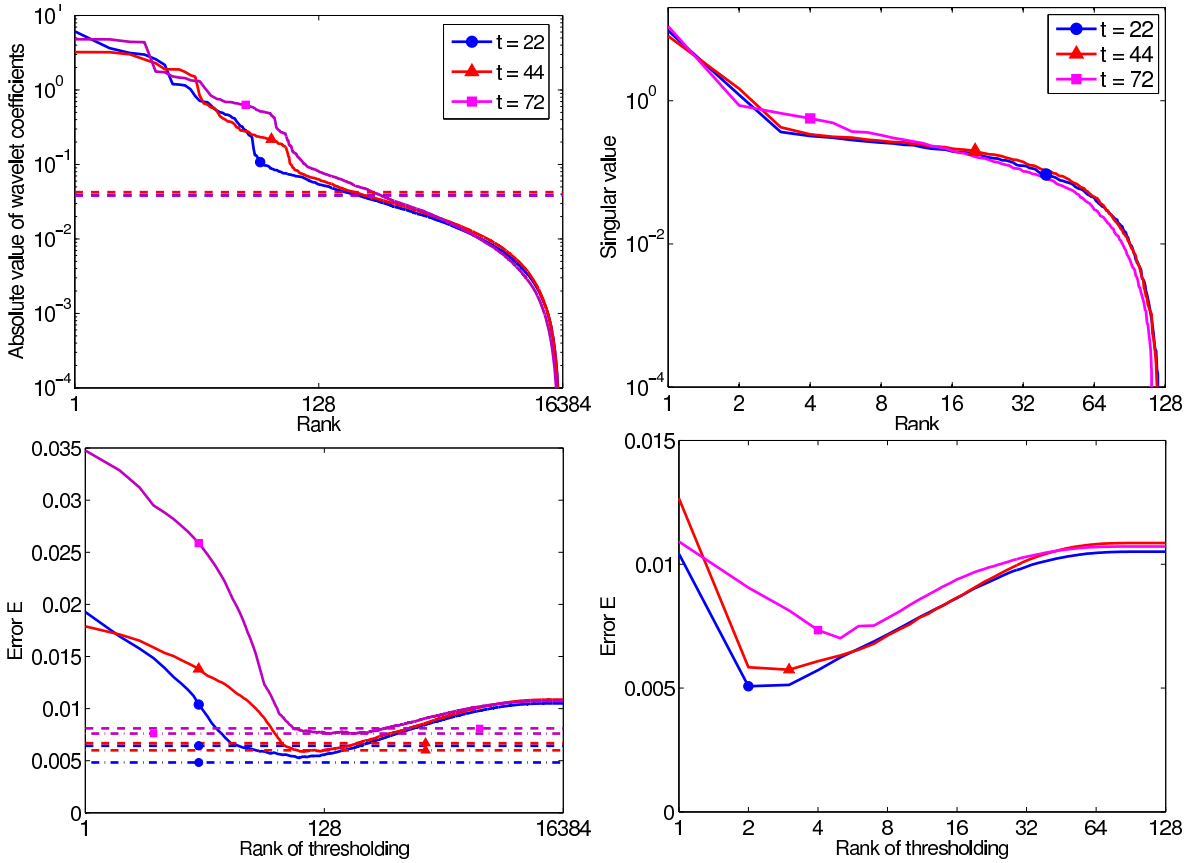


Figure III.1.3: Wavelet and POD analyses of collisional relaxation particle data at different fixed times, with $N_p = 10^5$. Top left: absolute values of the wavelet coefficients sorted by decreasing order (full lines), and thresholds given by the global wavelet shrinkage algorithm (dashed lines). Top right: singular values of the histogram used to construct f^P. Bottom left: error estimate $\frac{e^{1/2}}{N_g^2}$ with respect to the run for $N_p = 10^6$ as a function of the number of retained wavelet coefficients (full lines), error obtained when using the global wavelet shrinkage threshold (dashed lines), and error obtained using the WBDE method (dash-dotted lines). Bottom right: error estimate $\frac{e^{1/2}}{N_g^2}$ for f^P as a function of the number l of retained singular values.

	$t = 28$	$t = 44$	$t = 72$
f^H	0.14	0.17	0.12
f^P	0.068	0.090	0.094
f^W	0.064	0.094	0.088

Table III.1.2: Normalized root mean squared error e_0 (III.1.25) for the histogram, POD and WBDE estimates of the particle distribution function for $N_p = 10^5$ at three different times of the Maxwellian relaxation problem.

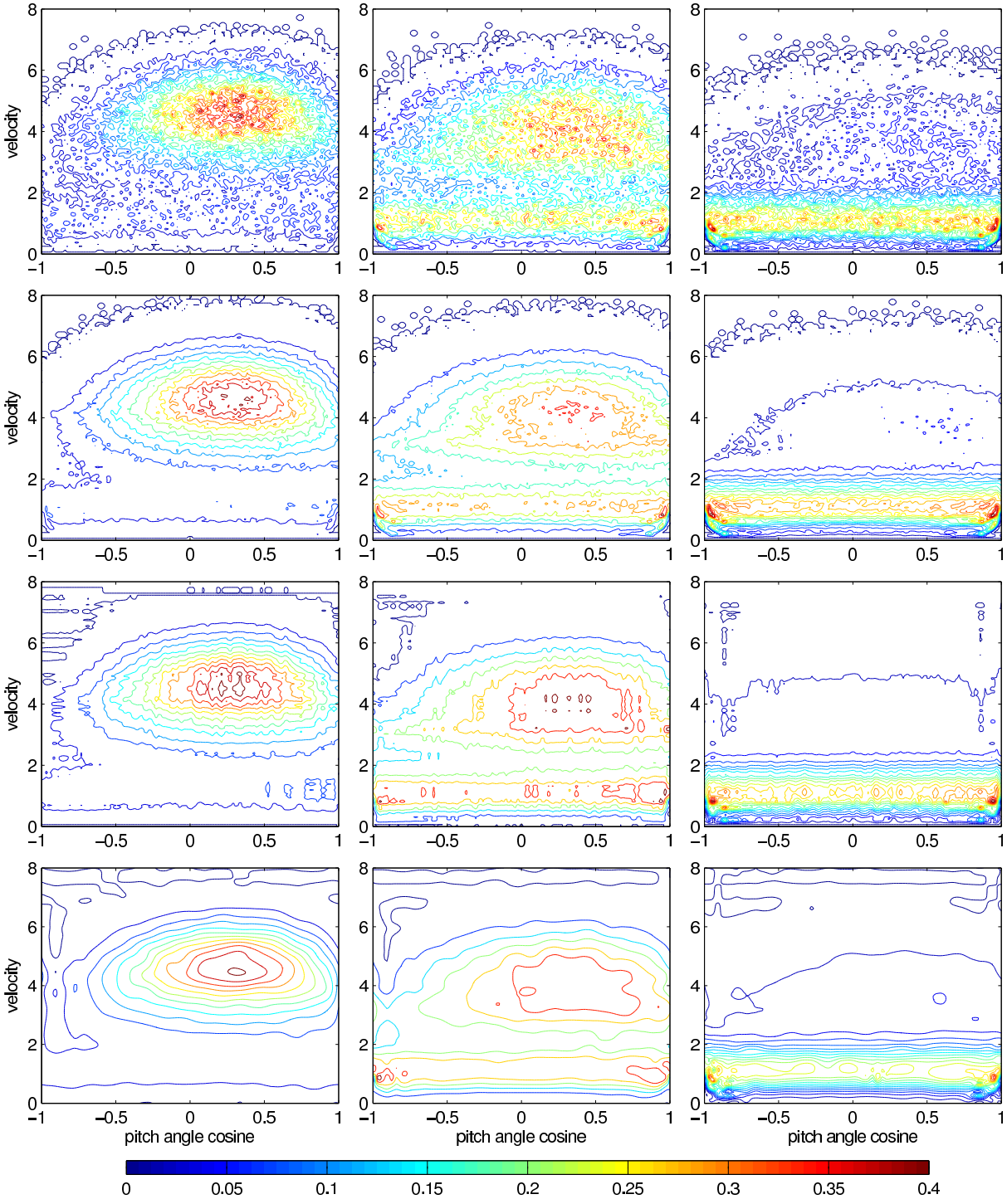


Figure III.1.4: Contour-plots of estimates of f for the collisional relaxation particle data. First row: Histogram method estimated using $N_p = 10^5$ particles. Second row: Histogram method estimated using $N_p = 10^6$ particles. Third row: POD method estimated using $N_p = 10^5$ particles. Fourth row: WBDE method estimated using $N_p = 10^5$ particles. The three columns correspond to $t = 28$, $t = 44$ and $t = 72$ respectively. The plots show twenty isolines, equally spaced in the interval $[0, 0.4]$.

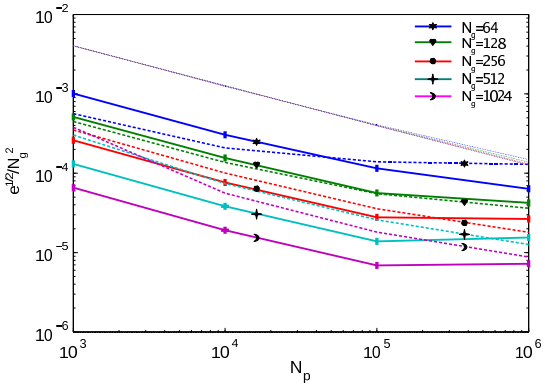


Figure III.1.5: Reconstruction error, $\frac{e^{1/2}}{N_g^2}$, as a function of N_p for the collisional relaxation particle data corresponding to the Maxwellian equilibrium state. Bold solid lines correspond to the WBDE method, bold dashed lines correspond to the POD method, and thin dashed lines correspond to the histogram method.

case the exponential decay starts after more than 100 modes, whereas in the POD case the exponential decay starts after only one mode.

The two panels at the bottom of Fig. III.1.3 show the square root of the reconstruction error normalized by N_g, $\sqrt{e}/N_g^2$, in the WBDE and POD methods. Because in this case we do not have access to the exact solution of the corresponding Fokker-Planck equation at the prescribed time, we used f^H computed using $N_p = 10^6$ particles as the reference density f^{ref} in Eq. (III.1.24). The error observed when applying a global threshold to the wavelet coefficients (bottom left panel in Fig. III.1.3) is minimal when around 100 modes are kept whereas in the POD case (bottom right panel in Fig. III.1.3) the minimal error is reached with about two or three modes. Fig. III.1.3 also shows the wavelet threshold obtained by applying the iterative algorithm based on the stationary Gaussian white noise hypothesis (Azzalini et al., 2004; Farge et al., 2006). The error corresponding to this threshold is larger than the optimal error because the noise in this problem is very non-stationary due to the lack of statistical fluctuations in the regions were particles are absent. In contrast, the error corresponding to the WBDE procedure (dash-dotted line in Fig. III.1.3) is typically smaller than the optimal error obtained by global thresholding. This is not a contradiction, because the WBDE procedure is not a global threshold, but a level dependent threshold.

Figure III.1.4 compares at different times the densities estimated with the WBDE and the POD (retaining only three modes) methods using $N_p = 10^5$ particles with the histograms computed using $N_p = 10^5$ and 10^6 particles. The key feature to observe is that the level of smoothness of f^W and f^P corresponding to $N_p = 10^5$ is similar, if not greater, than the level of smoothness in f^H computed using ten times more particles, i.e. $N_p = 10^6$ particles. Table III.1.2 summarizes the normalized reconstruction errors for $N_p = 10^5$ according Eq. (III.1.24) using f^H with $N_p = 10^6$ as f^{ref}. The WBDE and POD denoising methods offer a significant improvement, approximately by a factor 2, over the raw histogram method.

A more detailed comparison of the estimates can be achieved by focusing on the Maxwellian final equilibrium state

$$f_M(v) = \frac{2}{\sqrt{\pi}} v^2 e^{-v^2} , \qquad \text{(III.1.29)}$$

where, as in Eq. (III.1.28), the v^2 metric factor has been included in the definition of the distribution. For this calculations we considered sets of particles sampled from Eq. (III.1.29) in the compact domain $[-1,1]\times[0,4]$. Since f_M is an exact equilibrium solution of the Fokker-Plack equation, the ensemble of particles is in statistical equilibrium but it exhibits fluctuations due to the finite number of particles. Figure III.1.5 shows the dependence of the square root of the reconstruction error, e (normalized by N_g^2) on the number of particles N_p and the grid resolution N_g for the WBDE and POD methods. The main advantage of this example is that the exact density f^M can be used as the reference density f^{ref} in the evaluation of the error.

III.1.3.2 Collisional guiding center transport in toroidal geometry

The previous example focused on collisional dynamics. However, in addition to collisions, plasma transport involves external and self-consistent electromagnetic fields and it is of interest to test the particle density reconstruction algorithms in these more complicated settings. As a first step for solving this challenging problem, we consider a plasma subject to collisions and an externally applied fixed magnetic field in toroidal geometry. The choice of the field geometry and structure was motivated by problems of interest to magnetically confined fusion plasmas. The data was presented and analyzed using POD method in Ref. (del Castillo-Negrete et al., 2008). The phase space of the simulation is five dimensional. However, as in Ref. (del Castillo-Negrete et al., 2008), we limit attention to the denoising of the particles distribution function along two coordinates corresponding to the poloidal angle $\theta\in[0,2\pi]$ and the cosine of the pitch angle $\mu\in[-1,1]$. The remaining three coordinates have been averaged out for the purpose of this study. The θ coordinate is periodic, but the pitch coordinate μ is not.

An important issue to consider is that the data was generated using a δf code (DELTA5D). Based on an expansion on $\rho/L\ll 1$ (where ρ is the characteristic Larmor radius and L a typical equilibrium length scale) the distribution function is decomposed into a Maxwellian part f_M and a first-order perturbation δf represented as a collection of particles (markers)

$$\delta f(\mathbf{x})=\sum_n W_n\delta(\mathbf{x}-\mathbf{X}_n)\,, \tag{III.1.30}$$

like in Eq. (III.1.1) except that each marker is assigned a time dependent weight W_n whose time evolution depends on the Maxwellian background (Parker and Lee, 1993). The direct use of $\delta f(\mathbf{x})$ is problematic in the WBDE method because δf is not a probability density. To circumvent this problem the WBDE method was applied after normalizing the δ_f distribution so that $\int|\delta f|^H=1$, on a 128×128 grid.

Figure III.1.6 shows contour plots of the histogram f^H corresponding to $N_p=32\times10^3$, $N_p=64\times10^3$, and $N_p=1024\times10^3$ along with the WBDE and POD reconstructed densities. The POD reconstructions were done using $r=3$ modes, as in Ref. (del Castillo-Negrete et al., 2008). It is observed that comparatively high levels of smoothness can be achieved with considerably less particles by using either the WBDE or POD reconstruction methods. The WBDE method provides better results for the $\delta f\sim0$ contours. This is because POD modes are tensor product functions, that have difficulties in approximating the triangular shape of these contour lines. Note that the boundary artifacts due to periodization of the Daubechies wavelets do not seem to be very critical. The large wavelet coefficients associated

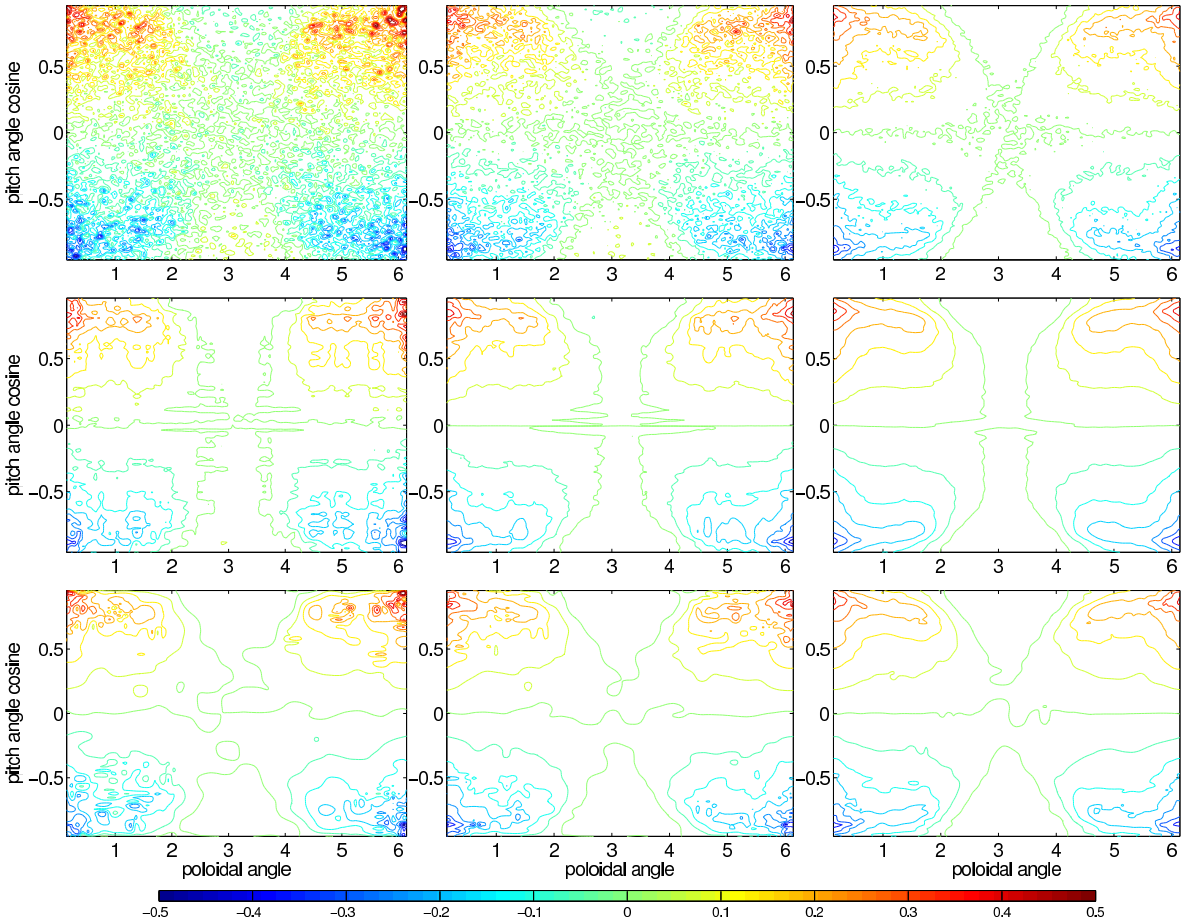


Figure III.1.6: Contour plots of estimates of f for the collisional guiding center transport particle data: Histogram method (first row), POD method (second row), and WBDE method (third row). The left, center and right columns correspond to $N_p = 32 \cdot 10^3$ (left), $N_p = 128 \cdot 10^3$ (middle) and $N_p = 1024 \cdot 10^3$ (right) respectively. The plots show seventeen isolines equally spaced within the interval $[-0.5, 0.5]$.

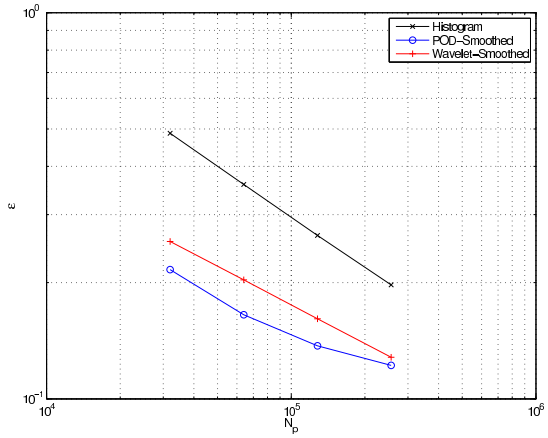


Figure III.1.7: Error estimate, $\frac{e^{1/2}}{N_g^2}$, for collisional guiding center transport particle data according to the histogram, the POD, and the wavelet methods.

with the discontinuity between the values of δf at $\mu = \pm 1$ are not thresholded, so that the discontinuity is preserved in the denoised function. Figure III.1.7 compares the reconstruction errors in the WBDE, POD, and histogram methods as functions of the number of particles. To evaluate the error we used f^H computed with $N_p = 1024 \times 10^3$ as the reference density f^{ref}. As in the collisional transport problem, the error is reduced roughly by a factor 2 for both methods compared to the raw histogram. Note that the scaling with N_p is slightly better for WBDE than for POD.

III.1.3.3 Collisionless electrostatic instabilities

In this section we apply the WBDE and POD methods to reconstruct the single particle distribution function from discrete particle data obtained from PIC simulations of a Vlasov-Poisson plasma. We consider a one-dimensional, electrostatic, collisionless electron plasma with an ion neutralizing background in a finite size domain with periodic boundary conditions. The dynamics of the distribution function is governed by the system of equations

$$\partial_t f + v\partial_x f + \partial_x \phi \partial_v f = 0 \tag{III.1.31}$$

$$\partial_x^2 \phi = \zeta \int f(x, v, t) dv - 1 \,, \tag{III.1.32}$$

where the variables have been non-dimensionalized using the Debye length as length scale and the plasma frequency as time scale, and ζ is the normalized length of the system. Following the standard PIC methodology (Birdsall and Langdon, 1985), we solve the Poisson equation on a grid and integrate the particle equations in time using a leap-frog method. The reconstruction of the charge density uses a triangular shape function. We consider two initial conditions: the first one leads to a bump on tail instability, and the second one to a two streams instability.

III.1.3.4 Bump on tail instability

To trigger a bump on tail instability we initialized ensembles of particles by sampling the distribution function

$$f_0(x, v) = \frac{2}{3\pi\zeta} \frac{1 - 2qv + 2v^2}{(1 + v^2)^2} \tag{III.1.33}$$

using a pseudo-random number generator. This equilibrium is stable for $q \leq 1$ and unstable for $q > 1$. The dispersion relation and linear stability analysis for this equilibrium studied in Ref. (del Castillo-Negrete, 1998) was used to benchmark the PIC code as shown in Fig. III.1.8. In all the computations presented here $q = 1.25$ and $N_p = 10^4$, 10^5 and 10^6. The spatial domain size was set to $\zeta = 16.52$ to fit the wavelength of the most unstable mode.

Since the value of q is relatively close to the marginal value, the instability grows weakly and is concentrated in a narrow band in phase space centered around the point where the bump is located, $v \approx 1$ in this case. In order to unveil the nontrivial dynamics we focus the analysis in the band $v \in (-3, 3)$, and plot the departure of the particle distribution function from the initial background equilibrium. The POD method is applied directly to $\delta f^H = f^H(x, v, t) - f_0(x, v)$, but the WBDE method is applied to the full $f^H(x, v, t)$, and $f_0(x, v)$ is subtracted only for visualization. Note that because we are considering only a subset of phase space, the effective numbers of particles, $N_p = 7318$, $N_p = 73143$ and

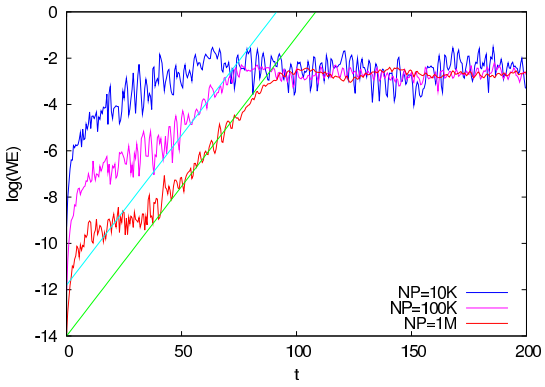


Figure III.1.8: Electrostatic energy as a function of time in the Vlasov-Poisson PIC simulations of the bump on tail instability for different numbers of particles. The straight lines denote the growth rate predicted by linear stability theory (del Castillo-Negrete, 1998).

	$N_p = 10^4$	$N_p = 10^5$
f^H	0.443	0.140
f^P	0.163	0.090
f^W	0.173	0.086

Table III.1.3: Comparison of normalized root mean squared errors e_0 (III.1.25) for the raw histogram and for the WBDE and POD methods, for the bump-on-tail instability at $t = 149$, depending on the number of particles. The simulation with $N_p = 10^6$ is used as a reference to compute the error.

$N_p = 731472$, are smaller than the nominal numbers of particles, $N_p = 10^4$, $N_p = 10^5$ and $N_p = 10^6$ respectively.

Figure III.1.9 shows contour plots of δf for different number of particles. Since the instability is seeded only by the random fluctuations in the initial condition, increasing N_p delays the onset of the linear stability and this leads to a phase shift of the nonlinear saturated regime. To aid the comparison of the saturated regime for different numbers of particles we have eliminated this phase shift by centering the peak of the particle distributions in the middle of the computational domain. A 256×256 grid was used in the WBDE method, and a 50×50 grid was used for the histogram and the POD methods. The thresholds for the POD method where $r = 1$, $r = 2$, and $r = 3$ for $N_p = 10^4$, $N_p = 10^5$ and $N_p = 10^6$, respectively. Except for the case where $N_p = 10^4$, both the POD and WBDE estimates are very smooth, in agreement with the expected behavior of f for this instability. It is observed that the level of smoothness of the histogram estimated using 10^6 particles is comparable to the level of smoothness achieved after denoising using only 10^5 particles. One should mention that for scales between L and J occurring in the WBDE algorithm we find that none of the wavelet coefficients are above the thresholds at each scale. In fact, a simple KDE estimate with a large enough smoothing scale would probably do the job pretty well for this kind of instabilities which do not induce abrupt variations in f. Table III.1.3 shows the POD and WBDE reconstruction errors for $N_p = 10^4$ and $N_p = 10^5$. The error is computed using formula (III.1.25), taking for f_{ref} the histogram obtained from the simulation with $N_p = 10^6$.

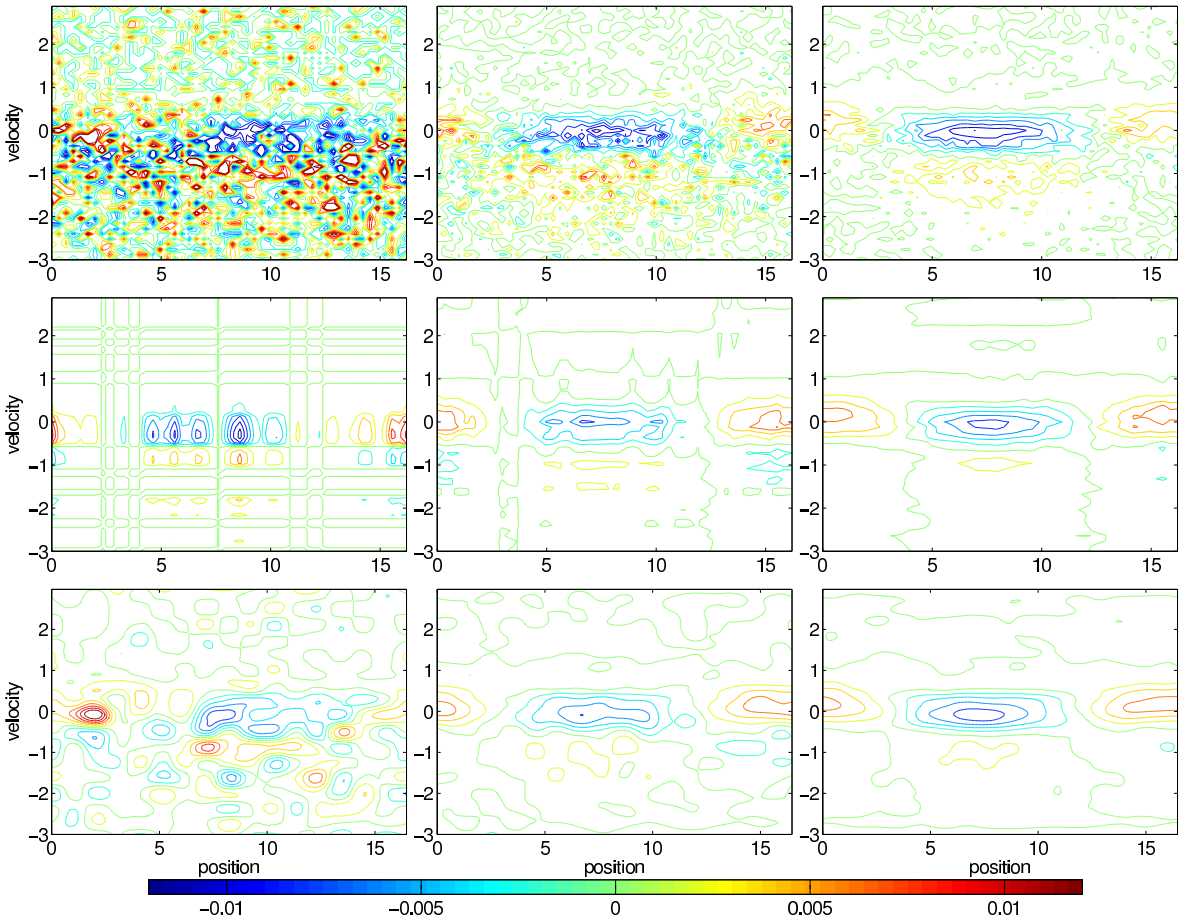


Figure III.1.9: Contour plots of estimates of δf for the bump-on-tail instability PIC data at $t = 149$: Histogram method (first row), POD method (second row), and WBDE method (third row), The left, center and right columns correspond to $N_p = 10^4$, $N_p = 10^5$ and $N_p = 10^6$ particles respectively. The plots show thirteen contour lines equally spaced within the interval $[-0.0120.012]$.

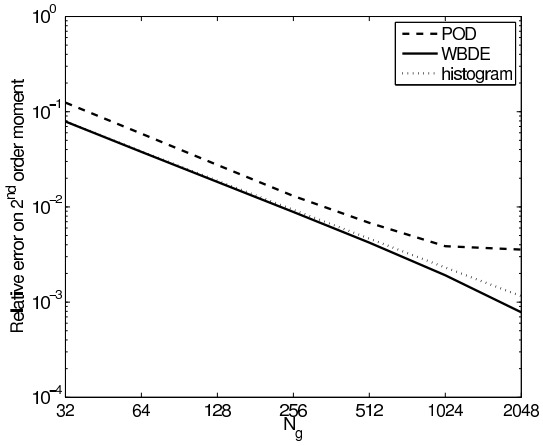


Figure III.1.10: Relative error on the second order moment as a function of the grid resolution, N_g, in the POD, WBDE, and histogram methods for the bump on tail instability particle data at $t = 149$, with $N_p = 10^6$ particles.

Figure III.1.10 shows the relative error on the second order moment :

$$\frac{|\mathcal{M}^W_{v,2} - \mathcal{M}^\delta_{v,2}|}{\mathcal{M}^\delta_{v,2}}$$

where $\mathcal{M}^W_{v,2}$ is defined by (III.1.19). A similar quantity is also represented for f^H and f^P. Time and number of particles are kept fixed at $t = 149$ and $N_p = 10^6$, and only grid resolution is varied. As expected, f^H and f^W conserve the second order moment with accuracy $O(N_g^{-1})$. The errors corresponding to f^P is of the same order of magnitude but seems to reach a plateau for $N_g \simeq 1024$. This may be due to the fact that for $N_g \geq 1024$, there is less than one particle per cell of the histogram used to compute f_P.

III.1.3.5 Two-streams instability

As a second example we consider the standard two-streams instability with an initial condition consisting of two counter-propagating cold electron beams initially located at $v = -1$ and $v = 1$. This case is conceptually different to the previous one because the initial condition depends trivially on the velocity. Therefore, there is no statistical error in the sampling of the distribution and the noise builds up only due to the self-consistent interactions between particles. In other words, there is initially a strong correlation between particles' coordinates, which will eventually almost vanish. This situation offers a way to test robustness of the WBDE method with respect to the underlying decorrelation hypothesis.

The analysis is focused on four stages of the instability, corresponding to $t = 40$, 60, 100, and 400. Fig. III.1.11 shows a comparison of the raw histogram, the POD and the WBDE reconstructed particle distribution functions at these four instants. Grid sizes were $N_g = 1024$ for the WBDE estimate, and $N_g = 128$ for the two others. For $t = 40$, no noise seems to have affected the particle distribution yet, therefore a perfect denoising procedure should conserve the full information about the particle positions. Although WBDE introduces some artifacts in regions of phase space that should contain no particles at all, it remarkably preserves the global structure of the two streams. This is possible thanks to the numerous wavelet coefficients close to the sharp features in f that are above the thresholds, in contrast to the bump-on-tail case. On the next snapshot at $t = 60$, the filaments have overlapped and the system is beginning to loose its memory due to numerical round-off errors. The fastest filaments still visible on the histogram are not preserved by WBDE, but the most active regions are well reproduced. At $t = 100$, the closeness between the histogram and the WBDE estimate is striking. To put it somewhat subjectively, one may say that WBDE did not consider most of the rough features present at this stage as 'noise', since they are not removed. Only with the last snapshot at $t = 400$ does the WBDE estimate begin to be smoother than the histogram, suggesting that the nonlinear interaction between particles has introduced randomization in the system.

The POD method is able to track very well the fine and coarse scale structures of the particle density using a small number of modes. In particular, for $t = 40$, 60, 100, and 400 only $r = 28$, $r = 27$, $r = 18$, and $r = 5$ modes were kept. The decrease of the number of modes with time results from the lost of fine scale features in the distribution function.

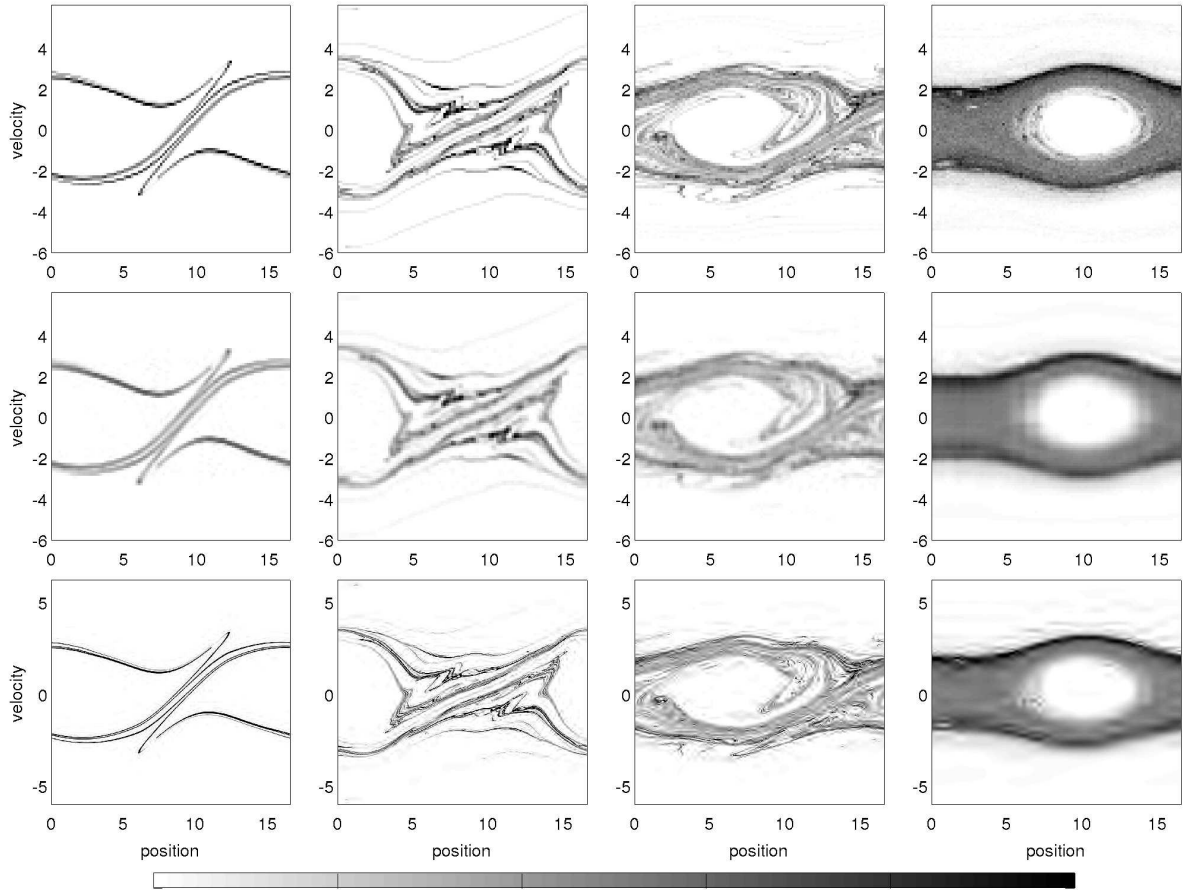


Figure III.1.11: Contour plots of estimates of f for the two streams instability PIC data at times $t = 40$, $t = 60$, $t = 100$ and $t = 400$ (left to right). Histogram method (first row), WBDE method (second row), and POD method (third row), The gray level tone varies uniformly in the interval $[0, A]$, where $A = 0.15$, $A = 0.08$, $A = 0.05$ and $A = 0.025$ in the first, second, third and fourth columns respectively.

III.1.4 Summary and Conclusion

Wavelet based density estimation was investigated as a post-processing tool to reduce the noise in the reconstruction of particle distribution functions starting from discrete particle data. This is a problem of direct relevance to particle-based transport calculations in plasma physics and related fields. In particular, particle methods present many advantages over continuum methods, but have the potential drawback of introducing noise due to statistical sampling.

In the context of particle in cell methods this problem is typically approached using finite size particles. However, this approach, which is closely related to the kernel density estimation method in statistics, requires the choice of a smoothing scale, h, (e.g., the standard deviation for Gaussian shape functions) whose optimal value is not known a priori. A small h is desirable to fit as many Debye wavelengths as possible, whereas a large h would lead to smoother distributions. This situation results from the compromise between bias and variance in statistical estimation. To address this problem we proposed a wavelet based density estimation (WBDE) method that does not require an a priori selection of a global smoothing scale and that its able to adapt locally to the smoothness of the density based on the given discrete data. The WBDE was introduced in statistics (Donoho et al., 1996). In this paper we extended the method to higher dimension and applied it for the first time to particle-based calculations. The resulting method exploits the multiresolution properties of wavelets, has very weak dependence on adjustable parameters, and relies mostly on the raw data to separate the relevant information from the noise.

As a first example, we analyzed a plasma collisional relaxation problem modeled by stochastic differential equations. Thanks to the sparsity of the wavelet expansion of the distribution function, we have been able to extract the information out of the statistical fluctuations by nonlinear thresholding of the wavelet coefficients. At late times, when the particle distribution approaches a Maxwellian state, we have been able to quantify the difference between the denoised particle distribution function and its analytical counterpart, thus demonstrating the improvement with respect to the raw histogram. The POD-smoothed and wavelet-smoothed particle distribution functions were shown to be roughly equivalent in this respect. These results were then extended to a more complex situation simulated with a δf code. Finally, we have turned to the Vlasov-Poisson problem, which includes interactions between particles via the self-consistent electric field. The POD and WBDE methods were shown to yield quantitatively similar results in terms of mean squared error for a particle distribution function resulting from nonlinear saturation after occurrence of a bump-on-tail instability. We have then studied the denoising algorithm during nonlinear evolution after the two-streams instability starting from two counter-streaming cold electron beams. This initial condition violates the decorrelation hypothesis underlying the WBDE algorithm, and thus offers a good way to test its robustness regarding this aspect. The WBDE method was shown to yield qualitatively good results without changing the threshold values.

Defining a 'mode' as one term in the decomposition of the particle distribution function into a sum, it appears that the number of modes used in the POD reconstruction algorithm is considerably less than the number of modes needed using wavelets. But to make a systematic comparison when it comes to storage requirements or data compression, one has to keep in mind that POD modes are empirical and problem dependent whereas wavelet modes are known a priori. Because of this, the specification of each POD mode requires $2N_g + 1$ components whereas the specification of each wavelet component only requires one

number: the amplitude of the mode. A potential limitation of the POD method is the the lack of systematic a priori thresholding criteria to determine the optimal number of modes. However, the information contained in the decay of the spectrum of the singular values can be used to determine optimal ranks. For example, in Ref. (del Castillo-Negrete et al., 2008) a criterion based on the relative rate of decay of the spectrum provided consistent good results for denoising Monte-Carlo particle data.

To test and compare the proposed denoising algorithms we have focused on the distribution function because it is in some sense the most basic and fundamental quantity in PIC and Monte-Carlo simulations. It would be of interest to explore the use of other fields like for example fluxes, charge densities, or derivatives of f rather than f itself. Lack of space did not low us to elaborate more on this issue that we look forward to address in a future publication. Regarding this, it is important to keep in mind that what to chose to compare the algorithms, beyond the obvious choice of the raw f, depends on the physics of the specific problem of interest. For example, calculations that require evaluation of quasi-linear fluxes in wave particle-interaction problems might benefit from denoising gradients of f. On the other hand, active denoising Vlasov-Poisson codes might only require denoising of the charge distribution (i.e. the integral of f in velocity) whereas denoising the current (first velocity moment of f) would be important in Vlasov-Maxwell codes. Other specific needs might arise when coupling particle codes to continuum codes, e.g. when performing extended MHD calculations. It may also be of interest to focus on the reconstruction error of the force fields, which determine the evolution of the simulated plasma. These forces depend on f through integrals, and statistical analysis of the estimation of f using weak norms (Victory and Allen, 1991) might be beneficial.

The computational cost of our method scales linearly with the number of particles and with the grid resolution. Therefore, WBDE is an excellent candidate to be performed at each time step during the course of a simulation. Once the wavelet expansion of the denoised particle distribution function is known, it is possible to continue using the wavelet representation to solve the Poisson equation (Jaffard, 1992) and to compute the forces. The moment conservation properties that we have demonstrated in this paper should mitigate the unavoidable dissipative effects implied by the smoothing stage. In Ref. (McMillan et al., 2008), a dissipative term was introduced in a global PIC code to avoid unlimited growth of particle weights in δf codes, and this was shown to improve long time convergence of the simulations. It would be of interest to assess if the nonlinear dissipation operator corresponding to WBDE has the same effect.

There are several potential extensions and applications of the techniques and results presented. Some of these include high dimensional problems, active denoising, and applications to more complex plasma models. The implementation of the WBDE algorithm to high (greater than two) dimensions is in principle straightforward. The POD method on the other hand can be more challenging since the SVD is applied to matrices. One simple way to circumvent this problem is to "fold" high dimensional data into matrices.However, this straightforward approach can be numerically inefficient. A promising alternative would be to use tensor decomposition techniques like the method of generalized low rank approximation of matrices used in (del Castillo-Negrete et al., 2008). The problem of active denoising, and the application of the reconstruction algorithms to more complex plasma models like Vlasov-Maxwell and gyrokinetics is a key follow up of the results presented here. Also, it would be of interest to explore applications to the problem of coupling particle and continuum codes in extended MHD and radio-frequency heating studies in fusion plasmas.

Acknowledgements

We thank Don Spong for providing the DELTA5D Monte-Carlo guiding center simulation data in Fig.6, originally published in Ref. (del Castillo-Negrete et al., 2008). We also thank Xavier Garbet for his comments on the paper and for pointing out several key references, and Bill Dorland for motivating discussions. MF and KS acknowledge financial support by ANR under contract M2TFP, Méthodes multiéchelles pour la turbulence dans les fluides et les plasmas. MF is grateful to the Wissenschaftskolleg zu Berlin for hospitality while revising this paper. DCN and GCH acknowledge support from the Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the U.S. Department of Energy under contract DE-AC05-00OR22725. DCN also gratefully acknowledges the support and hospitality of the École Centrale de Marseille for the three, one month visiting positions during the elaboration of this work. This work, supported by the European Communities under the contract of Association between EURATOM, CEA and the French Research Federation for fusion studies, was carried out within the framework of the European Fusion Development Agreement. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

III.2 Particle-in-Wavelets scheme for the 1D Vlasov-Poisson equations

The contents of this section have been submitted for publication in ESAIM:Proceedings, see (Nguyen van yen et al., 2010d)

Introduction

The state of a hot plasma out of local thermodynamic equilibrium is characterized by a particle distribution function depending on position, velocity, and time. When the effects of collisions are neglected, this distribution satisfies the partial differential equation (PDE) known as the Vlasov equation. In the presence of simplifying hypotheses and/or special symmetries, the complexity of the problem can be reduced to permit a resolution of the Vlasov equation using the Eulerian PDE discretization approach which requires a mesh in position and velocity. Recent progresses have pushed back the limit of the feasible to include for example the 5D gyrokinetic approximation (Grandgirard et al., 2006; Wang et al., 2006; Idomura et al., 2007). However the full equation with six variables plus time remains a major computational challenge, which could take as much advantage of better numerical methods as of further machine development and parallelization.

In the past, even more stringent limitations in memory size and computing power have driven physicists to develop specific simulation methods, which approximate the plasma by a collection of macro-particles (Birdsall and Langdon, 1985; Hockney and Eastwood, 1988) and can be linked to Monte-Carlo methods (Aydemir, 1994). Although the precision of Monte-Carlo methods is limited by sampling error, they are known to be competitive for problems involving integrals over many dimensions. For the specific problem of solving the Vlasov equations, it is well known that particle methods offer a more straightforward implementation and an easier parallelization than their Eulerian counterparts. However, the debate over the best compromise in terms of overall development and simulation cost versus precision remains unsettled. Current outstanding applications of particle methods include magnetic confinement fusion (Jolliet et al., 2007; Heikkinen et al., 2008), and laser-plasma acceleration (Cormier-Michel et al., 2008). But in the last decade, there has been a tendency to challenge some results obtained using legacy particle methods on the grounds that the simulations were underresolved (Nevins et al., 2005).

Plasma simulation using particles is dominated by derivatives of the finite size particle (FSP) method (Birdsall and Langdon, 1985; Hockney and Eastwood, 1988), to whom a vast amount of research has afforded the status of a reference. FSP schemes are obtained by convolving the empirical density corresponding to the distribution of marker particles by a fixed kernel, also known as charge assignment function (Hockney and Eastwood, 1988). Their numerical properties are well understood empirically and they also benefit from an elegant physical interpretation (Langdon and Birdsall, 1970). The most widely used FSP method, particle-in-cell (PIC), uses a triangular kernel. Higher order variants have been studied in the past (Okuda et al., 1979) but their computational cost has made their use relatively rare, since precision and reduction of noise were not the main concern at the time.

In (Nguyen van yen et al., 2010a), we have studied another way of estimating the density from the particle positions, called wavelet-based density estimation (WBDE). WBDE was

first introduced in the more general context of nonparameteric density estimation (Donoho et al., 1996), as an alternative to the classical kernel density estimation (KDE) method. Instead of being convolved with a kernel, the empirical density is projected onto a finite dimensional linear space spanned by a family of orthogonal wavelets. The projection space is determined from the data itself, which allows for a refined representation around sharp features, and could make the method more precise than FSP for a given computational cost.

The aim of this paper is to present a new scheme for the 1D Vlasov-Poisson equations, that we call particle-in-wavelets (PIW) because it relies crucially on wavelet expansions of the Dirac delta functions corresponding to every particle. WBDE is used for the density estimation step in the PIW scheme. The information pertaining to each particle thus gets spread to neighboring wavelets, in the same manner as it gets spread among neighboring grid points in the PIC case. The general philosophy of this work has been to build the scheme from scratch, taking as much advantage of wavelets as possible, so as to have a proof of principle of various properties specific to wavelets. We are focusing mostly on understanding the precision of the scheme and the conservation properties, as opposed to optimizing numerical efficiency.

Other methods have been proposed, and sometimes successfully implemented, to improve PIC simulations, but they all rely on the application of a linear operator to the particle distribution function in order to reduce the noise. Linear denoising operators are optimal when the regularity of the signal to be denoised is homogeneous in space. However, the distribution function within collisionless plasmas is known to develop sharp features. Therefore a nonlinear thresholding approach could be more efficient, and this work is a first step in this direction.

We first recall some background, on the Vlasov-Poisson equations in one space dimension on the one hand, and on wavelets on the other hand. Reference semi-Lagrangian Eulerian and PIC schemes are also described. After that we proceed to define the various steps of the PIW (particle-in-wavelets) wavelet-based scheme. Finally, numerical results from the three schemes are presented and the properties of the new PIW scheme are assessed in detail.

III.2.1 Background

III.2.1.1 Vlasov equations

The Vlasov equation governs the evolution of the particle distribution function $f(x,v,t)$ of a given species in a collisionless plasma. In the 1D case which we focus on, v varies in $\mathbb{R}$, and for simplicity, we assume periodicity in the x direction, i.e. $x \in \mathbb{T} := \frac{\mathbb{R}}{l\mathbb{Z}}$, where l is the period. $\Gamma = \mathbb{T} \times \mathbb{R}$ is the phase space for one particle in the plasma. We restrict ourselves to the case where the ion density is assumed to remain constant, and the magnetic field is neglected. The system is then fully described by the Vlasov-Poisson equations, which can be written in non-dimensional form as follows:

$$\partial_t f + v\partial_x f + \partial_x \phi \partial_v f = 0 \quad \text{(III.2.1)}$$

$$\partial_{xx}\phi + 1 - l\int_{\mathbb{R}} f(x,v,t)\mathrm{d}v = 0 \quad \text{(III.2.2)}$$

where $\phi(x,t)$ is the electric potential. In the following, we denote by $E(x,t) = -\partial_x\phi(x,t)$ the electric field, and by $\rho(x,t) = \int_{\mathbb{R}} f(x,v,t)\mathrm{d}v$ the charge density. Note that we have

adopted here the following slightly uncommon normalizations for ρ and f:

$$\int_\Gamma f(x,v,t)\mathrm{d}x\mathrm{d}v = 1 \quad \text{(III.2.3)}$$

$$\int_\mathbb{T} \rho(x,t)\mathrm{d}x = 1 \quad \text{(III.2.4)}$$

We shall also need the definitions of two physical quantities associated to the Vlasov-Poisson problem, namely the electrostatic energy

$$\mathcal{V}(t) = \frac{1}{2}\int_\mathbb{T} |E(x,t)|^2 \quad \text{(III.2.5)}$$

and the total energy

$$\mathcal{E}(t) = \frac{1}{2}\int_\Gamma v^2 f(x,v,t)\mathrm{d}x\mathrm{d}v + \mathcal{V}(t), \quad \text{(III.2.6)}$$

the latter being constant for solutions of the Vlasov-Poisson equations.

The goal of this paper is to describe a new numerical scheme for solving (III.2.1-III.2.2) together with an initial condition for f. We will use the fact that any solution of (III.2.1) is constant along the characteristic trajectories defined by:

$$\begin{cases} x'(t) = v(t) \\ v'(t) = -E\left(x(t), v(t), t\right) \end{cases} \quad \text{(III.2.7)}$$

Existing approaches to the same problem are divided into two main categories, that were already mentioned in the introduction:

- particle solvers, which integrate the characteristic equations (III.2.7) and discretize the potential on a grid in the x variable,
- Eulerian solvers where the full distribution function f is discretized on a phase space mesh or grid.

The first category is especially appealing when the phase space has dimension 4 or more, since it is then very costly to use an Eulerian approach. However, the discretization of f using particles is subject to statistical sampling noise which normally makes it of rather low precision (Nevins et al., 2005). In particular, for the 1D problem that we are considering here, we expect Eulerian solvers to be much more precise for a given computational cost. This allows us to take the solution obtained with a high precision Eulerian solver as a reference for measuring the convergence of particle methods. We have chosen a well validated semi-Lagrangian (SL) solver using cubic spline interpolation (Sonnendrücker et al., 1999).

It is a reasonable requirement that any new particle method should offer some improvement over the legacy PIC approach. As a contestant to compare with our newly proposed PIW method, we therefore take a PIC solver, which uses a triangular charge assignment function to estimate the charge density, and Fourier transforms to solve the Poisson equation.

When assessing the convergence of particle solvers, the loading scheme is especially important since it controls the initial amplitude of the noise. Provided that the initial data f_0 is factorized

$$f_0(x,v) = f_{0,x}(x) f_{0,v}(v)$$

$h_0 =$ 0.03858077774789	$h_2 =$ -0.07716155549577	$h_4 =$ 0.74568755893443
$h_1 =$ -0.12696912539621	$h_3 =$ 0.60749164138568	$h_5 =$ 0.22658426519707

Table III.2.1: Scaling function filter coefficients for the C^1 Coiflet family with 2 vanishing moments introduced in (Daubechies, 1993) and used throughout this paper.

we ensure that the initial noise is small using the following classical approach. First, compute a 2D Hammersley sequence (Hammersley et al., 1975), $(a_i, b_i)_{i\in\mathbb{N}}$, Then set as initial particle positions

$$\begin{aligned} x_i(0) &= F_{0,x}^{-1}(a_i) \\ v_i(0) &= F_{0,v}^{-1}(b_i) \end{aligned}$$

where $F_{0,x}^{-1}$ and $F_{0,v}^{-1}$ are the inverse mappings corresponding to the cumulative distribution functions $F_{0,x}$ and $F_{0,v}$ associated respectively to $f_{0,x}$ and $f_{0,v}$, and are in practice approximated using a pre-generated lookup table and polynomial interpolation. With this initialization scheme, called "quiet start", the initial error on the particle distribution function is expected to scale like N_p^{-1}, as opposed to $N_p^{-\frac{1}{2}}$ for a random start.

III.2.1.2 Wavelets

In this section we recall some notions on wavelets. We keep the same notations as in (Nguyen van yen et al., 2010a). More background may be found in the literature (see e.g. (Mallat, 1999)). The construction is done in $\frac{\mathbb{R}}{l\mathbb{Z}}$ with $l = 1$, and the adjustments for $l \neq 1$ will be made further down the moment they are required. Let ψ be a 1-periodic wavelet generating an orthogonal basis of $L^2(\frac{\mathbb{R}}{\mathbb{Z}})$, and φ be the associated scaling function.

The dilated and translated scaling functions and wavelets are denoted

$$\varphi_\lambda(x) = 2^{\frac{j_\lambda}{2}}\varphi(2^{j_\lambda}x - i_\lambda) \quad \text{(III.2.8)}$$

$$\psi_\lambda(x) = 2^{\frac{j_\lambda}{2}}\psi(2^{j_\lambda}x - i_\lambda) \quad \text{(III.2.9)}$$

where $\lambda = (j_\lambda, i_\lambda)$ is a multi-index characterizing the scale j_λ and position i_λ of a wavelet. We denote by V_J the 2^J-dimensional linear space spanned by $\left(\varphi_{(J,i)}\right)_{0\leq i\leq 2^J-1}$. An important property of the function φ is that it satisfies the two-scale recurrence equation (Mallat, 1999):

$$\varphi(t) = \sqrt{2}\sum_{i=0}^{S-1} h_i\varphi(2t - i) \quad \text{(III.2.10)}$$

where $(h_i)_{0\leq i\leq S-1}$ is called the scaling function filter.

The scaling function and wavelet coefficients of a function $\mathbf{u} \in L^2(\mathbb{T})$ are defined by

$$\overline{\mathbf{u}}_\lambda = \langle \mathbf{u} \mid \varphi_\lambda \rangle \quad \text{(III.2.11)}$$

$$\widetilde{\mathbf{u}}_\lambda = \langle \mathbf{u} \mid \psi_\lambda \rangle \quad \text{(III.2.12)}$$

i	0	1	2	3	4
$r_i^{(1)}$	0	-0.619368727403	-1.403784906417	2.427684222146	-0.839028531550
$r_i^{(2)}$	0.622642595976	-0.389477243531	0.179082676222	-0.326853461357	0.225926730678

Table III.2.2: Stencils for first and second derivative operators in the R-Coiflet 6 scaling functions basis. Only the values for positive k are given. The values for negative k are obtained from the conditions that $r^{(1)}$ is odd and $r^{(2)}$ is even.

where $\langle \cdot \mid \cdot \rangle$ denotes the usual scalar product in $L^2(\mathbb{T})$. For any positive integer L, $\mathbf{u}$ can be decomposed as follows:

$$\mathbf{u} = \sum_{i=0}^{2^L-1} \overline{\mathbf{u}}_{(L,i)} \varphi_{(L,i)} + \sum_{j=L}^{\infty} \sum_{i=0}^{2^j-1} \widetilde{\mathbf{u}}_{(J,i)} \psi_{(J,i)} \tag{III.2.13}$$

where the first sum is the projection of $\mathbf{u}$ on V_L, and the second sum contains all the additional details needed for the full reconstruction of $\mathbf{u}$.

In the following we shall also need to represent first and second order differential operators in the space V_J. The results necessary to do so were derived in (Beylkin, 1992), and we summarize them here for completeness. For any $u \in V_J$ we have, using (III.2.13):

$$u' = \sum_{i=0}^{2^J-1} \overline{\mathbf{u}}_{(J,i)} \varphi'_{(J,i)}$$

The Galerkin representation of the first order derivation operator in V_J is thus defined by the matrix G such that:

$$\begin{aligned} G_{i_1,i_2} &= \left\langle \varphi_{(J,i_1)} \mid \varphi'_{(J,i_2)} \right\rangle = 2^J \int_{\mathbb{T}} \mathrm{d}x \varphi\left(2^J x - i_1\right) \varphi'\left(2^J x - i_2\right) \\ &= \int_{\mathbb{T}} \mathrm{d}y \varphi\left(y + i_2 - i_1\right) \varphi'(y) := r^{(1)}_{i_2-i_1} \end{aligned}$$

In (Beylkin, 1992), it is shown that the matrix elements $r^{(1)}_{i_2-i_1}$ can be determined by solving a linear system with coefficients depending only on the h_i, and that they satisfy the anti-symmetry relation $r^{(1)}_{-i} = -r^{(1)}_i$, so that applying the matrix G is numerically equivalent to applying a centered finite difference operator. G is thus an antisymmetric matrix. The same procedure can be followed to obtain the Galerkin representation of the second derivative in V_J, defined by the matrix L such that:

$$L_{i_1,i_2} = \left\langle \varphi_{(J,i_1)} \mid \varphi''_{(J,i_2)} \right\rangle$$

and which corresponds to a symmetric filter $r^{(2)}_i$.

In the following we make use of only one orthogonal wavelet basis, namely the C^1 Coiflets with $S = 6$ introduced at the end of (Daubechies, 1993), which we dub "R-Coiflet 6". This wavelet is more regular than the normal Coiflet 6 wavelet, hence the R which stands for "regular". Note that 6 is the minimal filter length for which both $r^{(1)}$ and $r^{(2)}$ are well defined. The scaling function filter coefficients (h_i) are recalled in Table III.2.1. Except otherwise noticed, the results presented below also hold for any other orthogonal wavelet basis. The filters $r^{(1)}$ and $r^{(2)}$ corresponding to the R-Coiflet 6 wavelet basis are provided in Table III.2.2 for completeness.

III.2.2 Description of the PIW scheme

III.2.2.1 Wavelet-based density estimation

Let us briefly recall the wavelet-base density estimation (WBDE) algorithm for estimating the charge density from the position of a finite number N_p of particles. A more detailed account, illustrated by several applications, may be found in (Nguyen van yen et al., 2010a). WBDE, first introduced in (Donoho et al., 1996) belongs to the class of density estimators constructed from orthogonal projections (Silverman, 1986). Its main originality is that the members of the orthogonal family on which the density is projected are chosen in an adaptive manner, depending on the observed particle positions.

In this work, we have focused on the Vlasov-Poisson problem, for which only the spatial distribution of particles matters when estimating the electric field, while the velocity distribution is irrelevant. Therefore, we apply the WBDE algorithm to the spatial distribution only, which is one dimensional. In the units we are working with, this spatial distribution is identical to the charge density ρ. Defining the empirical density associated to the particle positions $(x_n)_{1\leq n\leq N_p}$ by

$$\rho^\delta(x) = \frac{1}{N_p}\sum_{n=1}^{N_p}\delta(x-x_n) \tag{III.2.14}$$

where δ is the Dirac measure, the idea of the WBDE algorithm is to project f^δ on the finite dimensional subspace spanned by all the wavelets at scales coarser than L and part of the wavelets at scales j such that $L \leq j \leq J$, where the scales L and J are defined as follows:

$$L = \left\lfloor \log_2 N_p^{\frac{1}{1+2r_0}} \right\rfloor \tag{III.2.15}$$

$$J = \left\lceil \log_2 \frac{N_p}{\log_2 N_p} \right\rceil \tag{III.2.16}$$

In these formulas, $\lceil\cdot\rceil$ and $\lfloor\cdot\rfloor$ denote respectively the lower and upper integer parts, and r_0 is the order of regularity of the wavelet, which is 1 in our case. The wavelets that are kept for $L \leq j \leq J$ are determined according to the amplitude of the corresponding coefficients in ρ^δ, in the spirit of the classical wavelet denoising algorithms, see (Donoho and Jonhstone, 1994; Donoho et al., 1996; Mallat, 1999). The threshold at scale j is defined by:

$$\Theta_j = K\sqrt{\frac{j}{N_p}}$$

where K is a constant which in principle depends on the regularity of the solution. In the following, we have fixed $K = 1$. The threshold was derived rigorously in (Donoho et al., 1996), and as detailed in (Nguyen van yen et al., 2010a), the $N_p^{-\frac{1}{2}}$ dependence can be justified heuristically from the asymptotic distribution of the wavelet coefficients when estimating a probability density from N_p independent realizations.

To treat the case $l \neq 1$, we can rescale ρ^δ to obtain a normalized distribution on $\frac{\mathbb{R}}{\mathbb{Z}}$:

$$x \rightarrow l^{-1}x \tag{III.2.17}$$

$$\rho^\delta \rightarrow l\rho^\delta \tag{III.2.18}$$

and notice that this change of variables leaves the scaling function and wavelet coefficients (III.2.11-III.2.12) unchanged. Therefore at this stage the algorithm can be applied harmlessly as if $l = 1$.

In summary, the following steps are required to actually implement the estimation:

(i) compute some approximation $\overline{\rho}^S_{(J,i)}$ of the scaling function coefficients of ρ^δ at scale J, by which the projection of ρ^δ on the space V_J can be expressed:

$$\rho^S = \sum_{i=0}^{2^J-1} \overline{\rho}^S_{(J,i)} \varphi_{(J,i)} \tag{III.2.19}$$

(ii) apply the fast wavelet transform algorithm to compute the wavelet coefficients of ρ^S at scales j such that $L \leq j \leq J$:

$$\rho^S = \sum_{i=0}^{2^L-1} \overline{\rho}^S_{(L,i)} \varphi_{(L,i)} + \sum_{j=L}^{J-1} \sum_{i=0}^{2^j-1} \widetilde{\rho}^S_{(j,i)} \psi_{(j,i)} \tag{III.2.20}$$

(iii) apply the nonlinear threshold operator :

$$\rho^W = \sum_{i=0}^{2^L-1} \overline{\rho}^S_{(L,i)} \varphi_{(L,i)} + \sum_{j=L}^{J-1} \sum_{i=0}^{2^j-1} \mathbb{1}_{[0,+\infty[} \left(|\widetilde{\rho}^S_{(j,i)}| - \Theta_j \right) \widetilde{\rho}^S_{(j,i)} \psi_{(j,i)} \tag{III.2.21}$$

where $\mathbb{1}_{[0,+\infty[}$ is the characteristic function of the interval $[0, +\infty[$.

The remaining difficulty in the above procedure is its first step, to which we dedicate the next section entirely. If steps (2-3) are omitted, we obtain a scheme which we call Linear PIW (L-PIW), and that we will use below for comparison purposes. In that case, the value of L has no influence, and we allow J to be chosen as a free parameter.

III.2.2.2 Approximation of empirical scaling function coefficients

To perform step (1) of the above WBDE algorithm, some approximate values $\overline{\rho}^S_\lambda$ of the coefficients $\overline{\rho}^\delta_\lambda$ must be defined. According to (III.2.11) and (III.2.14), their exact values are:

$$\overline{\rho}^\delta_\lambda = \frac{1}{N_p} \sum_{n=1}^{N_p} \varphi_\lambda(x_n), \tag{III.2.22}$$

which is difficult to compute because there is no analytical expression available for φ_λ. In (Nguyen van yen et al., 2010a) we proposed to use the approximation

$$\overline{\rho}^S_{(J,i)} = 2^{-\frac{J}{2}} \rho^H \left(2^{-J} i \right),$$

where ρ_H is a histogram constructed over a grid of size 2^J in each direction. To increase the accuracy in the computation, we adopt here the following alternative. We known that the scaling function φ is supported on the interval $[0, S-1]$, which can be discretized using $2^{J_{LT}}$ regularly spaced points

$$s_k = 2^{-J_{LT}} (S-1) k,$$

at which we can compute approximations of the values of φ

$$\check{\varphi}_k \simeq \varphi(s_k)$$

using for example the iterative algorithm described in (Daubechies and Lagarias, 1991). The $\check{\varphi}_k$ constitute a look-up table of approximations to the $\varphi(s_k)$, which needs to be computed only once for any chosen wavelet basis. Now assume that a particle is located at position x and that we want to approximate

$$\varphi_{(J,i)}(x) = 2^{\frac{j}{2}}\varphi(2^j x - i).$$

This can be nonzero only if $0 \leq 2^j x - i \leq S - 1$, that is, for S values of i. For each of these values of i we find the unique integer k such that

$$s_k \leq 2^j x - i < s_{k+1},$$

we let

$$\alpha = \frac{2^j x - i - x_k}{x_{k+1} - x_k}$$

and we compute a linear interpolation from the values stored in the lookup table,

$$\check{\varphi}_\lambda(x) = (1-\alpha)\check{\varphi}_k + \alpha\check{\varphi}_{k+1}.$$

There remains to sum over all the particles to compute the scaling function coefficient of the charge density which we use in the following:

$$\overline{\rho}^S_\lambda = \frac{1}{N}\sum_{n=1}^{N_p} \check{\varphi}_\lambda(x_n) \tag{III.2.23}$$

For each particle, we need to make S lookup operations in the table, so that the overall cost is proportional to SN_p. For the size of the table, we have used $J_{LT} = 18$. Note that when generalizing this algorithm to d dimensions, the required lookup table will remain one dimensional because the d-dimensional scaling functions are factorized into products of 1D scaling functions. Therefore there will be no increase in memory requirements for this stage of the algorithm.

III.2.2.3 Wavelet-Galerkin Poisson solver

Once the wavelet coefficients of the electron charge density ρ have been obtained by WBDE, the electric potential needs to be computed by solving the Poisson equation (III.2.2). A wavelet preconditioning technique has been previously developed to solve (III.2.2) in 3D using finite differences in the context of a PIC code (Terzić et al., 2007). Here, we prefer to use a full Galerkin discretization of (III.2.2) in the wavelet basis, as introduced in Section III.2.1.2:

$$l^{-2} L\overline{\phi}^S = 1 - \overline{\rho}^W \tag{III.2.24}$$

where $\overline{\phi}^S$ and $\overline{\rho}^W$ stand respectively for the column vectors $(\overline{\phi}^S_\lambda)$ and $(\overline{\rho}^W_\mu)$, ϕ^S stands for the unknown electric potential, and the l^{-2} prefactor comes from the rescaling of the domain (Eq. III.2.17).

A such, this system is ill-conditioned and solving it directly using an iterative method would be inefficient. Fortunately, there is a well known technique to improve the condition number. The main idea is to solve for the wavelet coefficients instead of the scaling function coefficients, and to use a diagonal preconditioner in wavelet space. First (III.2.24) is rewritten

$$l^{-2}\mathcal{F}L\mathcal{F}^{\dagger}\widetilde{\phi}^{S} = \mathcal{F}(1-\overline{\rho}^{W}) \tag{III.2.25}$$

where $\mathcal{F}$ is the matrix form of the wavelet transform operation, and $\cdot^{\dagger}$ stands for transposition. Now (III.2.25) can be efficiently solved using the conjugate gradient method with the diagonal preconditioner defined by:

$$D_{\lambda,\mu} = \delta_{\lambda\mu}2^{j_\lambda}$$

The wavelet coefficients of the potential at the previous time step are used as initial guess to speed-up the convergence. Once $\widetilde{\phi}^{S}$ has been obtained, an inverse wavelet transform is applied to yield $\overline{\phi}^{S}$.

Note that an alternative and non equivalent approach would be to write down a Galerkin discretization of (III.2.2) in the basis made of the wavelets whose coefficients are non-zero in the expansion (III.2.21) for ρ^{W}. We would thus obtain an approximation ϕ^{W} to the potential which has the same support in wavelet space, allowing for an adaptive strategy. In the present approach, ϕ^{S} has a priori a full support in wavelet space. As we will see below this has an important consequence on the self-force.

III.2.2.4 Interpolation and particle push

Like in the standard PIC method, it is essential that the interpolation method be compatible with the charge assignment scheme in order to avoid self-forces (Birdsall and Langdon, 1985). The scaling function coefficients of the electric field can be obtained by Galerkin projection of the gradient operator in the scaling function basis (see Sec. III.2.1.2):

$$\overline{E}^{S}_{\lambda} = \left(\varphi_\lambda, \varphi'_\mu\right)\overline{\phi}^{S}_{\mu} = lG\overline{\phi}^{S}$$

A natural first attempt is to directly interpolate E^{S} at the particle positions to obtain the forces exerted on the particles. For this purpose we could use the approximation:

$$E^{S}(x) = \sum_{\lambda\in\Lambda}\overline{E}^{S}_{\lambda}\check{\varphi}_\lambda(x) \tag{III.2.26}$$

but as we now show this does not guarantee the vanishing of self-forces. Indeed, let us assume for the rest of this section that there is a single particle located at position x. From (III.2.23), we see that, prior to the nonlinear thresholding step, the approximation by the scheme of the charge density is simply given by $\overline{\rho}^{S}_{\lambda} = \check{\varphi}_\lambda(x)$. Hence using the matrix notations defined above we have

$$E^{S}(x) = l\overline{\rho}^{S\dagger}G\overline{\phi}^{S} = l^{-1}\overline{\rho}^{S\dagger}GL^{-1}\overline{\rho}^{W} \tag{III.2.27}$$

Now since G and L are both circulating matrices, they commute, and using the fact that G is antisymmetric and L is symmetric, we obtain that GL^{-1} is also antisymmetric. If we would replace ρ^{W} by ρ^{S} above, we would immediately get that $E^{S}(x) = 0$, but in general

$\rho^W \neq \rho^S$ because of the nonlinear thresholding step. Fortunately a simple remedy exists to enforce the vanishing of self-forces in all cases. Denote by M the matrix of the linear operator defined diagonally in wavelet space by the mask corresponding to the thresholding operation. By definition we have that

$$\overline{\rho}^W = M\overline{\rho}^S$$

so that (III.2.27) is equivalent to

$$E^S(x) = l^{-1}\overline{\rho}^{S\dagger} G L^{-1} M \overline{\rho}^S$$

which leads us to define a new candidate for the electric field reconstruction:

$$\overline{E}^W = lMG\overline{\phi}^S \tag{III.2.28}$$

that is, E^W is obtained from E^S by discarding the same wavelet coefficients that were discarded when going from ρ^S to ρ^W. The self-force corresponding to E^W is obtained as before:

$$E^W(x) = l^{-1}\overline{\rho}^{S\dagger} M G L^{-1} M \overline{\rho}^S$$

but now, using the fact that M is symmetric and GL^{-1} is antisymmetric, we obtain as desired

$$E^W(x) = 0.$$

Once we know how to interpolate the electric field at the particle positions, the advancement in time of (III.2.7) is done using a classical Verlet integrator.

III.2.3 Numerical results

III.2.3.1 Landau damping

To begin with we consider a Landau damping test case, for which the initial condition is given by

$$f_0(x,v) = l^{-1}\left(1 + \varepsilon\cos(kx)\right)\frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}},$$

where $k = 0.5$, $l = 4\pi$, $\varepsilon = 10^{-3}$. The small value of ε implies that the evolution is dominated by the linear transport effect. The reference solution used for comparison was obtained with the SL solver, using 4097 grid points in the x and v directions. The timestep was $\delta t = 0.025$ for the reference solution, and $\delta t = 0.05$ for the others.

The time evolution of the electrostatic energy $\mathcal{V}(t)$ is shown in Fig. III.2.1 (left). Thanks to the high order interpolation properties of the wavelets, the PIW solution matches the decay of the electrostatic energy better than the PIC one. The time evolution of total energy $\mathcal{E}(t)$ for the PIW and reference scheme is shown in Fig. III.2.1 (right). Note that the kinetic energy appearing in $\mathcal{E}(t)$ is approximated by a discrete sum over the particles. Variations remain small for the two schemes, and are especially low for the PIW scheme with 2^{19} particles.

This classical test case serves to validate the PIW method, but since the solution remains very smooth it is not very well suited to provide more insight into the effect of nonlinear thresholding. We therefore move to the more challenging two streams instability.

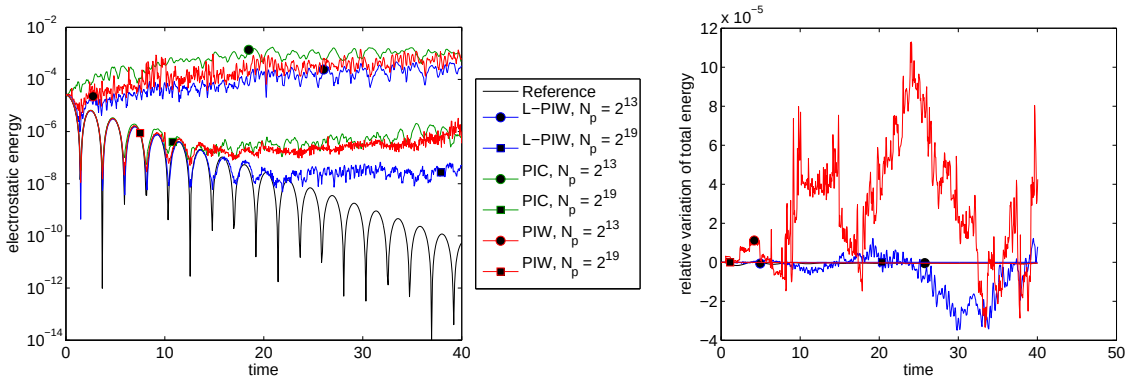


Figure III.2.1: Comparison of PIW, PIC and reference solvers for the Landau damping test case. Left: electrostatic energy as a function of time. Right: kinetic energy as a function of time.

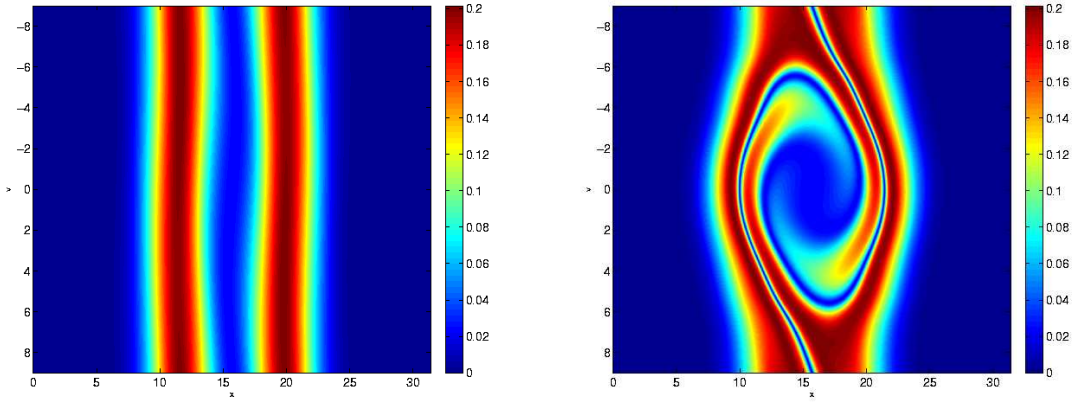


Figure III.2.2: Two streams in stability test case. Particle distribution function obtained using the reference solver at $t = 10$ (left) and $t = 30$ (right).

III.2.3.2 Two-streams instability

The initial condition for this test case is

$$f_0(x,v) = l^{-1}\left(1+\varepsilon\cos(kx)\right)\frac{1}{\sqrt{2\pi}}\left(e^{-\frac{(v-v_0)^2}{2}}+e^{-\frac{(v+v_0)^2}{2}}\right),$$

where $k = 0.2$, $l = 10\pi$, $\varepsilon = 10^{-2}$, and $v_0 = 2.4$. The reference distribution function is shown at $t = 10$ and $t = 30$ in Fig. III.2.2. It was obtained using the SL solver with 4097 grid points in the x and v directions. The timestep was $\delta t = 0.0125$ for the reference solution, and $\delta t = 0.05$ for the others.

For this test case the electrostatic energy (Fig. III.2.3, left) first oscillates for a few periods, roughly up to $t = 10$, and then undergoes an exponential growth which leads to a strongly nonlinear behavior. The PIC and PIW schemes with 2^{19} particles allow for a good approximation of the electrostatic energy up to $t = 30$.

To assess the quality of the approximations of the electric field in the PIC and PIW

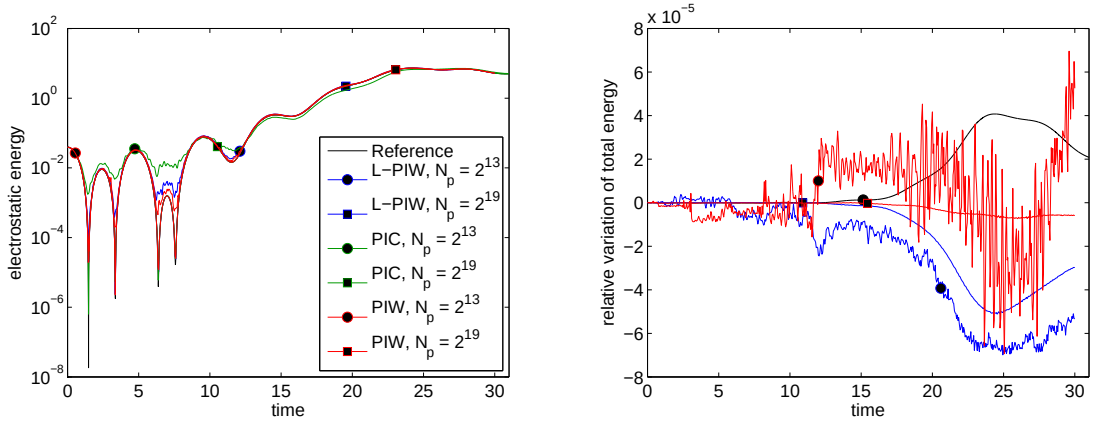


Figure III.2.3: Comparison of PIW, PIC and reference solvers for the two streams in stability test case. Left: electrostatic energy as a function of time (Eq. III.2.5). Right: total energy as a function of time (Eq. III.2.6).

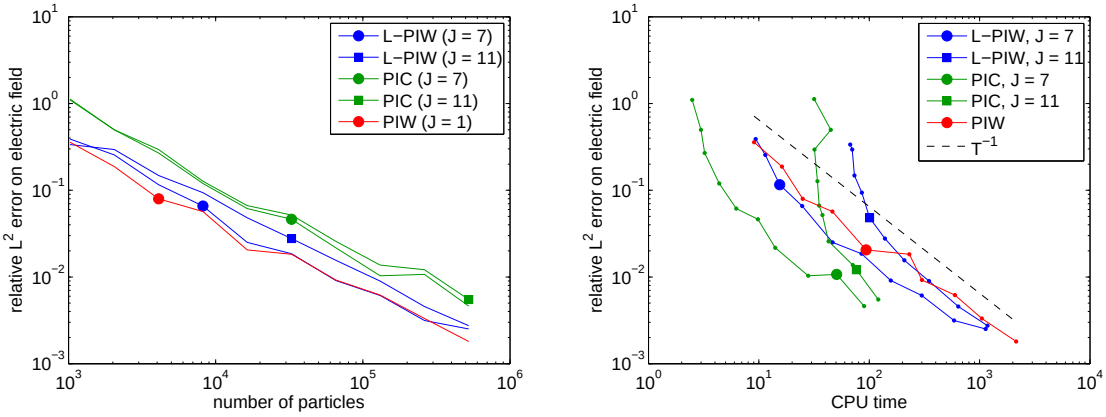


Figure III.2.4: Two streams instability test case. L^2 error on the electric field at $t = 30$, as a function of number of particles (left) and computing time (right).

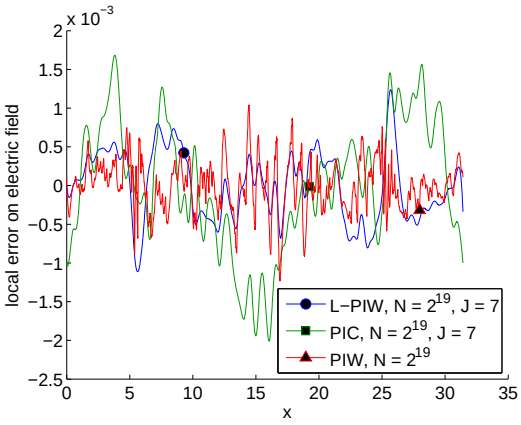


Figure III.2.5: Two streams in stability test case. Error on the electric field as a function of space at $t = 30$.

schemes, we focus on the relative L^2 error:

$$\delta_E(t) = \frac{\|E(x,t) - E_{ref}(x,t)\|}{\|E_{ref}(x,t)\|},$$

where $E_{ref}(x,t)$ is the electric field computed using the reference SL solver. To allow for the computation of $\delta_E(t)$, $E(x,t)$ is interpolated on the grid corresponding to $E_{ref}(x,t)$. In the PIW case, $E(x,t)$ is first reconstructed on a grid from its scaling function coefficients using the methods presented in (Genovese et al., 2008).

$\delta_E(t)$ is shown in Fig. III.2.4 (left) at $t = 30$, as a function of the number of particles. For all considered methods, the scaling lies between N_p^{-1} and $N_p^{-\frac{1}{2}}$. PIW offers an improvement of roughly a factor 3 over PIC, over a wide range of particle numbers. Linear PIW with $J = 7$ is a good candidate but starts to saturate for $N_p = 2^{19}$, probably because the spatial discretization error is reached. In contrast, nonlinear PIW remains uniformly better for any number of particles, thanks to its adaptive properties. In the right panel of Fig. III.2.4, $\delta_E(t)$ is plotted versus the total CPU time (in seconds) needed for the integration of the equations on the time interval $[0, 30]$. The same computer was used in all cases, but the comparison between the absolute values of the CPU time should be made with caution since the languages that were used are different (FORTRAN for the PIC code, and C++ for the PIW code). Anyway, an interesting observation which can be made is that the decrease of the error with computing time for the PIW code is uniformly N_p^{-1} over the whole range of considered parameters. For the PIC and L-PIW methods, in contrast, a saturation is observed when the number of particles is too low for a given spatial resolution.

To get a feeling of how the error we make on the electric field, we plot it as a function of x for $t = 30$ (Fig. III.2.5). It seems that the error can be qualitatively decomposed into the following three components:

(i) a large scale component, which is probably due to a phase shift between the PIW and reference solutions,

(ii) localized peaks, where some sharp features in the electric field are not properly resolved,

(iii) noise spread out over the whole domain.

Our goals with nonlinear thresholding was to preserve the sharp features that develop due to nonlinear interactions.

III.2.4 Discussion

The PIW approach was proposed as an alternative to PIC for plasma simulation using particles, and successfully implemented to solve the simple one dimensional Vlasov-Poisson toy model. The main conceptual difference between PIC and PIW is that the empirical distribution of particles is convolved with a kernel for the former, whereas it is projected onto a finite dimensional linear space for the latter. Thanks to the use of wavelets to generate the projection space, the denoising step which can be included in PIW and which we have studied is genuinely different from those that are developed for PIC schemes. Indeed, the wavelet representation can adapt locally to the regularity of the density, whereas linear approaches such as Krook operators (McMillan et al., 2008), Fourier filtering (Jolliet et al., 2007) or coarse graining (Chen and Parker, 2007), cannot.

The adaptive denoising strategy is especially useful when the density develops sharp features, which is known to occur generically as soon as nonlinear effects become important. To better understand this effect, we have compared a linear Landau damping test case, where the density remains homogeneously smooth, to a two stream instability test case. We have shown that for the first, PIW behaves as a higher order PIC approach, offering a constant improvement in terms of L^2 error, but no improvement in rate of convergence. In contrast, for the second test case, fine resolution was required to resolve sharp features developing in the electric field.as a result of nonlinear effects. The PIW scheme is able to automatically increase the resolution locally when the number of particles is sufficient to guarantee that the estimation of the density will not get corrupted by noise. This adaptivity guarantees that the error decays like the inverse square root of computing time over a wide range of particle numbers. The topic of discretization error control using wavelets has been studied elsewhere, see e.g. (Schneider and Vasilyev, 2010) for a review in the context of computational fluid dynamics.

For the simple 1D case considered here, the improvement observed was roughly a factor 3 of precision. The increase in computational cost of the PIW scheme with respect to the PIC scheme is proportional to the length of the wavelet filter, which we have fixed to 6 in this study. Extending the full PIW scheme to more realistic 2D or 3D geometries is a much more ambitious challenge. A first undertaking, complementary to the present study, could be to interface wavelet denoising using WBDE in an existing production code in order to further support the possible benefits of the approach.

Part IV

Regularization of inviscid equations

The answer, my friend, is blowin' in the wind.
The answer is blowin' in the wind.

Bob Dylan

Une première salve, un long froissement d'étoffe, soutenu, cinglant, puis le redoux des carresses pleines, lorsque le souffle se fait enveloppant et rond, jusqu'à l'empâtement. Ça se hausse, seconde salve, dans la véhémence et presque claquée. Un calme suit, qui aspire le corps vers l'avant. Puis troisième salve, forte puis decrescendo, jusqu'à ne plus onduler qu'avec les nuances d'une brise.

Alain Damasio
La horde du contrevent

Summary

IV.1 1D Burgers equation

This section is adapted from a published paper, see (Nguyen van yen et al., 2008).

IV.1.1 Introduction

The fully developed turbulent regime is described by solutions of the Navier–Stokes equations for two or three-dimensional incompressible fluids, in the limit where the kinematic viscosity becomes very small. By analogy, Burgulence is described by the solutions of Burgers equations for a one–dimensional fluid in the same limit, as first proposed by Burgers (1939, 1948) and advocated by von Neumann (1963). This toy model for turbulence has been extensively used since then (Kida, 1979; She et al., 1992; Vergassola et al., 1994; Avellaneda and Weinan, 1995; Gurbatov et al., 1997); Frisch and Bec (2001) have proposed to name it: *Burgulence*.

We consider the one-dimensional Burgers equation in a periodic domain of support $x \in [-1, 1]$, which describes the space–time evolution of the velocity $u(x, t)$ of a one–dimensional fluid flow :

$$\partial_t u + \frac{1}{2}\partial_x u^2 \;=\; \nu \partial_{xx} u \;, \tag{IV.1.1}$$

where ν denotes the kinematic viscosity. The solutions of (IV.1.1) can be computed analytically using the Cole-Hopf transformation (Hopf, 1950; Cole, 1954; Burgers, 1954). When $\nu \to 0$ the solutions of the viscous Burgers equation approach weak solutions of the inviscid problem. The uniqueness of these solutions stems from the condition that shocks have negative jumps, which guarantees energy dissipation. For Burgers equation, this condition is equivalent to an entropy condition (Germain and Bader, 1953; Lax, 1954; Oleinik, 1957; Kruzhkov, 1970).

The wavelet representation has been proposed for studying turbulence (Farge et al., 1992a), since it preserves both the spatial and spectral structures of the flow by realizing an optimal compromise in regard of the uncertainty principle. We have found that projecting the vorticity field onto a wavelet basis, and retaining only the strongest coefficients, extracts the coherent structures out of fully developed turbulent flows (Farge et al., 1999, 2001). We have then proposed a computational method for solving the Navier-Stokes equations in wavelet space (Farge et al., 1999). We have shown that extracting the coherent contribution at each time step preserves the nonlinear dynamics, whatever its scale of activity, while discarding the incoherent contribution corresponding to turbulent dissipation (Schneider and Farge, 2005). This is the principle of the CVS (Coherent Vortex Simulation) method we have proposed (Farge et al., 1999; Farge and Schneider, 2001).

The aim of the present paper is to apply the CVS filter to the inviscid Burgers equation and check if this is equivalent to solving the viscous Burgers equation. The outline is the following. First we recall the principle of CVS filtering and its extension using complex-valued translation-invariant wavelets. The numerical scheme is described briefly and the main part presents results of several numerical experiments, considering either deterministic or random initial conditions. Finally, we draw conclusions and propose some perspectives.

IV.1.2 Numerical method

The Burgers equation (IV.1.1) is discretized on N grid points using a Fourier spectral collocation method :

$$\frac{\partial U}{\partial t} + \frac{1}{3} D_N(U^2) + \frac{1}{3} U \cdot D_N(U) - \nu D_N^2(U) = 0 \ , \tag{IV.1.2}$$

where U approximates $(u(x_0,t), u(x_1,t), ..., u(x_{N-1},t))$, D_N stands for the Fourier collocation differentiation and $\cdot$ is the pointwise product of two vectors. The discretization of the nonlinear term in (IV.1.2) is chosen in order to conserve the kinetic energy $E = \frac{1}{2}\int_{-1}^{1} u^2(x,t)dx$ when $\nu = 0$ (Canuto et al., 1988). For time integration a fourth-order Runge-Kutta scheme is used.

At each time step, we will filter the solution using the CVS method which we now recall briefly. Given orthogonal wavelets (ψ_{ji}) and the associated scaling function at the largest scale φ, the velocity can be expanded into :

$$u(x) = \langle u \mid \varphi \rangle \varphi(x) + \sum_{j=0}^{J-1} \sum_{i=1}^{2^j} \langle u \mid \psi_{ji} \rangle \psi_{ji}(x) \ , \tag{IV.1.3}$$

where j is the scale index, i is the position index and the inner product is $\langle a \mid b \rangle = \int_{-1}^{1} a(x) \cdot b^*(x) dx$ with b^* denoting the complex conjugate of b. Since location in orthogonal wavelet space is sampled on a dyadic grid, this representation breaks the local translation invariance of (IV.1.1) which may impair the stability of the numerical scheme. Therefore we prefer using, instead of real-valued wavelets, complex valued wavelets (Kingsbury, 2001) which very closely preserve translation invariance. In this case, (IV.1.3) still holds as long as we replace the right hand side by its real part.

The CVS filter then consists in discarding the wavelet coefficients whose modulus is below a threshold T. In addition, wavelets coefficients at the finest scale are systematically filtered out. The resulting velocity u_T is a nonlinear approximation of u.

Because the velocity field decays in time, the threshold has to be estimated at each time step in a self-consistent way. To do this, we follow the iterative method introduced by Azzalini et al. (2004), which consists in imposing the ratio between the standard deviation of the discarded wavelet coefficients and the threshold itself:

$$T^2 = \frac{5}{N_T} \sum_{j=0}^{J-1} \sum_{i=1}^{2^j} |\widetilde{u}_{ji}|^2 H(T - |\widetilde{u}_{ji}|) \ , \tag{IV.1.4}$$

where H is the Heaviside step function and N_T is the number of wavelet coefficient below the threshold. The solution of (IV.1.4) is determined numerically using a fixed point iterative procedure (Azzalini et al., 2004), initialized with $T_0 = 5E/N$, where E is the total energy.

IV.1.3 Deterministic initial condition

We consider Burgers equation (IV.1.1) with the deterministic initial condition $u(t=0,x) = -\sin(\pi x)$. We begin by comparing three computations: a Galerkin-truncated inviscid

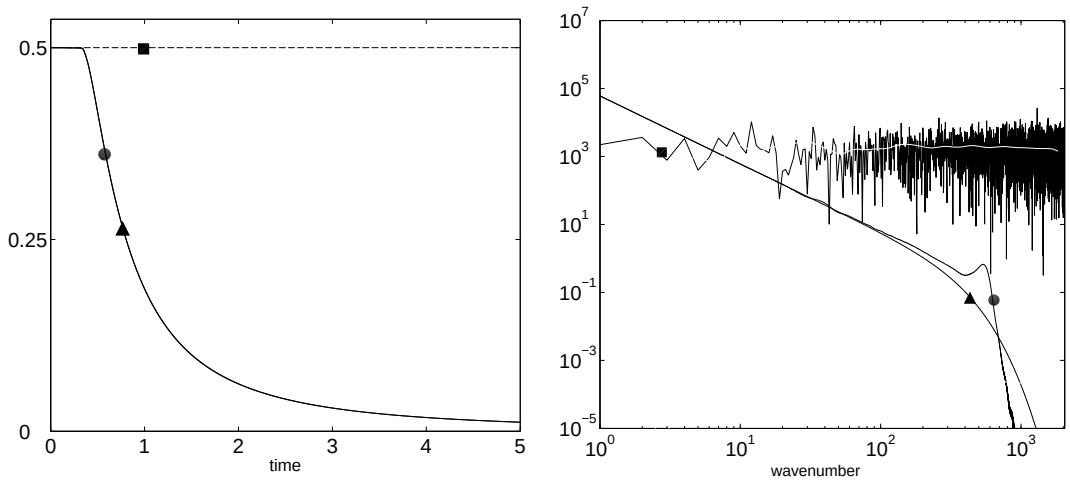


Figure IV.1.1: Deterministic initial conditions. Left: Time evolution of energy. Right: Energy spectrum at $t = 5$. We compare the Galerkin-truncated inviscid (square), viscous (triangle) and CVS-filtered inviscid (circle) cases. We observe that for the inviscid case (right) the wavelet spectrum (white line) better exhibits the energy equipartition than the Fourier spectrum (black line).

case ($\nu = 0$), a viscous case ($\nu = 10^{-4}$), and an inviscid case with the CVS filter applied at each time step. The solutions are computed up to time $t = 5$, using $N = 4096$ grid points.

By computing in the Galerkin-truncated inviscid case ($\nu = 0$), we check that our numerical scheme conserves energy (Fig. IV.1.1, left) as theoretically predicted. We observe that the final solution at $t = 5$ exhibits energy equipartition (Fig. IV.1.1, right) with a Gaussian velocity PDF, as expected. Notice that the the white line in Fig. IV.1.1 (right) corresponds to the wavelet energy spectrum, *i.e.*, the squared modulus of the wavelet coefficients computed with a complex-valued Morlet wavelet. It better exhibits the k^0 scaling, characteristic of the energy equipartition, than the highly oscillatory Fourier energy spectrum (black line). This illustrates the fact that the wavelet energy spectrum is more stable than the Fourier energy spectrum when we analyze only one realization of a stochastic process (Farge et al., 1992a).

For the viscous and CVS-filtered inviscid cases, the energy remains basically constant until the shock forms at $t = 1/\pi$, but then decays with a t^{-2} law. In Fig. IV.1.1 (right) the energy spectra of the viscous and CVS-filtered inviscid cases exhibit a power law behaviour with slope -2.

Fig. IV.1.2 shows the velocity at three time instants for the viscous and CVS-filtered inviscid cases. The CVS-filtered inviscid solution yields the same dynamics as the viscous one except for the small overshoot we observe at $x = 0$ after the shock has formed. This Gibbs phenomenon is stronger but less oscillatory for the CVS-filtered inviscid case than for the viscous case (see the insets in Fig. IV.1.2).

The time evolution of the percentage of retained wavelet coefficients is presented in Fig. IV.1.3 (left). It shows that with only relatively few coefficients (about $7\%N$) we are able to track the nonlinear dynamics of the flow and this number remains almost constant after the shock formation. At $t = 5$, the retained wavelet coefficients are located around $x = 0$,

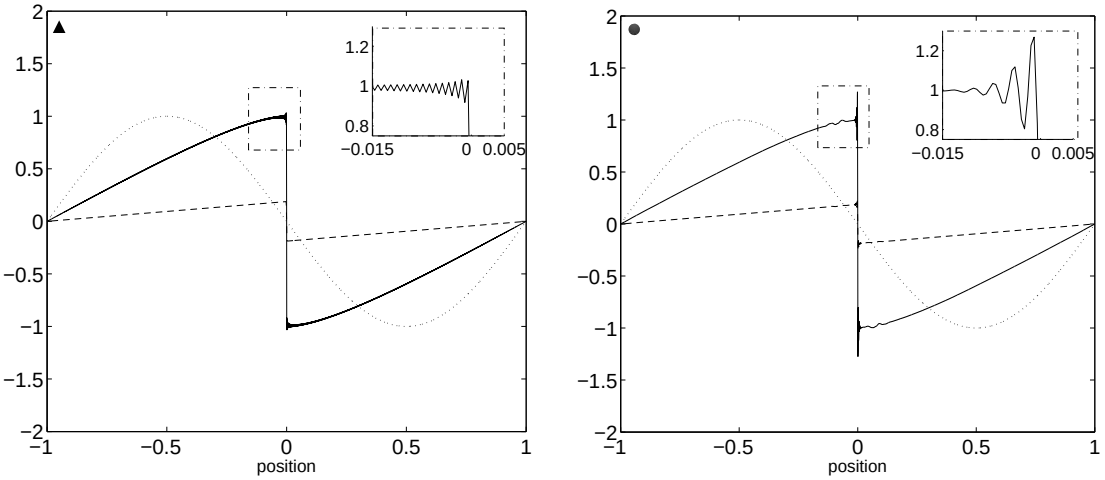


Figure IV.1.2: Deterministic initial conditions. Snapshots of velocity for the viscous (left) and the CVS-Filtered inviscid (right) cases at $t = 0$ (dotted line), $t = 0.5$ (solid line) and $t = 5$ (dashed line). The insets show the tip of the shock at $t = 0.5$.

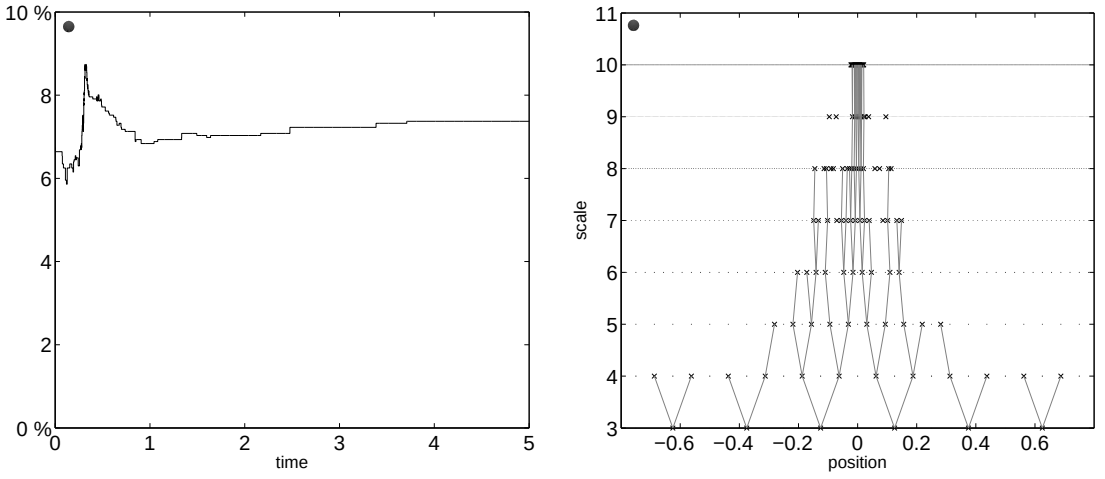


Figure IV.1.3: Deterministic initial conditions. Left: Time evolution of the percentage of wavelet coefficients retained after filtering. Right: Dyadic tree of the wavelet coefficients which are retained after filtering at $t = 5$. The crosses indicate the $7\%N$ retained wavelet coefficients, while the small dots correspond to the $93\%N$ discarded wavelet coefficients. The scale varies from coarse to fine, up the vertical axis.

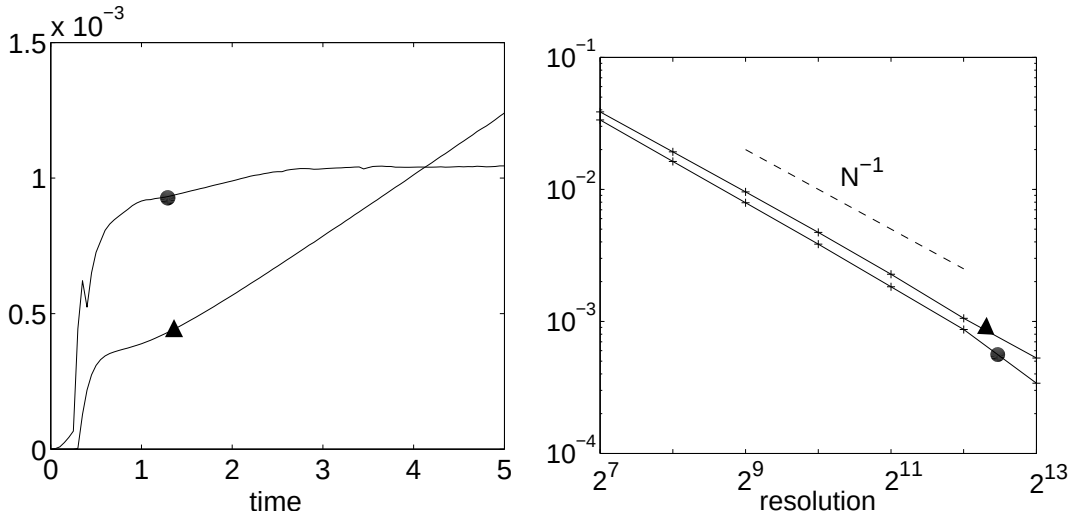


Figure IV.1.4: Deterministic initial conditions. Left: Time evolution of the relative mean squared error ϵ_N at $N = 4096$. Right: Relative mean squared error ϵ_N at $t = 5$ for different numerical resolutions, $N = 128$ to $N = 8192$. We compare the viscous (triangle) and CVS-filtered inviscid (circle) cases.

the position of the shock, and span all scales there, as illustrated in Fig. IV.1.3 (right).

We now show that, when N increases, the filtered solutions converge towards the entropy solution $u_{\rm ref}$ which solves the Burgers equation in the inviscid limit. For comparison we also consider the viscous solutions with viscosity depending on N ($\nu = 0.4096N^{-1}$) which are known to converge to $u_{\rm ref}$ everywhere, except at $x = 0$. The entropy solution $u_{\rm ref}$ is directly calculated using the method of characteristics.

First, we consider a global error estimate, the relative mean square error, defined as:

$$\epsilon_N(t) = \frac{\|u - u_{\rm ref}\|_2^2}{\|u_{\rm ref}\|_2^2} \ . \tag{IV.1.5}$$

On Fig. IV.1.4 (left) we plot $\epsilon_N(t)$ for $N = 4096$. The error for the CVS-filtered inviscid case is larger but saturates after $t \simeq 2$. In contrast, the error for the viscous case keeps increasing because the finite viscosity smooths the shock away. Considering now $t = 5$ and varying N, we find that for both the viscous and CVS-filtered inviscid cases ϵ_N decreases as N^{-1} (Fig. IV.1.4, right).

We now study the behaviour of the oscillations in the neighborhood of the shock when the resolution N is increased. The total variation of a function f on $[-1, 1]$ is defined by:

$$\|f\|_{TV} = \int_{-1}^{1} |\partial_x f| dx \ . \tag{IV.1.6}$$

To detect the presence of spurious oscillations, we compute the relative error on the total variation:

$$\epsilon'_N(t) = \frac{\|u(x,t)\|_{TV} - \|u_{\rm ref}(x,t)\|_{TV}}{\|u_{\rm ref}(x,t)\|_{TV}} \ , \tag{IV.1.7}$$

which is plotted as a function of N for $t = 5$ on Fig. IV.1.5 (left). For the viscous case, ϵ'_N is negative and converges towards zero when N increases. For the CVS-filtered inviscid

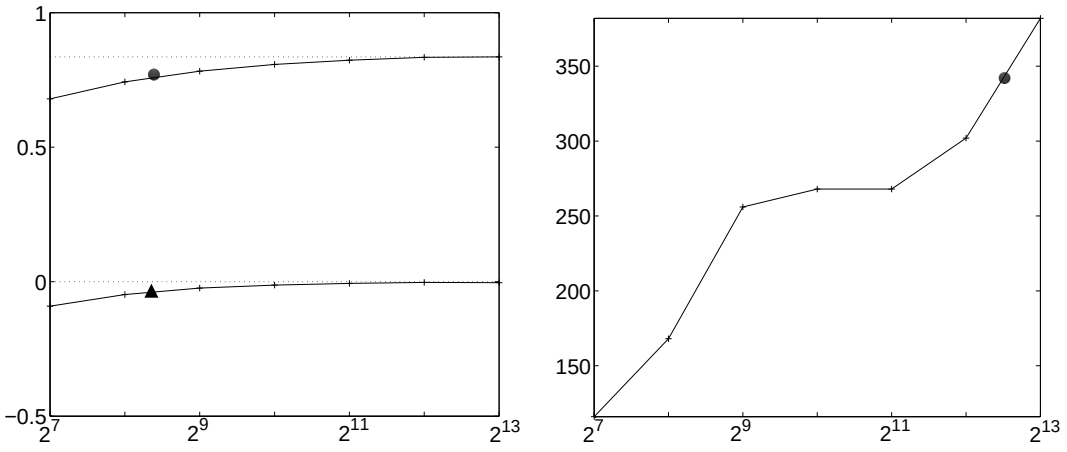


Figure IV.1.5: Deterministic initial conditions. Error on the relative total variation ϵ'_N (left), and number of retained wavelet coefficients (right), as functions of N at $t = 5$, for the viscous (triangle) and CVS-filtered inviscid (circle) solutions.

case, ϵ'_N tends to a finite positive value close to 0.84. The overshoot that could be seen on Fig. IV.1.2 persists but becomes more and more localized around the singularity when N increases, thus ensuring mean square convergence.

Let us end this section by a short discussion on the evolution of the compression rate when N increases. Fig. IV.1.5 (right) shows that the number of retained wavelet coefficients increases roughly logarithmically as a function of N. As a consequence, notice that for the filtered solution, the relative mean square error $\epsilon_N(t)$, if it is considered as a function of the number of retained coefficients only, converges to zero exponentially fast. However, to experience this promising rate of convergence in practice, we should compute the evolution of u using only the wavelet coefficients whose modulus remains above the threshold.

IV.1.4 Random initial condition

In the previous section we have demonstrated that the CVS-filtered inviscid Burgers equation exhibits an evolution similar to that of the viscous Burgers equation. We now would like to check if this is still verified in the context of *Burgulence* for both white noise (Avellaneda and Weinan, 1995) and Brownian motion (She et al., 1992).

White-noise initial condition We take as initial velocity one realization of a Gaussian white noise computed at resolution $N = 4096$, which corresponds to a random non intermittent initial condition. Since the CVS filter removes the non intermittent noisy contributions, if applied to a Gaussian white noise the latter would be completely filtered out. Therefore we first integrate the viscous equation with $\nu = 2 \cdot 10^{-5}$ without filtering, and wait until the flow intermittency has sufficiently developed before applying the filter. To check the flow intermittency we monitor the flatness of velocity gradient until it reaches the value 20, which happens at $t = 0.017$ for the realization described here. Then, we reset $t = 0$ and integrate up to $t = 5$, both the viscous equation with $\nu = 2 \cdot 10^{-5}$, and the CVS-filtered inviscid equation.

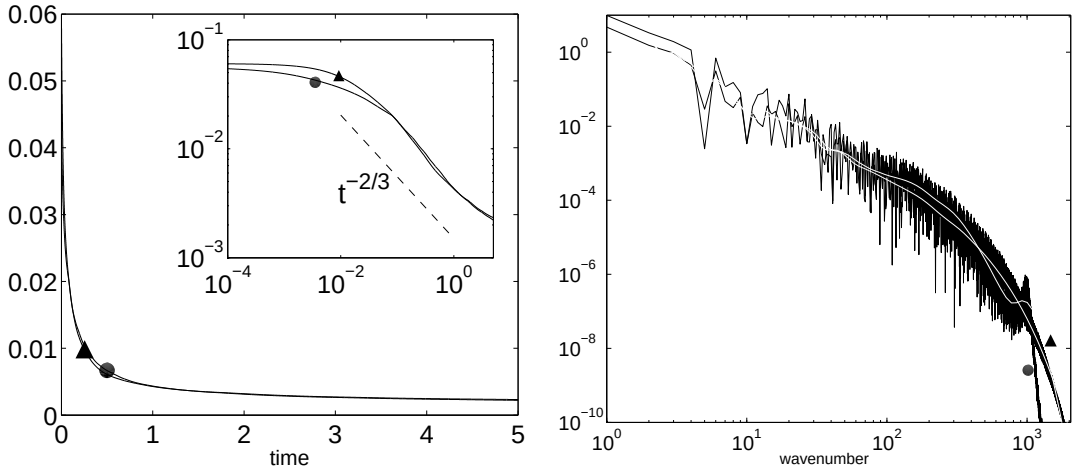


Figure IV.1.6: White noise initial conditions. Left: time evolution of energy. The inset shows the $t^{-2/3}$ decay in log-log coordinates. Right: energy spectrum at $t = 5$. We compare the viscous (triangle) and CVS-filtered inviscid (circle) simulations. We observe that the wavelet spectrum (white lines) better exhibits the k^{-2} scaling of energy than the Fourier spectrum (black lines).

In Fig. IV.1.6 (left) we show that the energy, for both the CVS-filtered inviscid solution and the viscous solution, decays with a $t^{-2/3}$ law, as found by Burgers (Burgers, 1954; She et al., 1992). In Fig. IV.1.6 (right) we observe at $t = 5$ that both energy spectra present the same k^{-2} scaling. Notice that the two white lines in Fig. IV.1.6 (right) correspond to the wavelet energy spectrum, which better exhibit the k^{-2} scaling of the energy than the highly oscillatory Fourier energy spectrum (black lines).

Finally, we show on Fig. IV.1.7 that the viscous and CVS-filtered inviscid solutions are almost identical in physical space, presenting a typical sawtooth profile as first noticed by Burgers (1954).

Brownian motion initial condition We use the same resolution $N = 4096$ as above, only the initial condition changes. Since we have chosen periodic boundary conditions we approximate the Brownian motion by the Fourier series:

$$u(x,0) = \sum_k u_k e^{ikx} \tag{IV.1.8}$$

where $k = -\frac{N}{2}+1, -\frac{N}{2}, ..., \frac{N}{2}-1$. We set $u_0 = 0$ and, for $k \neq 0$, we take for u_k a complex Gaussian random variable with standard deviation $1/|k|$.

The solution for the viscous case is computed with $\nu = 1.2 \cdot 10^{-4}$. For the CVS-filtered inviscid case, as we did for the white noise initial condition, we do not filter before enough intermittency has developed. We thus integrate the viscous equation with $\nu = 1.2 \cdot 10^{-4}$ for 0.05 time units and then switch viscosity off. This procedure provides the initial velocity which by construction is the same for both methods (Fig. IV.1.8).

The energy decay matches well between the CVS-filtered inviscid and the viscous solutions (Fig. IV.1.9, left). A k^{-2} power spectrum is also obtained for both at $t = 5$ (Fig. IV.1.9, right).

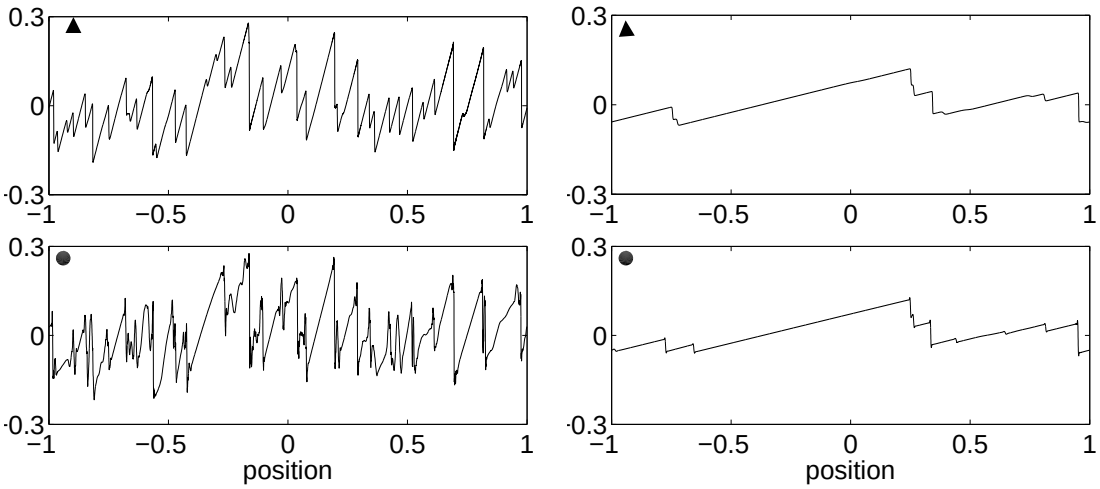


Figure IV.1.7: White noise initial conditions. Snapshots of velocity at $t = 0.3$ (left) and $t = 5$ (right). Top: viscous equation with $\nu = 2 \cdot 10^{-5}$ Bottom: CVS-filtered inviscid equation.

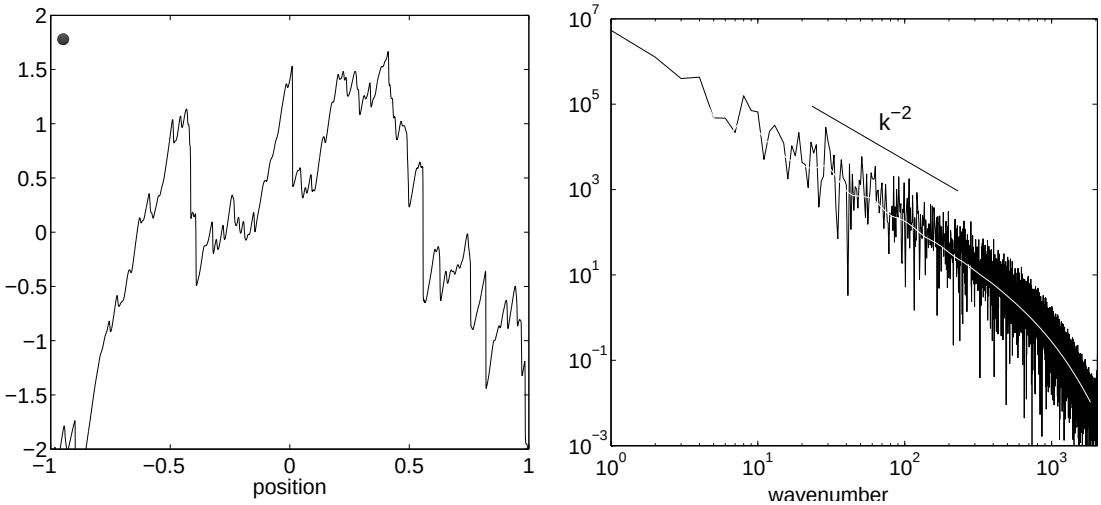


Figure IV.1.8: Brownian initial condition. Velocity at $t = 0$ (left), its Fourier energy spectrum (right, black line) and its wavelet energy spectrum (right, white line).

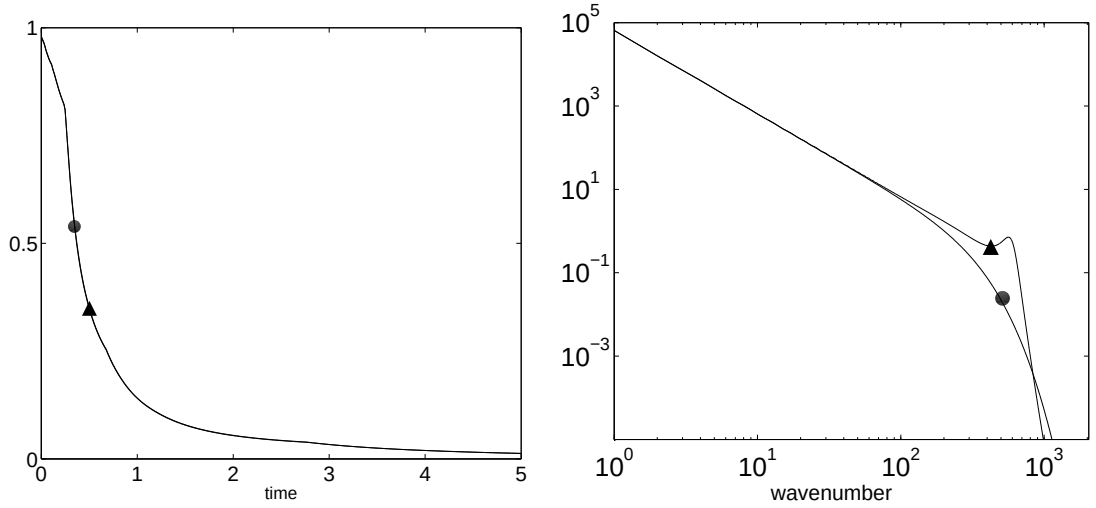


Figure IV.1.9: Brownian initial condition. Left: Time evolution of energy. Right: Energy spectrum at $t = 5$. We compare the viscous (triangle) and CVS-filtered inviscid (circle) cases.

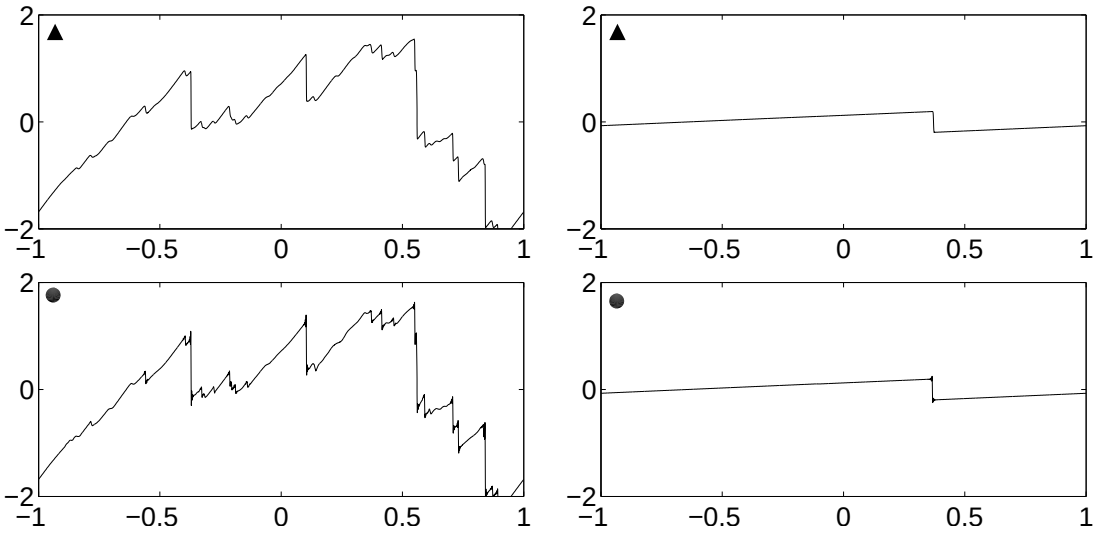


Figure IV.1.10: Brownian initial conditions. Snapshots of velocity at $t = 0.1$ (left) and $t = 5$ (right). Top: viscous equation with $\nu = 1.2 \cdot 10^{-4}$. Bottom: CVS-filtered inviscid equation.

At $t = 0.1$ numerous small shocks are present on the viscous solution (Fig. IV.1.10, top left). All of them are correctly reproduced by the CVS-filtered inviscid solution (Fig. IV.1.10, bottom left).

At $t = 5$ the single remaining shock that is still resolved in the viscous solution (Fig. IV.1.10, top right) is correctly reproduced on the CVS-filtered inviscid solution (Fig. IV.1.10, bottom right).

IV.1.5 Conclusion

We have shown that CVS-filtering at each time step the solution of the inviscid Burgers equation gives the same evolution as the viscous Burgers equation, for both deterministic and random initial conditions. As our contribution to Euler equation's 250th anniversary and Euler's 300th birthday, we conjecture that CVS-filtering of the Euler equation may be equivalent to solving the Navier–Stokes equation in the fully-developed turbulent regime, *i.e.*, when dissipation has become independent of viscosity. We predict that the retained wavelet coefficients would preserve Euler's nonlinear dynamics, while discarding the weaker wavelet coefficients would model turbulent dissipation and give Navier–Stokes solutions. Since in the fully-developed turbulent regime turbulent dissipation strongly dominates molecular dissipation, there is no reason to model turbulent dissipation by a Laplacian operator anymore. Indeed, turbulent dissipation is a property of the flow, while molecular dissipation is a property of the fluid and may no more play a role when turbulence is fully-developed. We think that in this regime the CVS filter could be a better way to model dissipation, replacing global by local smoothing while preserving nonlinear interactions. In this paper we have chosen the simplest toy model to test this conjecture, although Burgers' equation, in contrast to Euler's equation, is neither chaotic nor producive of randomness. Therefore we conjecture that the CVS-filter would work better for Euler/Navier–Stokes than for Burgers equation since CVS is based on denoising, which is justified when there is chaos and randomness.

Acknowledgments: We thank Uriel Frisch, Margarete Domingues, Claude Bardos, François Dubois and two anonymous referees for their useful comments. We acknowledge financial support from the ANR under contract M2TFP (Méthodes Multi-échelles pour la Turbulence Fluide et Plasma) and from the Association CEA-Euratom under contract V.3258.006. NK thanks the Université de Provence for supporting his stay in Marseille. MF is grateful to the fellows of Trinity College, Cambridge (U.K.), in particular Keith Moffatt, for their kind hospitality while revising this paper.

IV.2 Incompressible 2D Euler equations

This section is adapted from a published paper, see (Nguyen van yen et al., 2010a).

IV.2.1 Introduction

When they are dominated by nonlinear effects due to inertia, flows in Newtonian incompressible fluids enter the turbulent regime, in which their motions are very disordered and involve a wide range of scales. Nevertheless, they tend to possess certain robust phenomenological properties, e.g. the average energy dissipation rate, that have lead to their study in a specific framework known as the statistical theory of fully developed homegeneous isotropic turbulence (Frisch, 1995). This theory is not directly derived from the basic equations of fluid mechanics but requires additional statistical hypotheses that are difficult to check experimentally. Moreover, it doesn't take into account self-organization of the flow into coherent structures. On the other hand, current numerical methods tend to reproduce observed properties of turbulent flows. They offer rich opportunities for proposing new theoretical approaches and confronting them quantitatively.

Despite the chaotic character of turbulent flows, it is commonly accepted that, if the density is constant, their velocity field $\mathbf{u}(\mathbf{x},t)$ is well described by solutions of the Navier-Stokes equations:

$$\begin{cases} \partial_t \mathbf{u}_\nu + (\mathbf{u}_\nu \cdot \nabla)\mathbf{u}_\nu = -\nabla p_\nu + \nu \nabla^2 \mathbf{u}_\nu, \quad \mathbf{x} \in \mathbb{T}^d, \quad t \in]0, +\infty[\\ \nabla \cdot \mathbf{u}_\nu = 0 \\ \mathbf{u}_\nu(\cdot, 0) = \mathbf{u}_0 \end{cases} \tag{IV.2.1}$$

where ν is the kinematic viscosity of the fluid, t is time, and $\mathbf{u}_0$ is an initial flow on the d-dimensional torus $\mathbb{T}^d = (\frac{\mathbb{R}}{2\pi\mathbb{Z}})^d$. Here, the pressure p_ν is determined by $\mathbf{u}_\nu$ through the divergence-free condition and cannot be chosen independently. The equations have been written in dimensionless units, so that the Reynolds number can be simply defined by $\mathrm{Re} = \nu^{-1}$. Properly describing the behavior of solutions to (IV.2.1) when $\mathrm{Re} \gg 1$ remains a central problem for the understanding fully developed turbulence. A key mathematical difficulty is that the limit $\nu \to 0$, i.e. $\mathrm{Re} \to \infty$, is singular as the order of the equation changes. Indeed, many wind tunnel experiments (Batchelor, 1953; Sreenivasan, 1984), as well as numerical experiments (Kaneda et al., 2003), suggest that the limit $\mathbf{u}$ of $\mathbf{u}_\nu$ when $\mathrm{Re} \to \infty$ does not satisfy Euler's equations (IV.2.2),

$$\partial_t \mathbf{u} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\nabla p \tag{IV.2.2}$$

or at least not in the strong sense. This idea was elaborated upon by many authors following the seminal work of Onsager, about which a nice historical review has been recently published (Eyink and Sreenivasan, 2006).

When trying to approximate numerically turbulent flows, the traditional approach is to solve the Navier-Stokes equations and make the viscosity as small as is allowed by numerical discretization, which is ultimately limited by the available computational power (see e.g. (Kaneda et al., 2003) for state of the art results obtained using 4096^3 degrees of freedom). But if the dynamics in the inviscid limit, or at least some of its important features, become strictly independent of Re, it is much more desirable to solve the Euler equations instead of

the Navier-Stokes equations. A problem of parabolic type (eq. IV.2.1) is thereby replaced by a problem of hyperbolic type (eq. IV.2.2), as the higher order derivative disappears. In some cases, finite time singularities may then occur, and uniqueness of the solution is lost (Lax, 1973). This happens for example for the Burgers equation, or for the compressible Euler equation, and is conjectured to happen for the 3D incompressible Euler equations. Admittedly, the 2D Euler equations, that we are going to consider below, do not give rise to finite time singularities. Their solution remains smooth for all time, since we are considering only domains without boundaries, and smooth initial data (Wolibner, 1933; Kato, 1967). However, the vorticity gradients grow extremely fast in time. This phenomenon, which has been described as "slow collapse" (Yudovich, 2000), is almost as bad as finite time singularity as far as numerical simulation is concerned.

Due to the roughness of the exact solutions to these problems, numerical methods for solving hyperbolic partial differential equations must be accompanied by proper regularization mechanisms. If the solution is not unique, the responsibility even falls upon the numerical scheme to select the weak solution that it approximates. Physical considerations must hence be taken into account when designing the scheme so that it yields the proper solution. Existing methods include upwind characteristic (Osher and Solomon, 1982) or total variation diminishing (Harten, 1983) schemes, shock limiters (Sweby, 1984), spectral vanishing viscosity (Tadmor, 1989; Gottlieb and Hesthaven, 2001), hyperdissipation (Basdevant et al., 1981), and more recently inviscid regularization (Holm et al., 1998; Bardos et al., 2008; Khouider and Titi, 2008). The latter stands aside from the rest since it was introduced mainly as a tool for mathematical proofs. One may also refer to (Pulliam, 1986) for a comparison between several schemes.

In this paper, we apply a regularization method based on wavelet filtering to solve the 2D Euler equations written in vorticity-velocity formulation:

$$\begin{cases} \partial_t \omega + \mathbf{u} \cdot \nabla \omega = 0, \quad t \in]0, +\infty[, \quad \mathbf{x} \in \mathbb{T}^2 \\ \omega = \nabla \times \mathbf{u} \\ \nabla \cdot \mathbf{u} = 0 \\ \omega(\cdot, 0) = \omega_0 \end{cases} \tag{IV.2.3}$$

where ω denotes the vorticity component perpendicular to the plane. We compare the results with those obtained when the first equation is replaced by

$$\partial_t \omega_{\nu,\alpha} + (\mathbf{u}_{\nu,\alpha} \cdot \nabla)\omega_{\nu,\alpha} = -\nu_\alpha (-\Delta)^\alpha \omega_{\nu,\alpha}, \quad \alpha \geq 1 \tag{IV.2.4}$$

which corresponds to the 2D hyperdissipative Navier-Stokes equations with dissipativity α if $\alpha > 1$, or to the classical 2D Navier-Stokes equations if $\alpha = 1$, in which case we drop the index α.

Note that the physical relevance of the inviscid limit in 2D is questionable, since virtually all known flows become three dimensional at high enough Reynolds number. Nevertheless, the 2D case is chosen as a toy model to allow us access to high Reynolds number flows at an affordable computational cost. Another unphysical aspect of our study resides in the boundary conditions. Physically realistic flows are in contact with solid walls which need to be taken into account by introducing no-slip boundary conditions in the modeling. These boundary conditions cannot be satisfied in general by solutions of the Euler equations. Taking them into account properly in the inviscid limit hence introduces new difficulties that we reserve for future work. In the meantime, we use periodic boundary conditions.

The main goal of this paper is to extend to the 2D Euler equations the study that we did in a previous work (Nguyen van yen et al., 2008) concerning regularization of the 1D Burgers equation:

$$\begin{cases} \partial_t u + u\partial_x u = 0, \quad t \in]0, +\infty[, \quad x \in [-1, 1], \\ u(\cdot, 0) = u_0. \end{cases} \tag{IV.2.5}$$

In the first part, we explain our numerical method, which consists in classical discretization schemes and the original part of our work, namely, the wavelet filters. In the second part, we present some new numerical results obtained for the 1D Burgers equation, and we then study in detail the regularization method in the case of the 2D Euler equations. Finally, we draw some conclusions relative to those two sets of results and outline perspectives for extending the study to 3D flows.

IV.2.2 Numerical method

IV.2.2.1 Discretization

In (Nguyen van yen et al., 2008), the 1D Burgers equation was discretized by means of a Fourier collocation method, with pseudo-spectral evaluation of the nonlinear term (Canuto et al., 1988). In such schemes, the velocity is followed via its Fourier coefficients, which satisfy a system of nonlinear ordinary differential equations. To efficiently compute the convolution product that appears due to the quadratic nonlinear term, the solution is reconstructed on an evenly spaced grid, hence the term "pseudo-spectral". Conservation of energy was enforced thanks to a skew-symmetric formulation of the nonlinear term. A fourth order Runge-Kutta time discretization and a small timestep allowed us to observe numerical dissipation of energy remaining below 10^{-4} relative to the initial energy, even for zero viscosity.

Here, we use a slightly different approach which applies both to the 1D Burgers and 2D Euler equations. A Fourier expansion with pseudo-spectral evaluation of the nonlinear term is still utilized, but the product in physical space is now fully dealiased. In other words, the Fourier modes retained in the expansion of the solution are such that $|\mathbf{k}| \leq k_C$, where k_C is the desired cut-off wavenumber, but the grid has $N = 3k_C$ points in each direction, versus $N = 2k_C$ for a non-dealiased, critically sampled product. This dealiasing makes the pseudo-spectral scheme equivalent to a Fourier-Galerkin scheme up to round-off errors (Canuto et al., 1988), and thus conservative. For time discretization we stick to Runge-Kutta schemes, of order 4 for the 1D Burgers equation and of order 3 with a low storage formulation (Orlandi, 2000, page 20) for the 2D Euler equations.

All the required Fourier transforms are computed thanks to the FFTW library (Frigo and Johnson, 2005). This allows us to exploit at the same time OpenMP and the Message Passing Interface (MPI) for parallelization. To reach resolutions up to 8192×8192, 8 cores were utilized on a shared memory cluster node. For some of the simulations, we used 64 cores and 32 MPI processes on an IBM Regatta Power6 machine at the IDRIS-CNRS french computing center.

IV.2.2.2 Regularization

We call "filtering scheme" any operator that can be applied to the discrete numerical solution at the end of each timestep, before advancing to the next timestep. In this subsection, we attempt to describe the filtering algorithms for a generic solution f, knowing that they will later be applied either to $f = u$ for the 1D Burgers case or to $f = \omega$ for the 2D Euler case. Since the filters that we are interested in are applied in the wavelet domain, we start by computing the transform coefficients of the solution with respect to the space variable (see the Appendix for technical details). Given orthogonal wavelets $(\psi_\lambda)_{\lambda\in\Lambda}$ and the associated scaling function φ at the largest scale, the solution is thus expanded into (see e.g. (Mallat, 1999)):

$$f = \langle f \mid \varphi\rangle\varphi + \sum_{\lambda\in\Lambda}\langle u \mid \psi_\lambda\rangle\psi_\lambda \ , \tag{IV.2.6}$$

where λ is a multi-index giving the scale, position, and, in 2D, direction of each wavelet. Later we denote by $\tilde{f}_\lambda$ the wavelet coefficients $\langle u \mid \psi_\lambda\rangle\psi_\lambda$ of f. The inner product is defined in 1D by $\langle a \mid b\rangle = \int_{-1}^{1} a(x)\cdot b^*(x)\mathrm{d}x$, where $\cdot^*$ stands for complex conjugation, and in 2D by $\langle a \mid b\rangle = \int_{\mathbb{T}^2} a(\mathbf{x})\cdot b^*(\mathbf{x})\mathrm{d}x$.

We consider two wavelet families. The first one, well known to numerical analysts, is the Coiflet orthonormal family with supports of length 12 (Daubechies, 1993), proposed in (Farge et al., 1992a) to represent turbulent flows. The second one is the Qshift-B family associated with the dual-tree complex wavelet transform (DTCWT) (Kingsbury, 2001), that we propose to dub Kingslet. Although it was primarily designed for image processing, the latter also offers many attractive features from a computational point of view. Its main characteristics can be summarized as follows:

(i) the wavelets and scaling functions are complex-valued,

(ii) the family is not orthogonal, but it is a tight frame, which is a kind of overcomplete basis, having in d dimensions 2^d times more elements than an orthonormal basis at the same resolution,

(iii) near translation invariance is built into the transform, despite its dyadic structure,

(iv) in two and more spatial dimensions, the wavelets have much better directional selectivity than orthogonal wavelets.

Since the Kingslets constitute a tight frame, the reconstruction formula (IV.2.6) is still valid (Kingsbury, 2001), as long as the right hand side is replaced by its real part, f being real-valued.

Since this is the key point, let us explicit what we mean by translation invariance. Most multiscale approximation schemes involve the design of some sort of recursive grid or subdivision pattern. Normally these recursive structures depend on the choice of the origin and therefore are not invariant by translation. This is in particular the case for orthogonal wavelets, which are arranged as a dyadic tree. In most cases this tree has nothing to do with the phenomenon or signal one wants to study, but since it is built into the transform it tends to introduce some artifacts as soon as the coefficients are processed in some way. One solution consists in averaging the results over all possible shifts of the tree with respect to the signal. It is known as the undecimated wavelet transform (Coifman and Donoho, 1995).

Kingslets offer a more elegant solution to the same problem, namely, to average only a fixed number of trees (2^d in d dimensions), but under the condition that the wavelets associated to those trees are approximate Hilbert transform pairs. When this is done, the projectors on each approximation subspace almost commute with all possible signal shifts.

Once the wavelet coefficients of the solution have been obtained, we consider two kinds of filters: a linear one, and a nonlinear one. The linear filter consists in discarding only the wavelet coefficients at the finest scale. For the Kingslets case, this linear filter is equivalent to a simple convolution of the solution with the scaling function at the finest scale. It is nothing more than a lowpass filter, but with a special transfer function that happens to be the Fourier transform of a scaling function. This property does not hold for orthogonal wavelets, because the downsampling operation applied to the scaling function coefficients at the finest scale introduces aliasing effects (Kingsbury, 2001).

For the nonlinear filter, before discarding the wavelet coefficients at the finest scale, we apply a threshold to all wavelet coefficients: those whose moduli are below a certain value Θ are set to zero. The threshold Θ depends on the function itself and is computed by means of an iterative method (Azzalini et al., 2004). The resulting value is characterized by the following implicit relationship:

$$\Theta^2 = \frac{5}{N_I(\Theta)} \sum_{\lambda \in \Lambda} |\widetilde{f}_\lambda|^2 H(\Theta - |\widetilde{f}_\lambda|) \;, \tag{IV.2.7}$$

where H is the Heaviside step function and $N_I(\Theta) = \sum_{\lambda \in \Lambda} H(\Theta - |\widetilde{f}_\lambda|)$ is the number of wavelet coefficients below the threshold. Formula (IV.2.7) simply means that the threshold equals 5 times the standard deviation of the coefficients below it. Note that there is no adjustable parameter. The nonlinear filter will be referred to as CVS filter, where CVS stands for Coherent Vorticity Simulation (Farge et al., 1999). After applying either filter, the solution is reconstructed by inverse wavelet transform, and its Fourier coefficients by fast Fourier transform (FFT), so that the simulation can proceed with the next timestep.

The viscous or hyperviscous term, when present, is included in the computation without loss of accuracy thanks to an integrating factor method (Trefethen, 2000, page 111). The integrating factor is simply $e^{\nu k^2 t}$, with k the Fourier wavenumber modulus.

The six different regularization methods that we are going to compare are:

(i) Viscous: $\alpha = 1$ and $\nu_1 > 0$ is chosen high enough so that all scales of motion are resolved without applying any filter,

(ii) Hyperviscous: $\alpha = 4$, and $\nu_4 > 0$ is defined by $\nu_4 = \nu_1 k_C^{-6}$,

(iii) Real linear: $\nu = 0$, and we apply linear wavelet filtering of the finest scale using Coiflets,

(iv) Complex linear: $\nu = 0$, and we apply linear wavelet filtering of the finest scale using Kingslets,

(v) Real CVS: $\nu = 0$, and we apply nonlinear CVS filtering using Coiflets,

(vi) Complex CVS: $\nu = 0$, and we apply nonlinear CVS filtering using Kingslets.

Type	α	ν_α	Wavelet	Filter
Viscous **(i)**	1	>0	none	none
Hyperviscous **(ii)**	4	>0	none	none
Real linear **(iii)**	-	$=0$	Coiflet	linear
Complex linear **(iv)**	-	$=0$	Kingslet	linear
Real CVS **(v)**	-	$=0$	Coiflet	nonlinear
Complex CVS **(vi)**	-	$=0$	Kingslet	nonlinear

Table IV.2.1: Summary of all employed regularization methods.

A summary is given in table IV.2.1. We shall also compare the regularized solutions to the one obtained without any regularization mechanism, that is, by cautiouslessly applying our numerical scheme to the inviscid equation. The solution does not blow up even in this case since the numerical scheme is conservative. However, we are not claiming that it is the genuine solution of the – non-truncated – 2D Euler equations, and in fact it cannot be, due to the abrupt cut-off in Fourier space.

IV.2.2.3 Initial conditions

For the 1D Burgers equation, we choose the initial condition $u_0(x) = -\sin(\pi x)$, which induces the formation at $t = \frac{1}{\pi}$ of a single shock located in the middle $x = 0$ of the interval $[-1, 1]$. For the viscous simulations **(i)**, we choose $\nu_1 = 2N^{-1}$. A classical estimate says that the shock width is proportional to ν_1, hence ν_1 has to be larger than a constant times N^{-1} for the simulation to be properly resolved. The constant is adjusted by trial and errors. The value of ν_4 for all the hyperviscous simulations is given by $\nu_4 = \nu_1 k_C^{-6}$, where $k_C = \frac{N}{3}$.

For the 2D Euler equations, we consider two kinds of initial conditions. The first one is a vortex merger, made of two Gaussian vortices with positive circulation and one weaker Gaussian vortex with negative circulation (Kevlahan and Farge, 1997). The formula giving the initial vorticity is:

$$\omega(t=0,x,y) = \pi \left(e^{-\pi^2\left(\left(x-\frac{3\pi}{4}\right)^2+(y-\pi)^2\right)} + e^{-\pi^2\left(\left(x-\frac{5\pi}{4}\right)^2+(y-\pi)^2\right)} - \frac{1}{2} e^{-\pi^2\left(\left(x-\frac{5\pi}{4}\right)^2+\left(y-\pi\left(1+\frac{\sqrt{2}}{4}\right)\right)^2\right)} \right) \tag{IV.2.8}$$

where $(x, y) \in [0, 2\pi]^2$. For the viscous simulations starting from this initial condition, we have used $\nu_1 = 13.1072N^{-2}$.

The second initial condition is a correlated Gaussian noise. To generate it, we use the Fourier representation:

$$\omega(t=0,\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{N}^2,|\mathbf{k}|\leq k_C} \widehat{\omega}[\mathbf{k}] e^{i\mathbf{k}\cdot\mathbf{x}} \quad \text{where} \quad \widehat{\omega}[\mathbf{k}] = \frac{1}{4\pi^2}\int_{\mathbb{T}^2} \omega(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} \mathrm{d}^2\mathrm{x} \tag{IV.2.9}$$

and, defining $k = |\mathbf{k}|$, we then let

$$\hat{\omega}[t=0,\mathbf{k}] = \frac{1}{24\pi^2} \times \begin{cases} k e^{i\theta_{\mathbf{k}}} \text{ if } k \leq 6 \\ 6^2 k^{-1} e^{i\theta_{\mathbf{k}}} \text{ if } 6 < k \leq 42 \\ 0 \text{ otherwise} \end{cases} \tag{IV.2.10}$$

where the $\theta_{\mathbf{k}}$ are pseudo-random numbers uniformly and independently distributed in $[0, 2\pi[$. The $\theta_{\mathbf{k}}$ are drawn once using a Mersenne twister pseudo-random number generator (Matsumoto and Nishimura, 1998), and the same values are then used for all cases studied below. The initial isotropic enstrophy spectrum

$$Z[k] = 2\pi^2 \sum_{k \leq |\mathbf{k}| < k+1} |\widehat{\omega}[\mathbf{k}]|^2$$

is then approximately:

$$Z[t=0,k] \simeq \frac{1}{144\pi} \times \begin{cases} k^3 \text{ if } k \leq 6 \\ 6^4 k^{-1} \text{ if } 6 < k \leq 42 \\ 0 \text{ otherwise.} \end{cases} \tag{IV.2.11}$$

The energy of this initial condition is $E(t=0) = \frac{1}{2}\langle \mathbf{u}_0 \mid \mathbf{u}_0 \rangle \simeq 7.847 \cdot 10^{-2}$ and its enstrophy is
$Z(t=0) = \sum_{k=0}^{\infty} Z[k] = \frac{1}{2}\langle \omega_0 \mid \omega_0 \rangle \simeq 6.289$. For the viscous simulations starting from this initial condition, we have used $\nu_1 = 6.5536 N^{-2}$.

IV.2.3 Results

IV.2.3.1 1D Burgers

We first briefly recall the results for the 1D Burgers equation (Nguyen van yen et al., 2008), which are now confirmed by using a fully dealiased code, as described above. Compared to (Nguyen van yen et al., 2008), we thus have even less numerical dissipation. An upper bound is estimated by performing simulations with $\nu = 0$ without filtering, where we observe that for $N = 4096$ the fraction of the initial energy that has been dissipated at $t = \frac{5}{\pi}$ is less than 10^{-8}.

Recall that for Burgers equation, it is possible to compute the entropy solution to (IV.2.5) with a very high precision, using a Legendre transform method (Burgers, 1974). This procedure is particularly straightforward for the simple initial condition leading to a single shock that we have used. The obtained entropy solution will be our reference u_{ref} for this section. Note that the entropy solution dissipates energy, contrary to the Galerkin-truncated one, obtained by solving the inviscid Burgers equation using our conservative numerical scheme.

The solutions obtained with methods **(i)**, **(ii)**, **(v)** and **(vi)** for $N = 4096$ are shown in Fig. IV.2.1 (left). We observe that **(i)** and **(vi)** are quite similar, whereas **(v)** has much more pronounced oscillations. In Fig. IV.2.1 (right), the squared modulus of the Morlet wavelet transform (Grossmann and Morlet, 1984) of each solution is also shown using a logarithmic colorscale, which allows us to visualize the errors in space and scale. The artifacts due to aliasing in the Coiflet CVS filtering **(v)** are clearly visible on this figure. They are much less

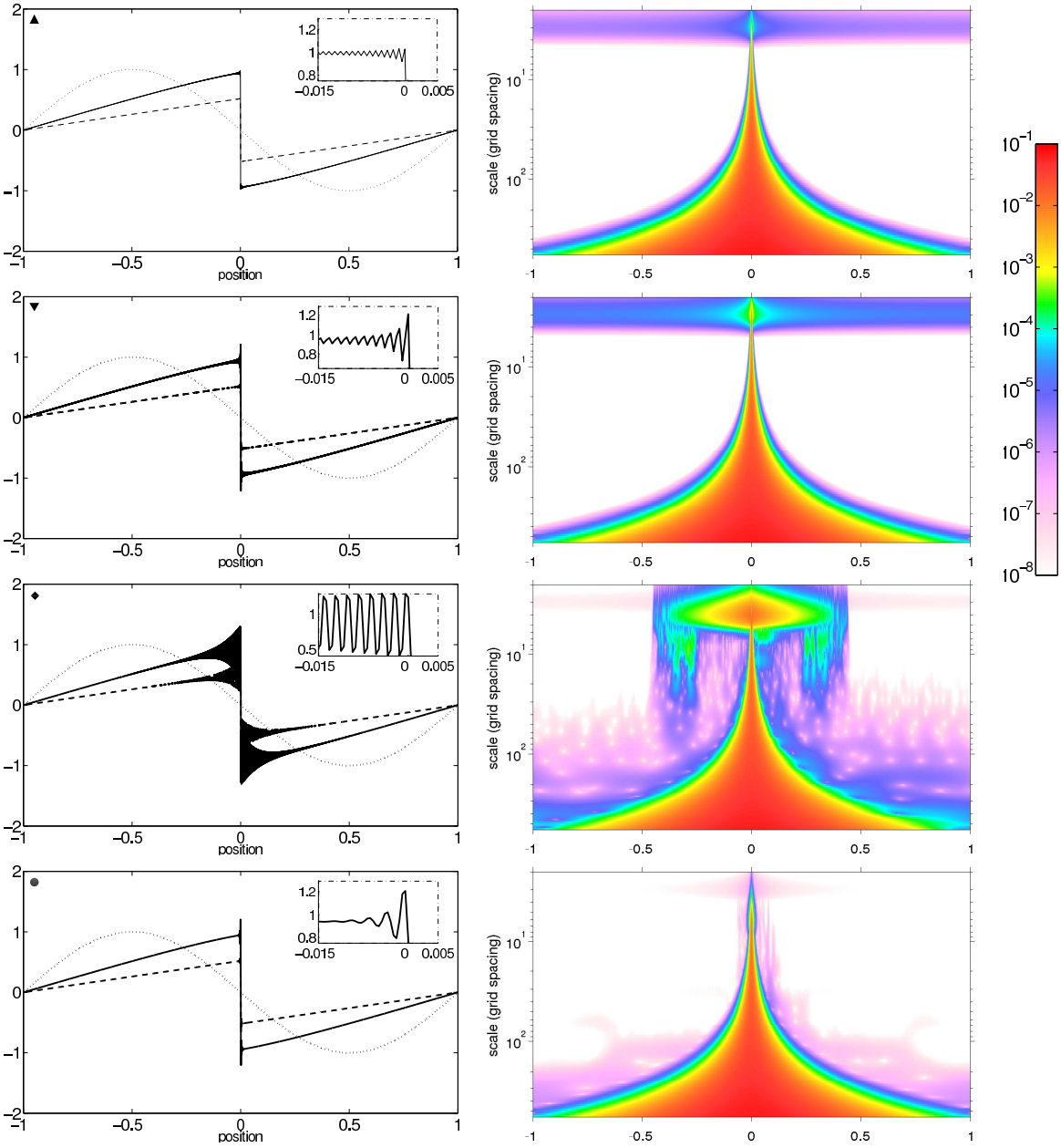


Figure IV.2.1: Left column: numerical solution to the 1D Burgers equation for methods **(i)**, **(ii)**, **(v)** and **(vi)** (top to bottom) with $N = 4096$. Dotted lines: $t = 0$, solid lines: $t = \frac{2}{\pi}$, dashed lines: $t = \frac{5}{\pi}$. The insets show a zoom around $x = 0$. Right column: squared modulus of the Morlet coefficients of each solution at $t = \frac{5}{\pi}$. The vertical scale is logarithmic and indicates the wavelet scale factor, normalized by the distance between two grid points. Note that the color scale is also logarithmic.

pronounced in the Kingslet case **(vi)**. The Gibbs phenomenon can also be noticed on the viscous **(i)** and hyperviscous **(ii)** solutions.

Two error measures allow us to quantify the difference between a numerical solution u_N and the reference entropy solution u_{ref}. The first error measure is defined as:

$$\epsilon_w = \frac{2\pi}{\|u_{ref}\|} \left(\sum_{|k|\leq 64} |\widehat{u}_N[k] - \widehat{u}_{ref}[k]|^2 \right)^{\frac{1}{2}} \tag{IV.2.12}$$

where $\|\cdot\|$ stands for the L^2 norm. This quantity is of practical interest, since it frequently occurs that only the low wavenumber behavior of the solution needs to be accurately described. From the mathematical point of view, it can be seen as an indicator of weak convergence. The second error measure is the total L^2 error:

$$\epsilon = \frac{\|u_N - u_{ref}\|}{\|u_{ref}\|} \tag{IV.2.13}$$

which measures strong convergence to the entropy solution.

Our first result is that in these simulations, weak convergence of the Galerkin-truncated inviscid simulations to the entropy solution does not occur (Fig. IV.2.2, left), in agreement with a remark made by Tadmor in (Tadmor, 1989). This confirms, if need be, that regularization is indispensable to solve Burgers equation with a spectral method. Post-processing the result of an unregularized simulation does not yield a physically relevant solution. On the contrary, the solutions given by all other methods appear to converge weakly to the entropy solution with a rate $O(N^{-1})$. The error associated to the viscous and Kingslets solutions is about one order of magnitude smaller than the error associated to the Coiflet solution.

The solutions obtained with every considered regularization methods also converge strongly towards the entropy solution with a rate $O(N^{-1/2})$ (Fig. IV.2.2, middle). This confirms the results of (Nguyen van yen et al., 2008) and extends them to several new cases, namely real-valued wavelets and linear filtering. As a sidenote, we confirm that the hyperdissipative 1D Burgers model does not improve in any way the results compared to the viscous model (Frisch et al., 2008).

In Fig. IV.2.2 (right), we have plotted the compression rate, defined as follows:

$$\rho = \frac{N}{N_C} \tag{IV.2.14}$$

where $N_C = N - N_I$ is the number of wavelet coefficients above the threshold. The subscripts C and I stand respectively for coherent and incoherent. The higher the compression rate, the better the compression.

The following conclusions may be drawn from the results in this section. To regularize the 1D inviscid Burgers equation, it is enough to filter out the wavelet coefficients at the finest scale. The results are better for complex-valued wavelets because of their reduced aliasing effects. The main advantage of CVS wavelet filtering compared to other regularization methods is that they allow us to compress the information needed to represent the solution. Here, we observe that the compression rate ρ is about 3 for $N = 1024$ and improves slowly for higher resolutions, a fact that we will comment more on in the conclusion.

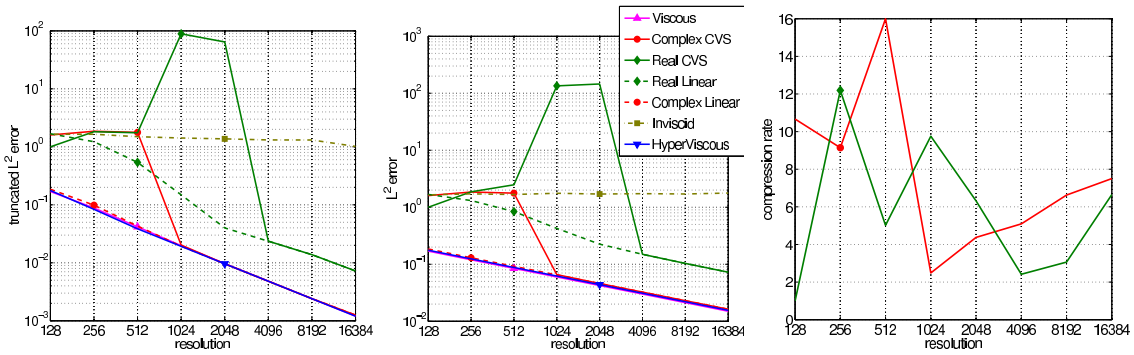


Figure IV.2.2: Comparison between the six regularization methods to solve the 1D Burgers equation. Left: normalized L^2-error for the Fourier modes with $|k| \leq 64$. Middle: total normalized L^2-error. Right: Compression rate $\rho = \frac{N}{N_C}$.

IV.2.3.2 2D Euler

As in the previous subsection, we study the results obtained with methods **(i)**-**(vi)**, but this time to solve the 2D Euler equations (IV.2.3). Five increasingly refined resolutions were used in each case: $N = 2^7, \ldots, 2^{11}$. Since we do not have an efficient way of computing the solution to (IV.2.3) with high precision, we use the viscous simulation **(i)** with $N^2 = 8192^2$ as a reference. Inviscid simulations at $N^2 = 2048^2$ show that the fraction of the initial enstrophy dissipated during the whole simulation is less than $5 \cdot 10^{-2}$, and less than 10^{-5} for the energy. Note that the numerical dissipation is due to the error coming from the time discretization of the equations, and should therefore be smaller in the viscous and filtered simulations, since the corresponding solutions are smoother in time.

Deterministic initial condition In this subsection, we consider the vortex merger initial condition (IV.2.8). The vorticity field at $t = 50$ for methods **(i)**, **(ii)**, **(v)** and **(vi)** are shown in Fig. IV.2.3. The reference vorticity field (top, left) and the result of the inviscid simulation (bottom, right) are also shown for comparison purposes. The gross dipolar structure of the vorticity field is well preserved by all methods. The main differences have to do with sharp gradients that are generated very rapidly during the merging of the two positively signed vortices. Each positive vortex corresponds to a nonuniform velocity field that causes the other positive vortex to rotate around it and deform, producing a spiral pattern that is still visible in Fig. IV.2.3. These sharp gradients are clearly not very well handled by method **(v)**, since artifacts have been introduced even in quiet regions of the domain.

The different regularization mechanisms induce a different behavior of the solution in a global sense, as can be seen on the time evolutions of energy and enstrophy, and on the enstrophy spectra (Fig. IV.2.4).

We now examine the difference between all the solutions and the reference solution in a quantitative manner. In Fig. IV.2.5 (left), we plot the truncated L^2-error, which is defined by (IV.2.12) with an additional 2π factor due to normalization. These curves illustrate that even the low wavenumber dynamics are affected by the regularization mechanism. Nevertheless, the error seems to converge to zero for all considered regularization methods. The rate is $O(N^{-2})$ for the viscous **(i)** simulations, consistent with known analytical results concerning the inviscid limit of Navier-Stokes equations (Kato, 1972). The arguments in (Kato, 1972)

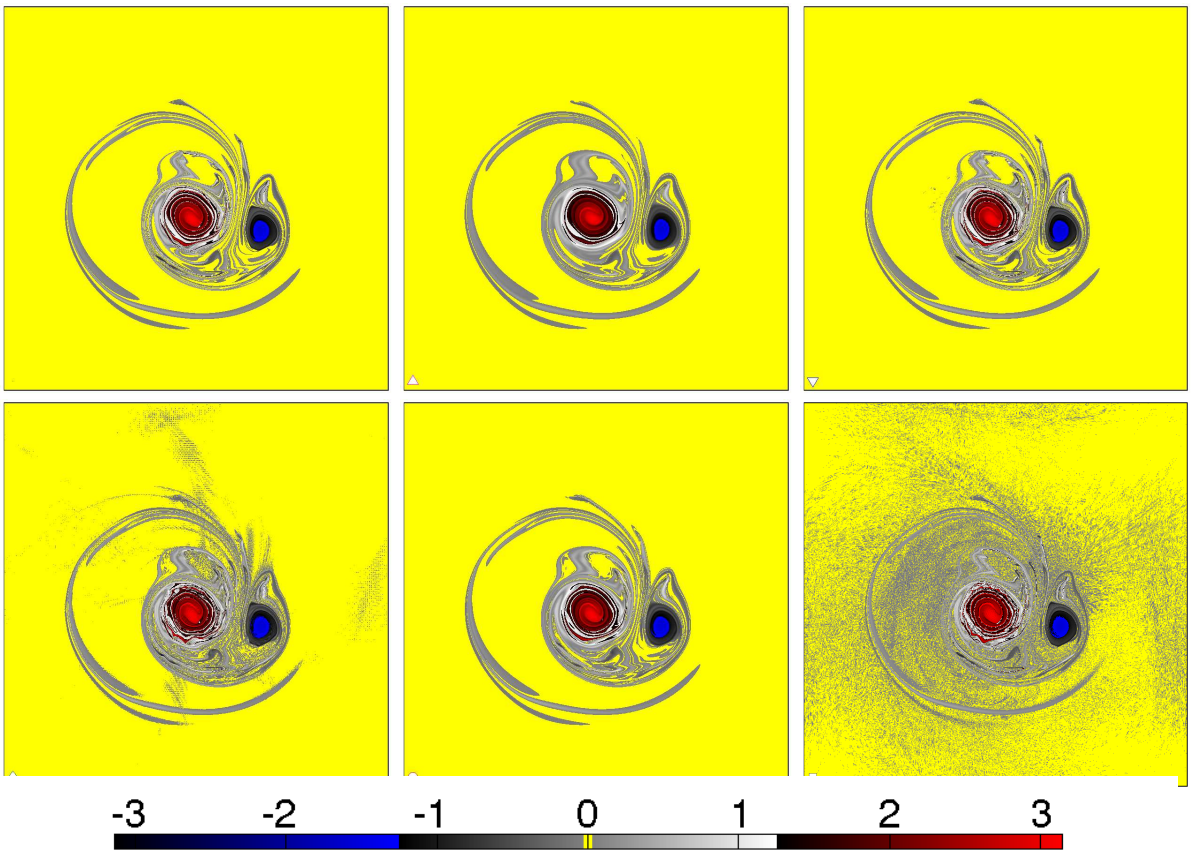


Figure IV.2.3: Vorticity fields at $t = 50$ for the deterministic initial condition. Top left: reference Navier-Stokes solution, $N = 8192$. Top middle: Navier-Stokes solution **(i)**, $N = 2048$. Top right: Hyperdissipative Navier-Stokes solution **(ii)**, $N = 2048$. Bottom left: Coiflet CVS filtered Euler solution **(v)**, $N = 2048$. Bottom middle: Kingslets CVS filtered Euler solution **(vi)**, $N = 2048$. Bottom right: Galerkin-truncated Euler solution, $N = 2048$.

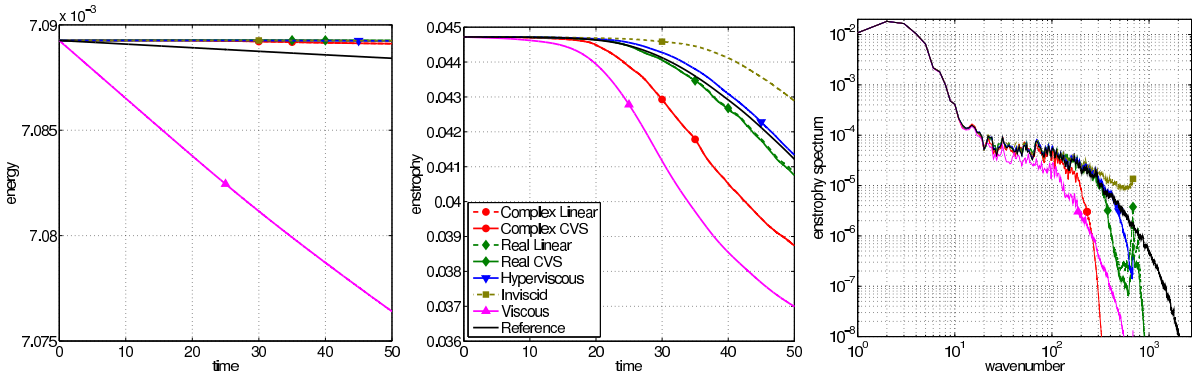


Figure IV.2.4: Deterministic initial condition: results for the different methods with $N = 2048$, and reference viscous solution with $N = 8192$. Left: time evolution of energy. Middle: time evolution of enstrophy. Right: enstrophy spectra at $t = 50$.

apply equally well to the hyperviscous Navier-Stokes equations, and we accordingly observe the same convergence rate for **(ii)**. The rate is less easy to determine for the other methods, but it seems to lie between $O(N^{-2})$ and $O(N^{-1})$.

Now the total L^2-error, defined similarly to (IV.2.13) above, is represented in Fig. IV.2.5 (right). The results are strikingly close to the previous ones, indicating that most of the L^2-error in fact comes from low wavenumbers. The results concerning the inviscid Galerkin-truncated Euler solution (square marker) are particularly interesting since they contrast with those observed for the 1D Burgers equation above. It was previously observed in (Cichowlas et al., 2005) "that the spectrally truncated [3d] Euler equations have long-lasting transients behaving just like those of the dissipative [3d] Navier-Stokes equations". This result is also supported by calculations using the eddy-damped quasi-normal Markovian closure theory (Bos and Bertoglio, 2006). Our results are unfortunately limited to two space dimensions, but they are stronger, since they indicate that the solution to the Galerkin-truncated 2D Euler equations converges to the solution of Euler equations in the L^2 norm.

By looking at the numbers presented in this section, one could draw the hasty conclusion that hyperviscous or even inviscid simulations do the job very well, so why bother with wavelet filtering ? But as we have already seen for Burgers equation, the advantage of wavelets has to do with compression. We now turn to random initial conditions, in order to demonstrate this for Euler equations.

Random initial condition By computing the flow evolution starting from the randomly generated initial condition (IV.2.10), we obtain the vorticity field at $t = 50$ for methods **(i)**, **(ii)**, **(v)** and **(vi)**, shown in Fig. IV.2.6. As above, the reference vorticity field (top, left) and the result of the inviscid simulation (bottom, right) are also given. Only the subdomain $[\pi, 2\pi] \times [\pi, 2\pi]$ is represented on the figure, in order to make the details more visible. The first observation that can be put forward is that the positions of all vortices match pretty well between all represented fields. Of course, this situation is bound not to last for long, since fully developed 2D turbulence is sensitive to initial conditions: the different simulations should separate sooner or later and end up having nothing left in common. But this occurs on a longer time scale than the one we have considered here. Note that the initial eddy turnover time $\tau := \frac{1}{\sqrt{2Z(t=0)}}$ is approximately 0.282 here.

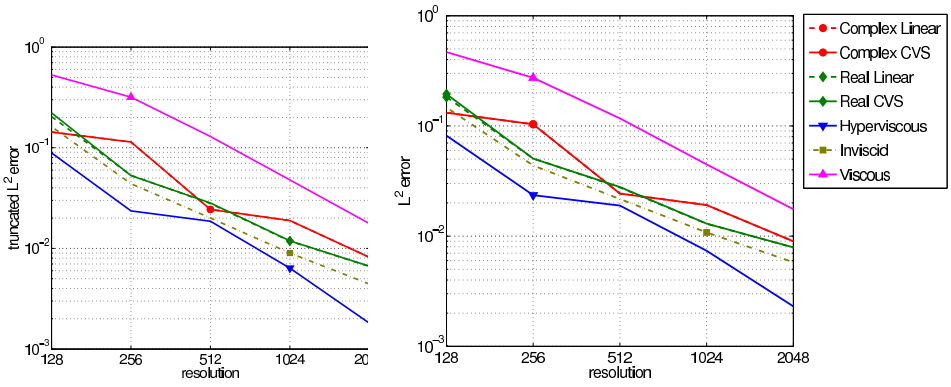


Figure IV.2.5: Left: truncated L^2 error ϵ_w (IV.2.12) at $t = 50.05$ with respect to the reference solution. Right: total L^2 error ϵ (IV.2.13) at $t = 50.05$ with respect to the reference solution.

We would like to outline two kinds of differences between the fields yielded by the different methods. First, their high wavenumber behavior is quite different. This can be checked on the enstrophy spectra (Fig. IV.2.8, right). In this respect, the similarity between the Kingslet-filtered **(vi)** (bottom, middle) and viscous (top, middle) vorticity fields is striking. Due to its better Fourier localization of the dissipation term, the hyperviscous simulation reproduces even better the high wavenumber features of the reference simulation, as is highlighted by the cuts in Fig. IV.2.7.

The second difference can be seen in Fig. IV.2.8 (left) and concerns only regularization by CVS filtering using Coiflets **(v)**, for which energy starts increasing again after $t \simeq 60$, contrary to all other methods. This effect suggests that the flow has been affected by the regularization method in an unphysical way, and was our main motivation for introducing Kinsglets in this study as a better alternative.

The time evolution of enstrophy is shown in Fig. IV.2.8 (middle). Even the reference simulation is still quite dissipative concerning enstrophy. This observation is consistent with a very slow decay of the enstrophy dissipation rate as a function of Reynolds number (Dmitruk and Montgomery, 2005; Tran and Dritschel, 2006; Lopes Filho et al., 2006), which implies a slow convergence to the Euler solution in the H^1 norm. Note that for the viscous solution, convergence is known to occur at a rate $O(\nu) = O(N^{-2})$ in any norm H^m (Kato, 1972), since the initial condition considered here consists in a finite number of Fourier modes and is therefore analytic. We believe that this rate is impossible to observe numerically because the constant in front of ν grows extremely fast in time. Although it would be interesting to consider also the H^1-error, we conclude that this quantity is out of our reach here. Indeed, because of the slow convergence, the viscous solution at $N = 8192$ cannot be used as a reference to compute the H^1 norm error between **(i)**-**(vi)** and the Euler solution, nor can any solution computed with current day methods and computational power.

In Fig. IV.2.8 (right), the enstrophy spectra of the different solutions are compared with the one of the reference. They all agree well at low wavenumbers, as can be seen in the inset.

Finally, we study the convergence in the inviscid limit, for fixed time and fixed initial condition (Fig. IV.2.9). As in the case of the deterministic initial condition, the truncated

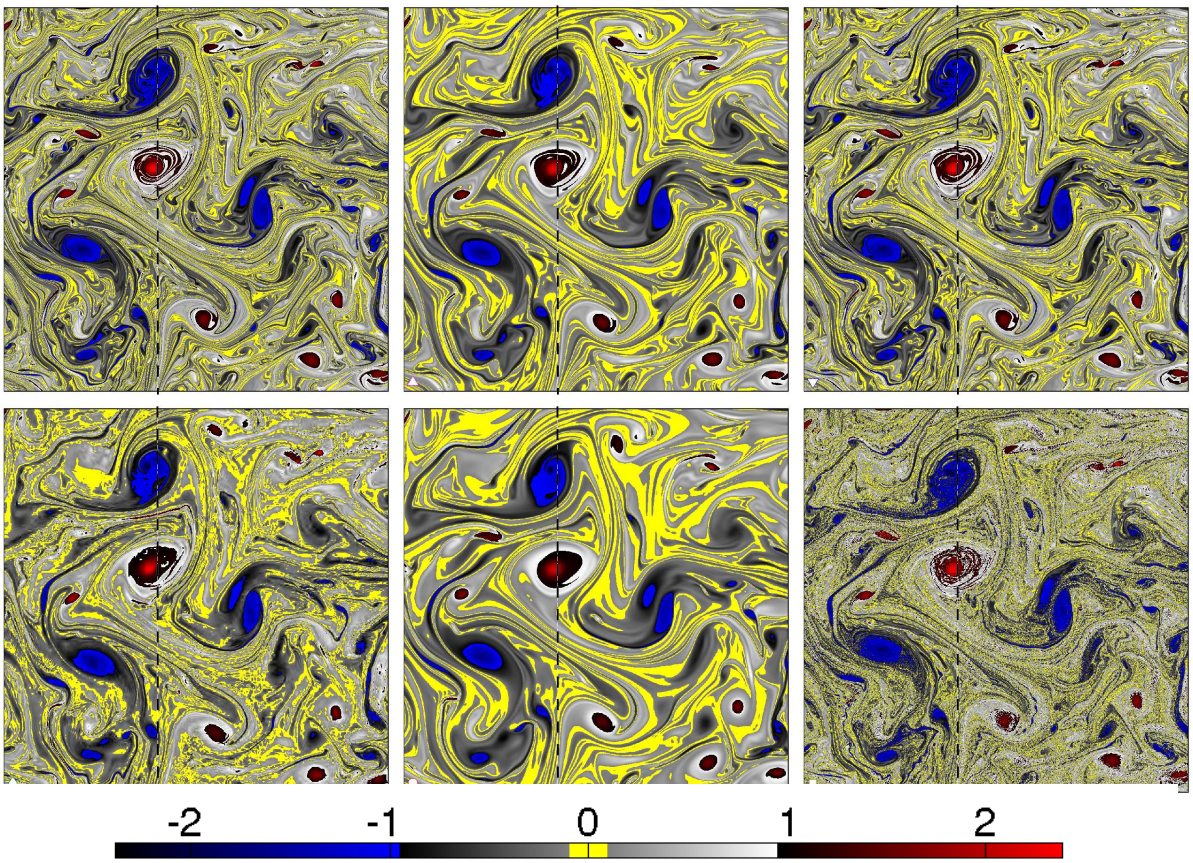


Figure IV.2.6: Vorticity fields at $t = 50$ for the random initial condition, restricted to the square $[\pi, 2\pi] \times [\pi, 2\pi]$. Top, left: reference Navier-Stokes solution with $N = 8192$. Top, middle: Navier-Stokes solution **(i)**, $N = 2048$. Top, right: Hyperdissipative Navier-Stokes solution **(ii)**, $N = 2048$. Bottom, left: Coiflet CVS filtered Euler solution **(v)**, $N = 2048$. Bottom, middle: Kingslets CVS filtered Euler solution **(vi)**, $N = 2048$. Bottom, right: Galerkin-truncated Euler solution, $N = 2048$. The black dashed lines indicates the location of the cuts shown in Fig. IV.2.7.

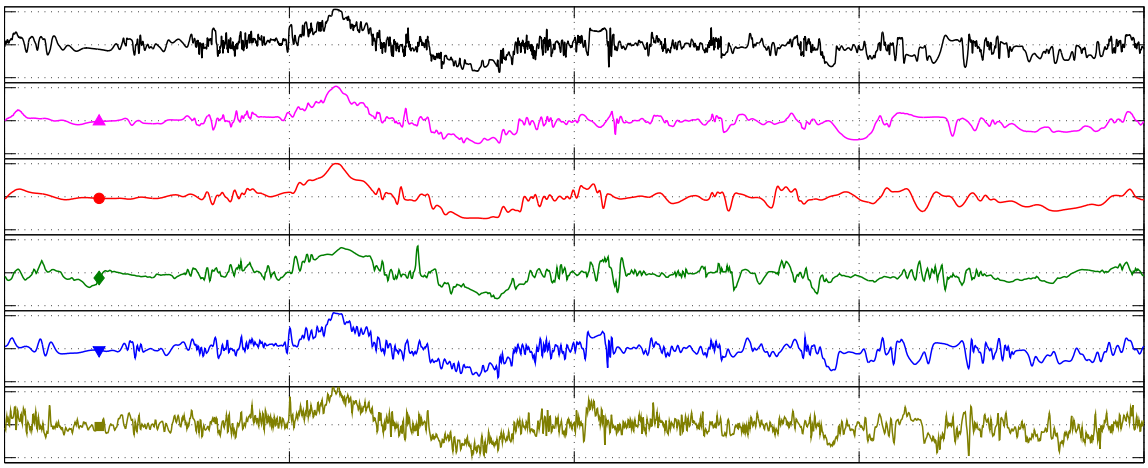


Figure IV.2.7: Cuts through the vorticity fields at $x = \frac{7\pi}{5}$ and $t = 50$, for the different methods with $N = 2048$, and for the reference viscous solution whith $N = 8192$. The legend can be found in Fig. IV.2.8.

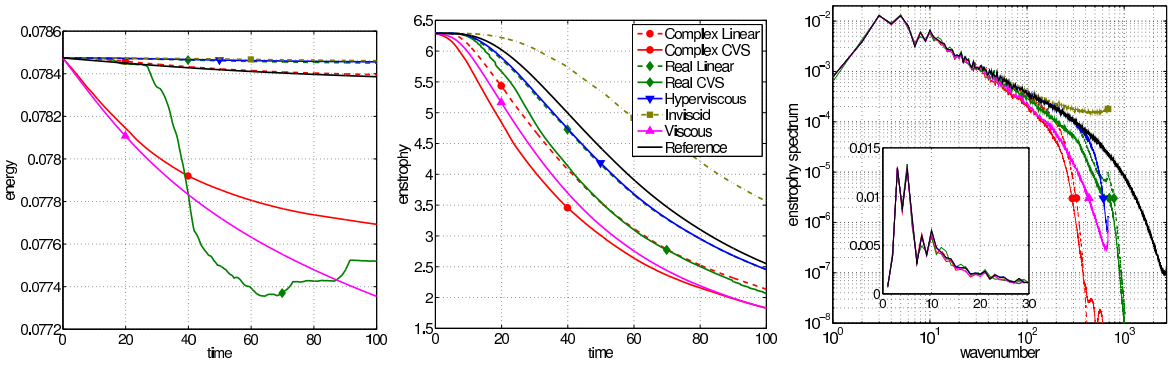


Figure IV.2.8: Random initial condition, advanced with different methods for $N = 2048$, and reference viscous solution for $N = 8192$. Left: time evolution of energy. Middle: time evolution of enstrophy. Right: enstrophy spectra at $t = 50$. The inset shows the low wavenumber part in lin-lin coordinates. The legend applies to all three graphs.

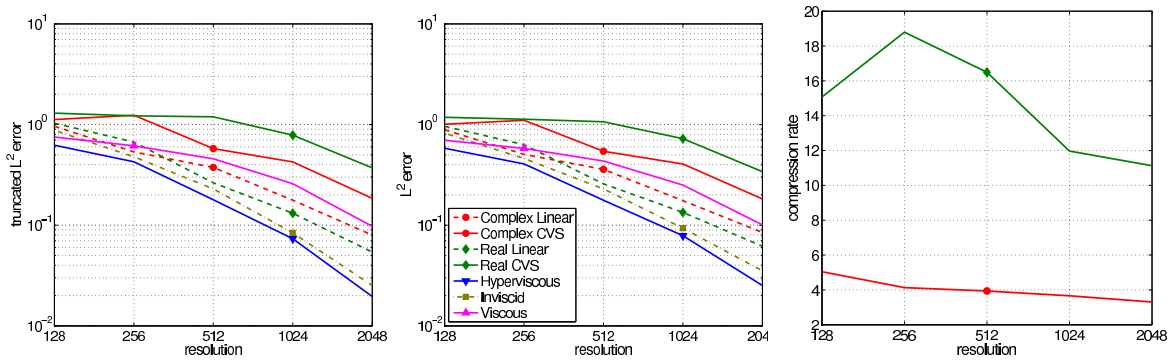


Figure IV.2.9: Random initial condition. Left: truncated L^2-error ϵ_w (IV.2.12) with respect to the reference solution. Middle: total L^2-error ϵ (IV.2.13) with respect to the reference solution. Right: compression rate (IV.2.14) achieved by nonlinear wavelet thresholding. The legend applies to all three graphs.

L^2-error ϵ_w and the total L^2-error ϵ are represented. The error decay is less pronounced in the considered range of Reynolds number than it was in the deterministic case above. It seems that the Reynolds number at which the analytically predicted $O(N^{-2})$ asymptotic behavior manifests itself has not been reached yet.

We now comment on the compression rate ρ (IV.2.14), plotted in Fig. IV.2.9, right. One can see that the CVS thresholding allows us to compress by a factor of about 3 using Kingslets, and about 10 using Coiflets. The computations presented here do not benefit from this compression, since the solution is reconstructed in Fourier space at each timestep. Nevertheless, the measured compression rate remains a good indicator of what can be achieved by adaptive wavelet-based solvers (Farge et al., 2001; Schneider and Vasilyev, 2010) using the thresholding rule that we have defined earlier.

IV.2.4 Conclusion and Perspectives

We have used a classical Fourier-Galerkin method to solve the 2D Euler equations, and we have proposed a new wavelet-based approach to regularize the solution, in order to obtain a good numerical approximation of the Euler solution. We have compared our method

with viscous and hyperviscous regularization, and also with the solution computed without regularization. The ability of all regularization methods to approach the inviscid dynamics was measured quantitatively by direct comparison with a high resolution viscous simulation. High resolution numerical simulations were made possible thanks to parallel implementations of the spectral code and of the wavelet transform.

The main findings concerning the 2D Euler equations can be summarized as follows:

- linear wavelet filtering approaches the inviscid dynamics as well as viscous or hyperviscous dissipative terms,
- nonlinear wavelet filtering with real-valued wavelets (Coiflets) introduces undesirable artifacts in the solution,
- nonlinear wavelet filtering with complex-valued wavelets (Kingslets) preserves the dynamics and offers at the same time a non negligible compression rate of about 3 for fully developed turbulence,
- there are strong indications that the solution to the Galerkin-truncated 2D Euler equations converges, at least in the weak sense, to the solution to the 2D Euler equations.

We can hence claim that the main goals of our wavelet regularization method have been attained. Indeed, we have shown that CVS filtering with Kingslets allows us to regularize the 2D Euler equations, and at the same time compress the solution. This is a strict extension of the results obtained for the 1D Burgers equation. One possible explanation for the disappointing results observed with Coiflets is the lack of a safety zone in wavelet space (Schneider et al., 2006). Since the orthogonal wavelet transform is not translation invariant, it cannot properly capture structures that are advected in a continuous fashion by the flow. The rather low compression rates observed when filtering solutions of the 1D Burgers equation should be seen as a limitation of the employed thresholding method, and not of the wavelets themselves. Indeed, for a function as simple as single shock, wavelets are in principle able to achieve much larger compression rates. Improving the threshold selection algorithm is a priority goal of our ongoing research work.

We would now like to draw attention to the striking differences between the results that we found for the 1D Burgers equation on the one hand and for the 2D incompressible Euler equations on the other hand. Indeed, the Galerkin-truncated dynamics of these two equations behave very differently in comparison to their non truncated counterparts: for the 2D Euler equations we have shown weak convergence of the solution to the Galerkin-truncated equations towards the inviscid solution[5] , while for the 1D Burgers equation no such convergence was observed, as Tadmor has shown using an elegant mathematical argument (Tadmor, 1989). Therefore, one should be extremely cautious when using Burgers equation as a toy model for understanding incompressible turbulence. One may argue that the 2D Euler dynamics is pathological since it does not have finite time singularity (FTS), and that Burgers is in this respect closer to 3D Euler. But even if the 3D Euler equations have FTS, these singularities cannot be of the same nature than the shocks occuring in the solutions of Burgers equation. The incompressibility condition plays a key role by introducing nonlocal effects that are completely absent in the Burgers model. To illustrate these differences further,

[5]A rigorous proof of strong convergence has been obtained since this paper was published (C. Bardos and E. Tadmor, private communication, 2010)

consider the results obtained with hyperviscous regularization. In (Frisch et al., 2008), it was argued that hyperviscosity introduces a quasi-thermalized range in the spectrum. As a consquence, all that is gained by the better wavenumber localization of the dissipation term is lost because of an enhanced bottleneck. These results were supported by 1D Burgers direct numerical simulations, and by the eddy-damped quasi-normal Markovian (EDQNM) closure for 3D turbulence. Our results are in agreement concerning 1D Burgers, but show that hyperviscous regularization offers very good performance for 2D Euler, as suggested by previous results (Kevlahan and Farge, 1997). It cannot be decided at present if the discrepancy is due to a shortcoming of the EDQNM closure, or if the hyperviscous approach breaks down when going from 2D to 3D. Checking this via direct numerical simulation in the 3D case is an interesting topic for future research.

In future work, we would also like to undertake the same kind of study with an adaptive wavelet-Galerkin solver instead of a Fourier-Galerkin solver. This will allow us to benefit at the same time from the compression and from the regularization offered by the CVS wavelet filter, and to measure the speed-up. Encouraging results concerning compression are already available in the literature on adaptive wavelet methods, see e.g. (Schneider and Vasilyev, 2010) and references therein.

Acknowledgements: We are very grateful to Nick Kingsbury for his help in developing the parallel wavelet transform, and his renewed hospitality in Trinity College, Cambridge, U.K. and in the Engineering Department, Cambridge University, U.K. We would also like to thank Claude Bardos, Uriel Frisch, Isabelle Gallagher, Thierry Paul and Laure Saint-Raymond for fruitful mathematical discussions. MF is very grateful to both Trinity College, Cambridge, and the Wissenschaftkolleg zu Berlin for hospitality.

IV.3 Remarks on Galerkin discretizations

For clarity we would like to make some remarks on the spatial discretization that we have encountered in the previous section, namely the Fourier-Galerkin discretization applied to the periodic Euler equations. The results are not new and serve only to bring attention to some points that were not completely elucidated when the contents of the two previous sections were published. We thank Professor Hammett for drawing our attention to these points.

Assuming that the spatial domain is $\Omega = \mathbb{T}^d$, and that the initial condition is smooth, the smooth solution $\mathbf{u}$ can be decomposed into its Fourier series:

$$\mathbf{u}(\mathbf{x}, t) = \sum_{\mathbf{k} \in \mathbb{Z}^d} \widehat{\mathbf{u}}_{\mathbf{k}}(t) e^{2\imath\pi \mathbf{k}\cdot\mathbf{x}},$$

where the Fourier coefficients are given by

$$\widehat{\mathbf{u}}_{\mathbf{k}}(t) = \int_{\mathbb{T}^d} \mathbf{u}(\mathbf{x}, t) e^{-2\imath\pi \mathbf{k}\cdot\mathbf{x}} \mathrm{d}\mathbf{x}.$$

The principle of the Fourier-Galerkin approach is to transform the equations for $\mathbf{u}$ into equations for $(\widehat{\mathbf{u}}_{\mathbf{k}})_{\mathbf{k}\in\mathbb{Z}^d}$, and to solve only those corresponding to $|\mathbf{k}| \leq K$, where K is a fixed integer. The remaining equations are replaced by the condition that $\widehat{\mathbf{u}}_{\mathbf{k}} = 0$ for $|\mathbf{k}| > K$. The equation satisfied by the velocity field $\mathbf{u}_K$ obtained from this system can be rewritten in a more condensed manner as:

$$\partial_t \mathbf{u}_K + \mathbf{P}_K(\mathbf{u}_K \cdot \boldsymbol{\nabla} \mathbf{u}_K) = 0 \tag{IV.3.1}$$

where $\mathbf{P}_K$ is the orthogonal projector defined in the Fourier basis by

$$\mathbf{P}_K(\mathbf{x} \to \boldsymbol{\gamma} e^{2\imath\pi \mathbf{k}\cdot\mathbf{x}})(\mathbf{x}) = \begin{cases} \left(\boldsymbol{\gamma} - \frac{\boldsymbol{\gamma}\cdot\mathbf{k}}{|\mathbf{k}|^2}\mathbf{k}\right) e^{2\imath\pi \mathbf{k}\cdot\mathbf{x}} \text{ if } |\mathbf{k}| \leq K, \\ 0 \text{ otherwise.} \end{cases}$$

where the parenthesized term in the first case incorporates the incompressibility condition. By dotting (IV.3.1) with $\mathbf{u}_K$ and integrating on $\mathbb{T}^d$ we get:

$$\mathrm{d}_t \|\mathbf{u}_K\|^2 + 2 \int_{\mathbb{T}^d} \mathbf{u}_K \cdot \mathbf{P}_K(\mathbf{u}_K \cdot \boldsymbol{\nabla} \mathbf{u}_K).$$

Now since $\mathbf{P}_K$ is orthogonal it is also self-adjoint, and by transfering it to the first $\mathbf{u}_K$ factor under the integral, we get $\mathbf{P}_K \mathbf{u}_K$. On the other hand, assuming that the initial condition satisfies $\mathbf{P}_K \mathbf{u}_K(0) = \mathbf{u}_K(0)$, this property is preserved by (IV.3.1), so that $\mathbf{P}_K \mathbf{u}_K(t) = \mathbf{u}_K$ for any t. Thanks to the fact that $\mathbf{u}_K$ is divergence free, we thus have:

$$\mathrm{d}_t \|\mathbf{u}_K\|^2 = 0,$$

which shows that the "energy" of the solution to the Galerkin-truncated system is constant.

Using the same algebra it is possible to show that for $d = 2$ the Fourier-Galerkin discretization also conserves enstrophy. Moreover, for $d = 2$, the conservation of enstrophy holds for Galerkin discretization in *any* orthogonal family. Let us prove it, denoting $(\psi_\lambda)_\Lambda$ the orthogonal family, P_Λ the orthogonal projector on the corresponding subspace of $L^2(\mathbb{T}^2)$, and ω_Λ the discrete solution. By definition

$$\partial_t \omega_\Lambda + P_\Lambda(\mathbf{u}_\Lambda \cdot \boldsymbol{\nabla} \omega_\Lambda) = 0,$$

where $\mathbf{u}_\Lambda = \boldsymbol{\nabla}^\perp \Delta^{-1} \omega_\Lambda$. Dotting this equation with ω_Λ yields:

$$\mathrm{d}_t \|\omega_\Lambda\|^2 + 2 \int_{\mathbb{T}^2} \omega_\Lambda P_\Lambda (\mathbf{u}_\Lambda \cdot \boldsymbol{\nabla} \omega_\Lambda),$$

but as above, P_Λ being self-adjoint and $\mathbf{u}_\Lambda$ being divergence free, the integral vanishes, and $\|\omega_\Lambda\|^2$ is thus conserved.

With this property in mind, the results of the two preceding sections must be interpreted with a little bit of caution. Indeed, it was shown that wavelet filtering the discrete solution in a Fourier-Galerkin discretization of the 2D incompressible Euler equations dissipates enstrophy (and similarly for the energy in the inviscid 1D Burgers equation). But this dissipation of enstrophy is tied to the fact that a Fourier projector and a wavelet projector are applied alternatively, and that these two projectors do not commute. In fact the product of two such projectors is not even a projector itself. Hence there is no contradiction with what we have just shown, namely that a genuine Galerkin discretization of the 2D Euler equations over a fixed wavelet family would conserve enstrophy.

Now if the set set Λ itself changes in time, as would be the case in an adaptive wavelet-Galerkin solver, that could imply some change in the enstrophy. It could be of interest to explore this issue, starting for example by the 1D inviscid Burgers equation as a toy-model.

Part V

Dissipation at vanishing viscosity

> Je ne vois donc pas, je l'avoue, comment on peut expliquer par la théorie, de manière satisfaisante, la résistance des fluides. Il me paraît au contraire que cette théorie, traitée et approfondie avec toute la rigueur possible, donne, au moins dans plusieurs cas, la résistance absolument nulle ; paradoxe singulier que je laisse à éclaircir aux Géomètres.
>
> Jean le Rond d'Alembert
> *Opuscules Mathématiques*, vol. 5, chap. 34 (1768).

Summary

V.1 Volume penalization

In this section, we introduce the volume penalization method for the enforcement of boundary conditions in incompressible viscous flows, which we will use in the next section. The method was first proposed by Arquis and Calgagirone (1984), in the context of porous media computations. Indeed is is based on the physical consideration that a solid wall can be approximated by a porous wall with very small porosity. For that purpose, the fluid domain Ω is first embedded into the torus $\mathbb{T}^d$, by resorting to a suitable rescaling if necessary. Then, the following equation for a velocity field $\mathbf{u}(\mathbf{x},t)$, $\mathbf{x}\in\mathbb{T}$, $t\in[0,T]$ is considered:

$$\partial_t\mathbf{u}+\mathbf{u}\cdot\boldsymbol{\nabla}\mathbf{u}=-\boldsymbol{\nabla}p+\nu\boldsymbol{\nabla}^2\mathbf{u}-\frac{1}{\eta}\chi_0\mathbf{u} \tag{V.1.1}$$

where χ_0 is the indicator function of $\mathbb{T}\backslash\Omega$, and the real $\eta>0$ is called the penalization parameter. The incompressibility condition is imposed as for the NSE. Any solution of (V.1.1) satisfies the following energy equation:

$$\frac{\mathrm{d}E}{\mathrm{d}t}=-2\nu Z-\frac{1}{\eta}\int_{\mathbb{T}^d}\chi_0\mathbf{u}^2. \tag{V.1.2}$$

Angot et al. (1999) and Carbou and Fabrie (2003) showed that the solution $\mathbf{u}_\eta$ to the initial value problem for (V.1.1) converges in L^2 to the solution $\mathbf{u}$ of the NSE with the same initial data, and more precisely, that:

$$\|\mathbf{u}_\eta-\mathbf{u}\|\le C(\nu,u_0,\chi_0)\sqrt{\eta},$$

where C does not depend on η. Note the important point that C depends on ν, and in fact the best known estimates blow up exponentially when $\nu\to 0$.

The discretization of (V.1.1) is not an easy problem because χ_0 is a discontinuous function. Various methods have been proposed, for example finite volumes (Sarthou et al., 2008) and Fourier collocation (Schneider and Farge, 2005). Here we adopt a different approach, which is however closely connected to the one of (Schneider and Farge, 2005). The idea is to approximate χ_0 by a smooth function χ prior to discretization using a Fourier-Galerkin scheme. We look for χ with the following properties:

(i) the Fourier coefficients of χ satisfy $\widehat{\chi}_{\mathbf{k}}=0$ for $|\mathbf{k}|\ge K$,

(ii) $\chi\ge 0$,

(iii) χ is close to χ_0 in $L^1(\mathbb{T}^d)$

(iv) χ is of the order of the numerical round-off error inside Ω, except perhaps close to $\partial\Omega$.

Condition (i) ensures that the product $\chi\mathbf{u}$ has only a finite number of nonzero Fourier coefficients when $\mathbf{u}$ does, which is necessary for the Galerkin discretization, while (ii) preserves the property deduced from (V.1.2) that the energy can only decrease and (iii) keeps the solution close to the one of the original problem. Condition (iv) avoids oscillations of χ inside Ω, which would imply an unacceptable perturbation of the solution.

To enforce (i), we look for χ as a convolution $\chi=\chi_0*\Phi:=\int_{\mathbb{R}^d}\chi_0(\mathbf{y}-\mathbf{x})\Phi(\mathbf{y})\mathrm{d}\mathbf{y}$. where Φ is in $L^2(\mathbb{R}^d)$, and the Fourier transform of Φ vanishes for $|\mathbf{k}|\ge K$. In the following

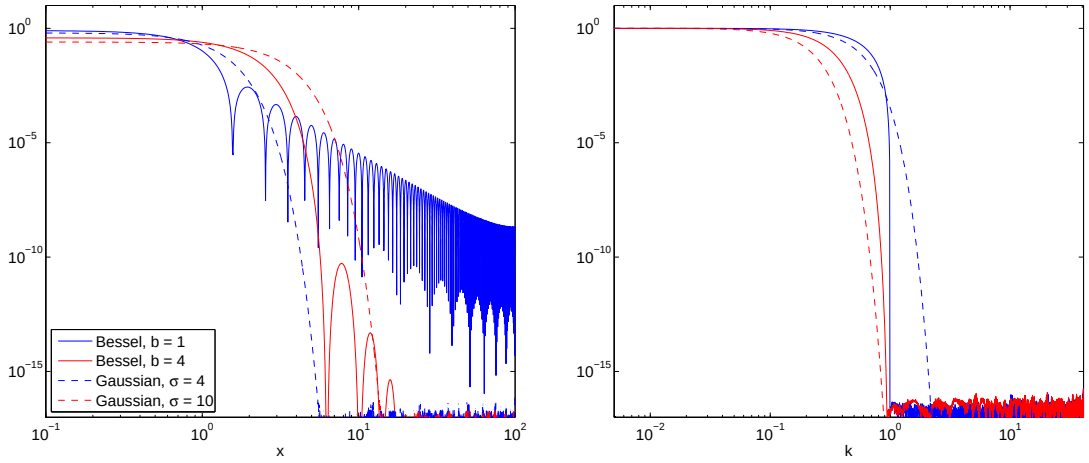


Figure V.1.1: Some putative mask smoothing filters, in physical space (left) and Fourier space (right).

we denote the Fourier transform on $\mathbb{R}^d$ by $\widehat{\Phi}(\mathbf{k})$ with parenthetized argument, whereas $\mathbf{k}$ appears as index to Fourier coefficients on $\mathbb{T}^d$. Condition (ii) implies that Φ should be positive, and (iii) means that $\widehat{\Phi}$ should be as close to 1 as possible. A first idea would be to take $\widehat{\Phi}(\mathbf{k}) = 1$ if $|\mathbf{k}| \leq K$ and 0 otherwise, but that would contradict (ii) and moreover the resulting χ would have a very bad localization which enters in conflict with condition (iv). In fact, condition (iv) imposes that Φ should decay fast at infinity. In (Ehm et al., 2004), mollifiers having optimal localization (in the sense that their variance is minimal in a certain class of functions) were constructed as follows. Let J_0 be the zeroth order Bessel function of the first kind, j_0 its smallest positive root, and define

$$\widehat{\phi}(\mathbf{k}) = \begin{cases} J_0(j_0|\mathbf{k}|) \text{ for } |\mathbf{k}| \leq 1 \\ 0 \text{ for } |\mathbf{k}| > 1 \end{cases} .$$

By construction, the function ϕ is well localized in the sense that its variance $\int_{\mathbb{R}^d} |x|^2 \phi(\mathbf{x}) d\mathbf{x}$ is as small as possible, but ϕ decays quite slowly at infinity. Indeed, $\widehat{\phi}$ is continuous on the circle $|\mathbf{k}| = 1$ but not C^1, therefore $\phi(\mathbf{x})$ decays at best like $|\mathbf{x}|^{-1.5}$. Moreover, ϕ is not positive. Both issues can be dealt with by taking

$$\Phi(\mathbf{x}) = C\phi\left(\frac{K\mathbf{x}}{2b}\right)^{2b}$$

where $b > 1$ and C is a normalization constant. Such a Φ satifies all the conditions that we have imposed, as long as b is chosen properly. In the following, we call Φ a Bessel mollifier.

To choose b we resort to a qualitative judgement, based on Fig. V.1.1, which shows the profiles of Φ and $\widehat{\Phi}$ for $b = 1$ and $b = 4$. For simplicity K has been scaled to 1. We see that $b = 4$ is a reasonnable choice, since Φ decays quickly below the round-off error. For comparison, Gaussian filters are also shown, for which $\Phi(\mathbf{k}) = \exp\left(-\frac{1}{2}\sigma^2\mathbf{k}^2\right)$. Gaussian filters can be empoyed but their localization in physical space is not as good for a given cut-off wavenumber.

In practice, the multiplication $\chi\mathbf{u}$ is computed is physical space by first reconstructing $\mathbf{u}$ and χ on a grid. To avoid aliasing errors, this grid should have at least $N = 3K$ points in

each direction. From Fig. V.1.1 (left) it appears that the Bessel mollifier with $b = 4$ roughly extends over 18 grid points. This procedure can therefore be applied only if K is sufficiently large, as is the case in the next section.

V.2 Molecular dissipation in the presence of walls

The contents of this section have been submitted for publication in Physical Review Letters, see (Nguyen van yen et al., 2010b).

V.2.1 Introduction

Flows whose energy dissipation rate is almost independent on small microscopic coupling coefficients are of central interest in fluid dynamics. For example, they are observed with Landau damping in collisionless plasmas, with reconnection in magnetized fluids or plasmas, and with fully developed turbulence in incompressible neutral fluids. Despite its ubiquity, the phenomenon is still often called "anomalous dissipation". One way of approaching the problem has been to look for generic structures whose existence in the flow is sufficient to imply nonzero energy dissipation. Shocks in compressible flows are a relatively well understood example of such energy dissipating structures, but their well established theory (Lax, 1973) is still an exception.

In this Letter, we focus on two-dimensional incompressible flows in contact with solid boundaries, for which the understanding is, comparatively, at an embryonic stage. Working with the inviscid potential flow equations, later generalized as the Euler equations, d'Alembert came up in 1768 with the famous paradox bearing his name (le Rond d'Alembert, 1768) that the flow would exert no drag force onto solid obstacles. The Navier-Stokes equations (NSE) were then derived during the 19th century by including molecular friction effects. It was gradually realized that the paradox came from the singular nature of the vanishing viscosity limit, mostly due to the no-slip boundary condition imposed along the solid boundary. In 1904, Prandtl (Prandtl, 1904) resolved the paradox in the very special case of flows in which the effects of viscosity are confined to a boundary layer of thickness proportional to $\mathrm{Re}^{-1/2}$ in the neighborhood of the wall, where Re is the Reynolds number. He was able to compute a drag coefficient, and hence also an energy dissipation rate, which are both proportional to $\mathrm{Re}^{-1/2}$ in the vanishing viscosity limit. Prandtl's theory does not apply when the boundary layer detaches from the wall, because the Euler equations can then no more be used to describe the flow, even far from the solid boundary. In 1984 Kato (Kato, 1984) proved mathematically that, in the vanishing viscosity limit, the energy dissipation rate tends to zero if and only if the solution of the NSE converges to the solution of the Euler equations with the same initial data. He also proved in the same paper that, for dissipation to occur anywhere in the flow at any time, at least some dissipation had to occur within a very thin boundary layer of thickness proportional to Re^{-1} in the neighborhood of the wall.

V.2.2 Model and numerical method

The elementary event that we focus on in this Letter consists in the collision of a vorticity dipole into a solid wall, a classical object of experimental studies (see e.g. (Flor and van Heijst, 1994)), which has also been used for benchmarking several numerical methods (Orlandi, 1990; Clercx and Bruneau, 2006; Keetels et al., 2007; Kramer et al., 2007). A first specific study of energy dissipation appeared in (Clercx and van Heijst, 2002), where the NSE with no-slip boundary conditions were approximated using Galerkin discretization over a basis of Chebychev polynomials, both in the wall parallel and wall normal directions. The numerical solutions satisfied the no-slip boundary conditions to machine accuracy but it was later found out (Clercx and Bruneau, 2006) that they were not converged beyond the collision time.

In contrast, we have focused on using a numerical scheme which resolves scales at least as fine as Re^{-1}, to cope with Kato's theorem. This stringent requirement on resolution has not been enforced, to our knowledge, in any previous numerical experiments at similar Reynolds numbers. The best way that we have found to comply with it was to work with a numerical model known as volume penalization. The counterpart, as we shall show in detail below, is that the no-slip boundary conditions are replaced by Navier boundary conditions with a slip length tending to zero when $\mathrm{Re} \to \infty$.

The initial value problem for the NSE with volume penalization (PNSE) can be written as follows (Arquis and Calgagirone, 1984):

$$\begin{cases} \partial_t \mathbf{u} + (\mathbf{u} \cdot \boldsymbol{\nabla})\, \mathbf{u} = -\boldsymbol{\nabla} p + \nu \boldsymbol{\nabla}^2 \mathbf{u} - \frac{1}{\eta} \chi \mathbf{u} \\ \boldsymbol{\nabla} \cdot \mathbf{u} = 0, \qquad \mathbf{u}(\cdot, t=0) = \mathbf{u}_0, \end{cases} \tag{V.2.1}$$

where $\mathbf{u}(\mathbf{x}, t)$ and $p(\mathbf{x}, t)$ are respectively the velocity and pressure fields defined for $\mathbf{x}$ in the unit torus $\mathbb{T}^2 = (\mathbb{R}/\mathbb{Z})^2$ and $t \geq 0$, ν is the kinematic viscosity, χ is a mask function which equals 0 in the fluid part Ω of $\mathbb{T}^2$ and 1 elsewhere, and η is the penalization parameter which needs to be sufficiently small in order to impose that the velocity almost vanishes outside Ω. In the following we shall take $\Omega =]0.05, 0.95[\times \mathbb{T}$, thus modeling a periodic channel. To be consistent with previous studies (Orlandi, 1990; Keetels et al., 2007), the Reynolds number Re is defined as $\mathrm{Re} = \frac{UL}{\nu}$, where L is the half-width of the channel ($L = 0.45$ in our units) and U is the initial rms velocity. Note that both the NSE and PNSE are well posed in two dimensions, see (Foiaš et al., 2001).

Our main quantity of interest will be the kinetic energy $e(\mathbf{x}, t) = \frac{1}{2}|\mathbf{u}(\mathbf{x}, t)|^2$ which evolves pointwise according to

$$\partial_t e + \boldsymbol{\nabla} \cdot ((e + p)\mathbf{u}) = -\nu |\boldsymbol{\nabla} \mathbf{u}|^2 - \frac{1}{\eta} \chi \mathbf{u}^2 + \nu \Delta e, \tag{V.2.2}$$

where the term $\nu |\boldsymbol{\nabla} \mathbf{u}|^2$ will be called local energy dissipation rate, as in (Kato, 1984). The total energy $E = \int_{\mathbb{T}^2} e$, satisfies

$$\frac{\mathrm{d}E}{\mathrm{d}t} + 2\nu Z + \frac{1}{\eta} \int_{\mathbb{T}^2} \chi \mathbf{u}^2 = 0, \tag{V.2.3}$$

where $Z = \frac{1}{2} \int_{\mathbb{T}^2} |\boldsymbol{\nabla} \mathbf{u}|^2$ is the enstrophy.

For $\eta \to \infty$, the PNSE reduces to the NSE, and the solution then satisfies the following scaling for energy dissipation (Beale and Majda, 1981), which we call wall-less scaling:

$$E(t_2) - E(t_1) \propto \mathrm{Re}^{-1}. \tag{V.2.4}$$

Re	985	1970	3940	7880
N	2048	4096	8192	16384
η	$4 \cdot 10^{-5}$	$2 \cdot 10^{-5}$	10^{-5}	$0.5 \cdot 10^{-5}$

Table V.2.1: Parameters of numerical experiments.

For $\eta \to 0$ and fixed Re, it has been proven mathematically (Angot et al., 1999; Carbou and Fabrie, 2003) that the PNSE approaches the NSE in Ω with no-slip boundary conditions. This convergence has been checked numerically (Keetels et al., 2007) for the dipole-wall collision at $\mathrm{Re} = 1000$.

Initial data that lead to a dipole-wall collision were introduced in (Orlandi, 1990). It is conveniently defined by its vorticity field $\omega_0 = \boldsymbol{\nabla} \times \mathbf{u}_0$, for which the analytical expression is

$$\omega_0(\mathbf{x}) = \omega_e \sum_{i=0}^{1} (-1)^i \left(1 - \frac{|\mathbf{x} - \mathbf{a}_i|^2}{r^2}\right) e^{-\frac{|\mathbf{x}-\mathbf{a}_i|^2}{r^2}}, \tag{V.2.5}$$

where $\mathbf{a}_0 = (0.445, 0.5)$, $\mathbf{a}_1 = (0.555, 0.5)$, $r = 0.045$ and $\omega_e = 299.5$, so that $U = 0.443$. Note that in addition to the numerical scheme, two things differ from (Clercx and van Heijst, 2002): the initial distance between the two vortices is slightly larger, and the domain is a channel instead of a square container.

To benefit from optimal efficiency, we work with the vorticity formulation of the PNSE in divergence form. But for our purpose, it is important to use a spatial discretization which satisfies the energy balance equations (V.2.2) and (V.2.3) to round-off accuracy. Therefore, we approximate χ by a positive function whose Fourier coefficients vanish for wave vectors having their modulus higher than a given cut-off wavenumber K. The mask is thus mollified over a scale proportional to K^{-1}, which corresponds in practice to about 20 grid points. The mollified equations are then discretized using a classical pseudo-spectral scheme (Schneider and Farge, 2005). To avoid aliasing errors, the grid over which the products are evaluated has $N = 3K$ points in each direction. For time discretization we have used an explicit third order low storage Runge-Kutta scheme, with an integrating factor method (Trefethen, 2000, page 111) to take into account the viscous term. The duration δt of each time step is adjusted to ensure stability. The parameters of the four numerical experiments whose results are reported in this Letter are listed in Table V.2.1. Our choices of resolution are dictated by the requirement to resolve scales as fine as Re^{-1}, and we take η as small as possible but such that it does not constrain δt more than the CFL condition does.

V.2.3 Results

Fig. V.2.1 depicts the vorticity field in the subdomain $[0.708, 0.962] \times [0.5, 0.754]$, at $t = 0.36$, 0.4, 0.45 and 0.495, for $\mathrm{Re} = 7880$, computed with $N^2 = 16384^2$. Since the vorticity field is antisymmetric with respect to $x_2 = 0.5$, we need only describe its evolution in the upper half of the domain. We observe that, as the positive vortex of the impinging dipole propagates rightwards, a sheet of negative vorticity is created on the wall. The sheet starts to roll-up after having reached an intensity of about 17 times the initial vorticity maximum, and a strong vortex is hence produced. The produced vortex then pairs with the impinging vortex to form a secondary dipole which bounces back from the wall.

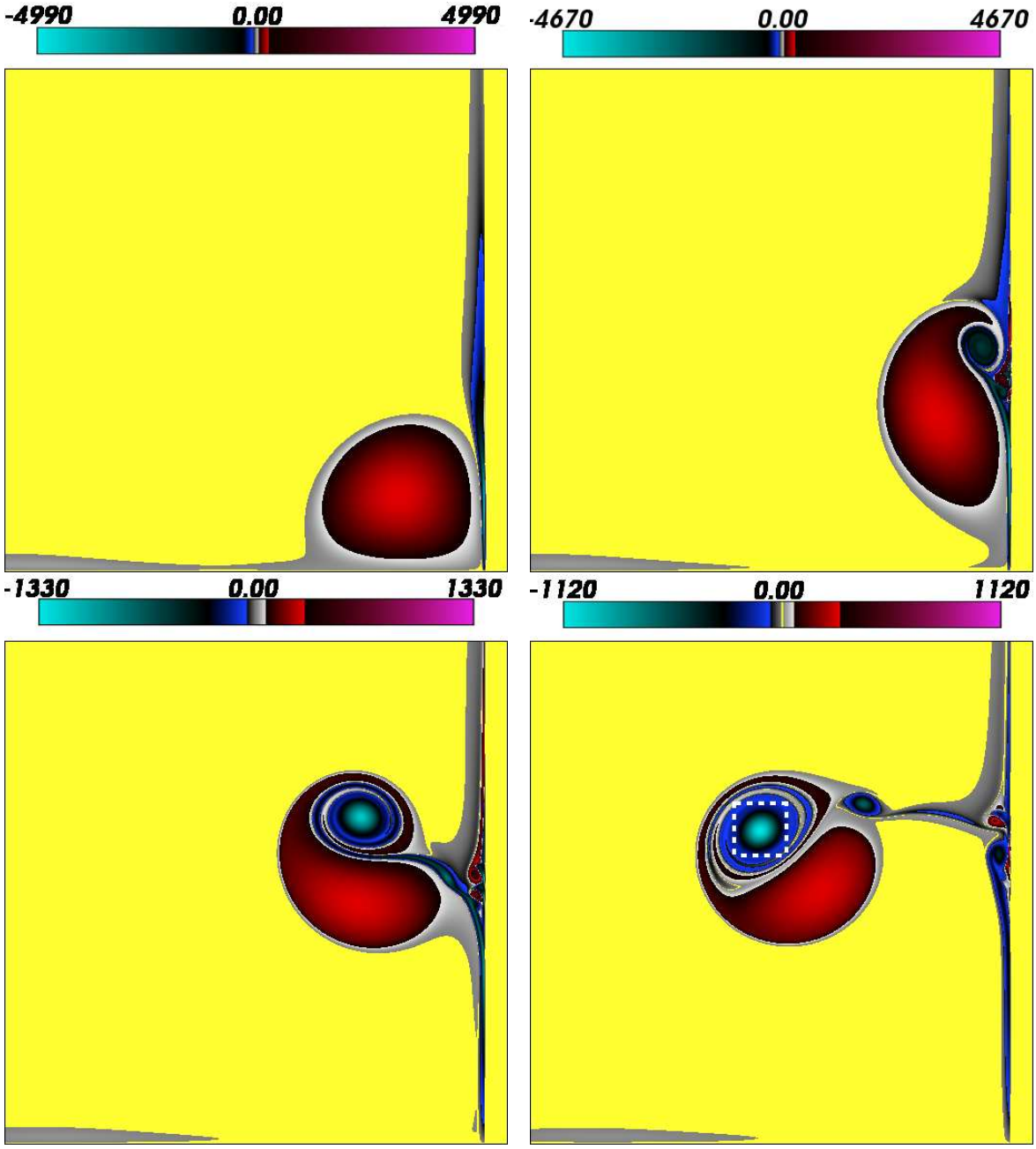


Figure V.2.1: Vorticity field in the subdomain $[0.708, 0.962] \times [0.5, 0.754]$ at $t = 0.36$, 0.4, 0.45 and 0.495 (left to right) for $\mathrm{Re} = 7880$. The white dotted box at $t = 0.495$ frames region B (see text). Black pixels correspond to $\omega = \pm 300$ in all four pictures.

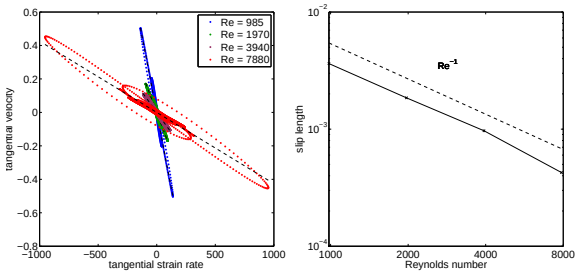


Figure V.2.2: Left: tangential velocity u_2 versus tangential strain rate $\partial_1 u_2$. The dashed lines are least squares linear fits. Right: slip length α defined by (V.2.6) as a function of Re.

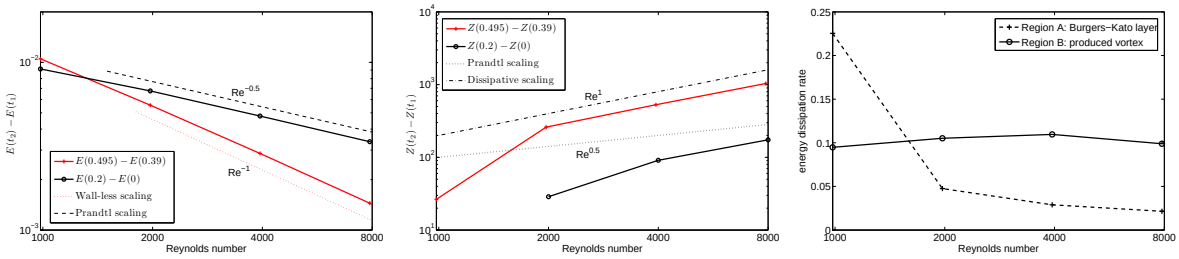


Figure V.2.3: Left and middle: energy dissipation (respectively enstrophy increase) as a function of Re for $t \in [0, 0.2]$ and $t \in [0.39, 0.495]$. Right: instantaneous energy dissipation as a function of Re at $t = 0.495$ in regions A and B (see text).

Before presenting further results we would like to dwell on the issue of boundary conditions. We concentrate on the wall facing the impinging dipole and, to avoid grid effects, we define its position as the isoline $\chi = 0.02$ along which we interpolate the fields. The value 0.02 is somewhat arbitrary, but its order of magnitude can be justified by balancing the penalization term $\frac{1}{\eta}\chi\mathbf{u}$ with the viscous term $\nu\Delta\mathbf{u}$ in the PNSE. With this definition, the wall-normal velocity u_1 is smaller than 10^{-3} independently of Re, so that to a good accuracy there is no flow through the boundary, as expected. But the wall-parallel velocity u_2 reaches values of order 0.1, to be compared to the initial rms velocity $U = 0.443$. Plotting u_2 as a function of $\partial_1 u_2$ for $t = 0.495$ along the wall (Fig. V.2.2, left) reveals a relationship of the form:

$$u_2 + \alpha(\mathrm{Re}, \eta, N)\partial_1 u_2 \simeq 0 \qquad \text{(V.2.6)}$$

with a correlation coefficient higher than 0.98 in all cases. Hence, restricted to the domain $\chi < 0.02$, $\mathbf{u}$ almost satisfies the NSE with Navier boundary conditions and with a slip length α. Values of α obtained from least squares fits are shown in Fig. V.2.2 (right) as a function of Re, where we observe that α is approximately proportional to Re^{-1}.

We now come to our main results concerning energy dissipation. In Fig. V.2.3 (left), $E(0.2)$–$E(0)$ and $E(0.495)$–$E(0.39)$ are plotted versus Re. During the time interval $[0, 0.2]$, the effect of the boundary is still negligible and the wall-less scaling (V.2.4) is recovered. During $[0.39, 0.495]$, we find that the Prandtl scaling $E(0.495)$–$E(0.39) \propto \mathrm{Re}^{-1/2}$ nearly holds. To give more weight to the effects of the boundary, we now compute the increase in enstrophy between $t = 0$ and, respectively, $t = 0.39$ and $t = 0.495$ (Fig. V.2.3, middle). We observe in both cases that $Z(t)$–$Z(0) \propto \mathrm{Re}$, suggesting that the global energy dissipation rate $2Z\mathrm{Re}^{-1}$ does not go to zero in the vanishing viscosity limit. This can not be seen on the curve of $E(0.495)$–$E(0.39)$ because the effect of the bulk dissipation is still dominant at the Reynolds numbers considered here. However, since $E(0.495)$–$E(0.39) \approx 2\nu \int_{0.39}^{0.495} Z(t)\mathrm{d}t$,

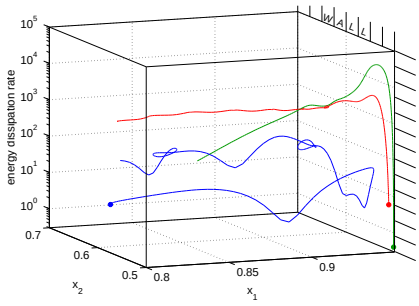


Figure V.2.4: Energy dissipation rate $\nu|\nabla\mathbf{u}(x_1(t), x_2(t), t)|^2$ versus particle position $(x_1(t), x_2(t))$ for $t \in [0.3, 0.495]$ along three selected Lagrangian trajectories, at $\mathrm{Re} = 3940$. The circles indicate the positions at $t = 0.3$.

we expect that at higher Re, the dissipative scaling $E(0.495)$–$E(0.39) \propto \mathrm{Re}^0$ holds. To provide more conclusive evidence of this, we isolate two regions where energy is actually dissipated: region A, a vertical slab of width $10N^{-1}$ inside the fluid domain along the isoline $\chi = 0.02$, and region B, a square box of side length 0.025 around the center of the main structure that has detached from the wall at $t = 0.495$ (dotted box in Fig. V.2.1, right). The energy dissipation rate, as defined by (V.2.2), is integrated respectively over the domain A or B and plotted versus Re in Fig. V.2.3 (right). It can be seen that in both cases the dependence on Re becomes weak for $\mathrm{Re} \gtrsim 2000$.

The understanding of these results is facilitated by following some Lagrangian particles in the flow. Particles were advected using a bicubic interpolation of the velocity field and third order Runge-Kutta time integration. We now describe the evolution of the energy dissipation rate $\nu|\boldsymbol{\nabla}\mathbf{u}(\mathbf{x}, t)|^2$ along three selected trajectories for $\mathrm{Re} = 3940$ (Fig. V.2.4). The first striking feature is that it displays a strong maximum for two particles which start from the wall (green and red curves), occurring when they are still in region A. In contrast, there is little dissipation along the third trajectory (blue curve), which starts away from the wall and never enters region A. At later times, energy dissipation goes back to much smaller values for one of the trajectories that approached the wall (red curve), while it is still one order of magnitude larger for the other one (green curve), because the particle is trapped inside the strong vortex produced at the wall (corresponding to region B at $t = 0.495$).

V.2.4 Conclusion

In summary, we have shown that the dipole-wall collision, a well known elementary event in two-dimensional (2D) incompressible flows, is a good candidate for having nonzero energy dissipation in the vanishing viscosity limit. We have imposed Navier boundary conditions with a slip length roughly proportional to Re^{-1}. Note that the convergence of solutions to the NSE under Navier boundary conditions with a slip length varying as $\alpha = \mathrm{Re}^{-\gamma}$ to solutions of the Euler equations has been proven mathematically for $\gamma < 0.5$ (Xin, 2010), but that the case $\gamma \geq 0.5$ is still open. Our results suggest that convergence does not hold for $\gamma \approx 1$, due to the same scenario that is implied by Kato's theorem, namely, that energy keeps being dissipated in a very thin boundary layer of thickness proportionnal to Re^{-1}, orders of magnitude thinner than the Prandtl boundary layer, and which was also predicted by Burgers (Burgers, 1923) to explain the observed behavior of the drag force. We propose to call it the Burgers-Kato boundary layer. In addition, we have shown that

energy dissipation continues within spiral structures detaching from the boundary, and is likely to occur only along Lagrangian trajectories that "crash" into the wall when $\mathrm{Re} \to \infty$. One should be aware that the perfect fluid model by itself is not sufficient to reproduce this important physical effect. We argue further that energy dissipating structures may be common in 2D wall-bounded flows, with slip or no-slip boundary conditions, as long as the slip length is small enough. They could perhaps be observed in soap film experiments or in oceanic flows. These structures cannot be described in the framework of classical statistical theories of 2D flows (Onsager, 1949; Kraichnan, 1967; Batchelor, 1969). Understanding the statistical properties of flows containing energy dissipating structures is thus an open question of great relevance to a theory of 2D wall-bounded turbulence (Eyink, 2008).

We heartily thank Claude Bardos whose guidance was essential for the completion of this work. We are also grateful to Gregory Eyink, Dmitry Kolomenskiy, Anna Mazzucato, Helena Nussenzveig-Lopes, Qiu Yanqi, Xin Zhouping and Yin Qizheng for fruitful discussions. MF and RNVY are thankful to the Wissenschaftskolleg zu Berlin for its hospitality. We thank IDRIS-CNRS for providing computing facilities. This work was carried out within the framework of the European Fusion Development Agreement and the French Research Federation for Fusion Studies. It is supported by the European Communities under the contract of Association between Euratom and CEA.

V.3 Turbulent dissipation in 2D homogeneous turbulence

The contents of this section have been submitted for publication in Physica D, see (Nguyen van yen et al., 2010c).

V.3.1 Introduction

Fluid flows are said to be in the fully developed turbulent regime when they are highly disordered and dominated by nonlinear effects due to inertia, which occurs when their Reynolds number Re is large. Competing tendencies to form and to dislocate coherent structures then create motion over a wide range of spatial and temporal scales. In the special case of periodic two dimensional (2D) decaying turbulence, on which we focus in this paper, merging processes tend to win the game and the average size of structures increases. Kraichnan has shown that this basic phenomenology could be understood with the help of a dual cascade paradigm, in which enstrophy and energy are sent respectively towards scales finer and coarser than the integral scale of the flow ((Kraichnan, 1967), hereafter K67). Within the simplified setting where energy and enstrophy are injected around a characteristic wavenumber k_I, he established that the enstrophy cascade would occur between k_I and a cut-off wavenumber $k_D \propto \mathrm{Re}^{\frac{1}{2}}$, and that for large Re the energy spectrum would become Re-independent and approach the universal scaling $E(k) \propto k^{-3}$. Such an energy spectrum is associated to a k-independent enstrophy flux from small wavenumber modes to large wavenumber modes. The range of k between k_I and k_D is called the inertial range, while wavenumbers k larger than k_D constitute the molecular dissipation range. The relevance of the enstrophy cascade for the phenomenological description of decaying 2D turbulence is supported by numerical experiments (Batchelor, 1969; Chasnov, 1997).

The K67 theory is completely statistical, in the sense that it does not claim to predict anything about the time evolution of a flow, but invokes only generic features of all turbulent flows satisfying certain hypotheses, that is, of all elements of a large ensemble. It may be criticized on the grounds that realistic flows, such as a jet or the wake behind an obstacle, typically do not satisfy its underlying assumptions. Even in the restricted setting where the theory is approximately valid, one may be interested in the history of the flow and not only in properties of a global attractor. Another shortcoming of the K67 theory is that it is inconsistent with a mathematical theorem. Indeed, as pointed out in (Dritschel et al., 2007), one of its essential hypotheses is that the enstrophy dissipation rate does not go to zero in the limit $\mathrm{Re} \to \infty$. But it is well known (Golovkin, 1966) that in the limit $\mathrm{Re} \to \infty$, smooth solutions of the 2D Navier-Stokes equations (NSE) with smooth forcing converge in the enstrophy norm to smooth solutions of the 2D Euler equations (corresponding to $\mathrm{Re} = \infty$), which have constant enstrophy. This paradox, already predicted using a closure model (Pouquet et al., 1975) and pointed out in (Vallis, 1985), was recently revisited by high resolution direct numerical simulation (DNS) of the 2D NSE (Dmitruk and Montgomery, 2005; Tran and Dritschel, 2006). It was shown that even though the enstrophy dissipation rate is bounded by $C(t)\mathrm{Re}^{-1}$ when $\mathrm{Re} \to \infty$ (a mathematical theorem!), the decay with Re that can be observed in practice is only logarithmic, because $C(t)$ increases extremely fast. The implications of this result on the statistical theory of 2D turbulence were discussed in (Dritschel et al., 2007) (see also (Davidson, 2008) for a different viewpoint). It was argued

that no Reynolds independent quantity was available to characterize dissipation, and that the hope of a Reynolds independent description of 2D turbulence had to be given up.

This conclusion is somewhat disappointing, given the already mentioned convergence of smooth solutions to the 2D Navier-Stokes equations in a periodic domain towards solutions to the 2D Euler equations (Golovkin, 1966). Convergence occurs at a relatively fast rate $O(\mathrm{Re}^{-1})$ in the energy norm, so that Re-dependent effects can be seen as perturbations on top of the inviscid behavior. In the presence of walls imposing no-slip boundary conditions, the 2D Euler equations are also well posed, but their solutions behave very differently from those of the Navier-Stokes solutions even at large Re, and the problem is much more delicate, see e.g. (Bardos and Titi, 2007) for a discussion. The 3D case also presents formidable difficulties. But in the 2D wall-less case, on which we focus in this paper, the initial value problem for the incompressible Euler equations is a solid foundation from which features of 2D turbulence should be deduced. In fact, the use of inviscid equations, sometimes along with an ad-hoc regularization mechanism, is widespread in numerical models, for example in geophysical fluid dynamics (Pedlosky, 1987). Hence the main obstacle may lie in our inability to ask the right questions.

In the search for a reduced description of hydrodynamic turbulence, the detailed study of reference solutions obtained from well validated numerical methods remains an important ingredient. Therefore, although our long term goal is to build a reduced description of 2D turbulence based on the wavelet representation, we limit ourselves in the present contribution to the analysis of fully developed decaying 2D turbulent flows starting from first principles. From this a priori analysis, we are able to study several wavelet-based models and to establish their essential features. We hope that the resulting picture of 2D turbulence will provide a way towards more predictive approaches.

In the first section, we recall the general notion of incomplete statistical equilibrium, and explain how it was studied by Kraichnan and others in the context of turbulence. We then describe the mathematical setting and the numerical tools that we have chosen for our specific study of 2D decaying turbulence. After a brief reminder on wavelet theory, we present a set of results about the statistics of 2D turbulence. We then recall the coherent vorticity extraction procedure, which aims to split the degrees of freedom of a turbulent flow into a noisy part and a deterministic part. Subsequently, we study the transfer of enstrophy between these two parts, both from a Fourier view point and from a wavelet view point. In the last section, we study the retroaction of the noisy vorticity component onto the rest of the flow. Finally, we discuss the overall results and draw some conclusions.

V.3.2 Conditional statistical modelling

Predictable quantities in turbulent flows can only be defined in a statistical sense, as was already remarked by Burgers (Burgers, 1923) and Taylor (Taylor, 1935). In the 2D case, absolute statistical equilibria assuming only conservation of energy and enstrophy were derived by Kraichnan (Kraichnan, 1967) and observed numerically by Basdevant and Sadourny (Basdevant and Sadourny, 1975) after a long time in simulations of the Galerkin truncated 2D Euler equations. Solutions to the full 2D Euler equations never reach these absolute equilibria, due to their lack of truncation at fine scales, which allows enstrophy to escape towards infinitely high wavenumber regions of the spectrum. Instead, these solutions tend to follow the phenomenology associated to the already introduced K67 theory, but no rigorous

statistical ensemble has been constructed to explain this observation. In a stationary setting, one may try to derive the K67 theory by studying the invariant measure of the dynamical system associated to the Euler equations (see e.g. (Ruelle and Takens, 1971) and related papers). But the fact that the same phenomenology can be observed in freely decaying flows suggests that the stationarity hypothesis is superfluous. In any case, as argued by many authors, even a detailed understanding of the invariant measure may yield little useful information about the behavior of a particular solution, because of the slow and non-uniform sampling of the attractor (see e.g. (Bradshaw, 1994)). However, if one considers a flow that is evolving in time, we already know that it is a solution to the Navier-Stokes equation, which are completely deterministic, and there is no objective way to introduce a statistical ensemble (see (Bricmont, 1996) for more along this line).

To understand the difficulty, let us first consider the relatively easier case of classical kinetic theory (Balian, 2006). In the fluid approximation, many-body systems are considered to be in a state of local thermodynamic equilibrium and are described by a few macroscopic fields, like velocity and temperature. The local Maxwellian distribution of particle velocities can be recovered by maximizing entropy with the constraint that the macroscopic fields takes their known value in each point. Therewith the statistical ensemble containing possible realizations of the microscopic degrees of freedom is defined as a direct product of local ensembles corresponding to each fluid particle. As long as the hypothesis of local thermodynamic equilibrium holds, the equations governing the evolution of macroscopic quantities depend on the particular realization of the microscopic motion only through a stochastic forcing term, which is neglected in practice. There is thus a perfect separation between microscopic and macroscopic motions. If one changes, by thought experiment, the sign of the velocity of a single molecule in such a fluid, the microscopic motion soon takes an entirely different trajectory but there is no measurable influence on the macroscopic velocity field. The only signatures of the microscopic properties of the system are the friction terms in the fluid equations. When energy is transferred to microscopic motion, it is dissipated, or "thrown away" according to the Latin etymology, evocative of an irreversible loss of information.

Now due to the apparent disorder of the macroscopic velocity field itself, it seems desirable to refine the separation by distinguishing between two classes of macroscopic degrees of freedom, those which we want to predict, and those which are to be replaced by statistical distributions or, in other words, "thrown away" to join their microscopic comrades. Kraichnan (Kraichnan, 1974) gave an intuitive argument in favor of such an intermediate description over a fully statistical one. Comparing the characteristic time needed to spread energy in space within a given scale, and the characteristic time needed to transfer energy between scales, he noted that they were of the same order of magnitude. In other words, active flow elements will not completely forget their location in space before they start getting distorted, and, conversely, will not completely forget their shape before they move to another location in space. These active flow elements may be amenable to a deterministic treatment, while others call for a statistical description. Kraichnan and Chen (Kraichnan and Chen, 1989) went further, asserting that "Turbulence [...] is an interplay of order and disorder, associated with strong departure from absolute statistical equilibrium". They noted that advanced statistical models such as the direct interaction approximation (DIA) (Kraichnan, 1958) could not faithfully describe the dynamics of very simple dissipative systems. To develop a better description, they proposed the concept of "constrained decimation", which consists in splitting the degrees of freedom of the flow into two parts:

- those that are modelled statistically, which collectively constitute the "dissipated" flow,
- those whose time evolution is described by an initial value problem, which we call "explicit", following (Kraichnan and Chen, 1989).

The notion of "degrees of freedom" is used here in a loose sense that will be made more rigorous later on. What is important is that the split is not necessarily static, but can depend on the instantaneous flow, or even on its history. Indeed, this approach, to which we refer as conditional statistical modelling, is time-dependent by construction, and in it no assumptions are made about the stationarity of the solution. To each definition of the explicit flow corresponds a notion of dissipation, or transfer of energy (or enstrophy) from explicit to dissipated degrees of freedom. The same idea was later elaborated by several authors, see the review by Kraichnan (Kraichnan, 1988) and references therein. Due to the nonlinear term, it is likely that, contrary to the microscopic degrees of freedom, the dissipated flow retroacts strongly onto the explicit flow and cannot be neglected or even treated perturbatively. Hence the frontier between the two can a priori be chosen subjectively, but the strength and qualitative properties of the retroaction of the dissipated flow on the explicit flow determines a posteriori the practical relevance of the split.

In fact the idea of extending the notion of dissipation to macroscopic degrees of freedom in turbulent flows goes back at least to Richardson and Gaunt (Richardson and Gaunt, 1930). The classes they had in mind can be roughly termed "coarse scale motion" and "fine scale motion", and therefore the dissipation so defined was formally equivalent to an eddy viscosity, for which Prandtl had given a formula five years earlier (Prandtl, 1925), and which has remained the dominant paradigm ever since. More advanced methods have also been developed based on the same splitting between coarse and fine scales, for example large eddy simulation (LES) (Deardorff, 1970) and nonlinear Galerkin (Dubois et al., 1993). In the last ten years, more advanced statistical physics concepts have been developed in this context (Ellis et al., 2000), and applied for example to describe multiscale atmospheric flows (Turkington et al., 2001) or solutions of dispersive nonlinear wave equations (Eisner and Turkington, 2006). But from equilibrium statistical physics we know that the correctness of the predictions depends highly on the choice of the right variables to describe the system. It is therefore an important research topic to explore different concepts of dissipation in turbulent flows. The goal of this paper is to explore two alternatives to the coarse scale/fine scale split. They are both based on the wavelet representation of the vorticity field, which has been advocated since the late 1980s as an improvement over the more classical Fourier representation (Farge and Rabreau, 1988a; Meneveau, 1991; Farge et al., 1992a). Multiscale expansions can be seen as a natural follow up on the wavenumber band expansions studied by Kraichnan in (Kraichnan, 1974), with the essential improvement of maintaining some space locality in the description. Interestingly, they were also encouraged by Jaynes (see the remarks at the end of (Jaynes, 1985)).

On the long list of remaining issues, there is the matter of how to choose the statistical model for the dissipated flow. By analogy with the Gibbs entropy of equilibrium statistical mechanics, the entropy S_F of the flow can be defined as the Shannon entropy (Shannon, 1948) computed from the statistical distribution of the dissipated degrees of freedom (see section V.3.7.1). Following the ideas of Jaynes (Jaynes, 1957), the distribution that should be chosen to make the most unbiased predictions about the future evolution of the flow is the one that maximizes S_F. We should keep in mind that the goal of entropy maximization is to avoid biased results, and not to optimize the predictions of a model. Maximal entropy

predictions are not necessarily good, in fact they may even be completely worthless if the split between explicit and dissipated degrees of freedom has been ill chosen.

In the following, we want to consider a single solution to the initial value problem for the Euler equation, and see how these ideas fit together to describe it.

V.3.3 Mathematical framework and numerical method

In any modelling effort there is, a priori, a phase of analysis, where information is gathered either from experiments or from a well established underlying model, and, a posteriori, a phase of synthesis, where the new model is validated. Although our long term goal is to help improve existing statistical models and computational methods, we focus here on the a priori analysis stage, having in mind the word of wisdom of Meneveau & Katz (Meneveau and Katz, 2000): "a posteriori tests typically do not provide much insight into the detailed physics of models and the reasons that they do or do not work". Therefore we set out to study numerical solutions to the 2D Navier-Stokes equations, an approach commonly known as direct numerical simulation (DNS), and since we are interested in time dependent properties of the flow and not only in stationary statistics, we work on the initial value problem:

$$\begin{cases} \partial_t \omega + (\mathbf{u}\cdot\boldsymbol{\nabla})\omega + \nu(-\Delta)^\alpha \omega = 0 \\ \boldsymbol{\nabla}\cdot\mathbf{u} = 0, \;\; \omega = \boldsymbol{\nabla}\times\mathbf{u}, \;\; \omega(\cdot,t_0) = \omega_0 \end{cases} \tag{V.3.1}$$

where the unknown vorticity $\omega(\mathbf{x},t)$ is a scalar field defined on $\mathbb{T}^2\times[t_0,t_1]$, $\mathbb{T}^2 = (\mathbb{R}/\mathbb{Z})^2$ is the unit torus, t_0 and t_1 are respectively the initial and final time, $\mathbf{u}$ is the velocity field, ν is the fluid viscosity, $\alpha\in\mathbb{N}^*$ and ω_0 is a smooth initial vorticity field. Classical results guarantee existence and uniqueness of a solution $\omega(\mathbf{x},t)$ to problem (V.3.1) in a suitable function space, see (Ladyzhenskaya, 1963). These results extend to the case of the incompressible Euler equations, corresponding to $\nu = 0$ in (V.3.1), see (Bardos and Titi, 2007). When $\alpha = 1$, (V.3.1) are called the Navier-Stokes equations (NSE), and when $\alpha > 1$ they are called the hyperdissipative NSE (HNSE) (Basdevant et al., 1981). Hyperdissipation is an ad-hoc regularization mechanisms for the incompressible Euler equations, which aims to approach the inviscid dynamics better than classical dissipation ($\alpha = 1$), for given computational resources. Although its widespread use has been criticized (Frisch et al., 2008), we have made a good case for it in the restricted context of the 2D NSE with periodic boundary conditions in a previous paper (Nguyen van yen et al., 2009). Here, we shall consider only two choices, namely $\alpha = 1$ and $\alpha = 2$.

In the following, $\langle\cdot \mid \cdot\rangle$ is the classical scalar product in $L^2(\mathbb{T}^2)$, the space of square-integrable functions on $\mathbb{T}^2$. $\|\cdot\|$ is the associated norm, and the Fourier transform of a field f on $\mathbb{T}^2$ is defined by

$$\hat{f}[\mathbf{k}] = \int_{\mathbb{T}^2} f(\mathbf{x}) \exp\left(-2\imath\pi\mathbf{k}\cdot\mathbf{x}\right) \mathrm{d}\mathbf{x},$$

where $\imath = \sqrt{-1}$ and $\mathbf{k}\in\mathbb{Z}^2$. When $\nu = 0$ and when the initial data are smooth, an infinity of integral quantities are preserved by the flow, among which the energy $E = \frac{1}{2}\|\mathbf{u}\|^2$ and the enstrophy $Z = \frac{1}{2}\|\omega\|^2$. When $\nu > 0$, Z decays in time according to:

$$\frac{\mathrm{d}Z}{\mathrm{d}t} + 2\nu P = 0 \tag{V.3.2}$$

where $P = \frac{1}{2}\|\boldsymbol{\nabla}\omega\|^2$ is called the palinstrophy.

To use as initial condition we construct a random vorticity field ω_r by letting

$$\hat{\omega}_r[\mathbf{k}] = \begin{cases} \frac{|\mathbf{k}|}{6} e^{\iota\theta_{\mathbf{k}}} \text{ if } |\mathbf{k}| \leq 6 \\ (\frac{|\mathbf{k}|}{6})^{-1} e^{\iota\theta_{\mathbf{k}}} \text{ if } 6 < |\mathbf{k}| \leq 42 \\ 0 \text{ otherwise} \end{cases} \tag{V.3.3}$$

where the $\theta_{\mathbf{k}}$ are pseudo-random numbers drawn uniformly in $[0, 2\pi[$. If $E[k]$ denotes the energy spectrum of the flow

$$E[k] = \frac{1}{2} \sum_{k \leq |\mathbf{k}| < k+1} |\hat{\mathbf{u}}[\mathbf{k}]|^2$$

we have for ω_r

$$E_r[k] \simeq \begin{cases} \frac{\pi k}{36} \text{ if } |k| \leq 6 \\ 36\pi k^{-3} \text{ if } 6 < |k| \leq 42 \\ 0 \text{ otherwise} \end{cases} \tag{V.3.4}$$

which peaks at $k = 6$. Taking ω_r as initial data for the HNSE (V.3.1), we shall consider the family of solutions with the parameters mentioned in Table V.3.1 on the time interval $[0, 200]$. Since the initial spectrum slope is compatible with the K67 prediction, a quick development of fine scales is favored, but at the same time dissipative effects do not start to play a role too early. The initial vorticity field is the same for all runs, only ν and α are varied, so that we approach a single solution to the Euler equations, which is our object of study for the rest of this paper. Its initial energy is $E(0) \simeq 2 \cdot 10^{-3}$ and its initial enstrophy is $Z(0) \simeq 0.16$. With that the initial eddy turnover time, defined by $\tau = Z(0)^{-\frac{1}{2}}$, is approximately 2.5. Since the initial data is fixed, the only dimensionless parameter playing a role in the NSE is the Reynolds number, which we choose to define by

$$\mathrm{Re} = \frac{Ul}{\nu} \tag{V.3.5}$$

where $l = \frac{1}{6}$ is the integral scale of the initial flow and $U = l^{-1}\sqrt{2E(0)}$ is the initial RMS velocity. When $\alpha > 1$, we do not attempt to define an equivalent Reynolds number.

The HNSE are discretized in space using a classical fully dealiased Fourier pseudo-spectral method, which is equivalent to a Fourier-Galerkin scheme up to round-off accuracy (Canuto et al., 1988). This ensures that (V.3.2) is satisfied by the semi-discrete solution up to round-off accuracy. We denote by K the maximum wavenumber, and the necessary number of grid points in each direction is then $N = 3K$. ν is chosen as small as possible with the constraint that the solution remains well resolved, which in practice means that it should be proportional to $K^{-2\alpha}$ (see Table V.3.1). For time discretization we employ a third order Runge-Kutta scheme (Orlandi, 2000, p.20), together with an integrating factor to accommodate the viscous term. The duration of each timestep is adjusted according to the CFL condition. Under these conditions, numerical dissipation due to time-discretization was shown to be negligible in (Nguyen van yen et al., 2009). All computations are made in double precision, and using OpenMP for parallelization.

α	1	1	1	1	1	2
$\nu \cdot 10^7$	6.3	1.6	0.40	0.010	0.0025	$8.4 \cdot 10^{-11}$
$\mathrm{Re} \cdot 10^{-3}$	17	66	266	1062	4248	HNSE
N	512	1024	2048	4096	8192	8192

Table V.3.1: Parameters of reference numerical experiments.

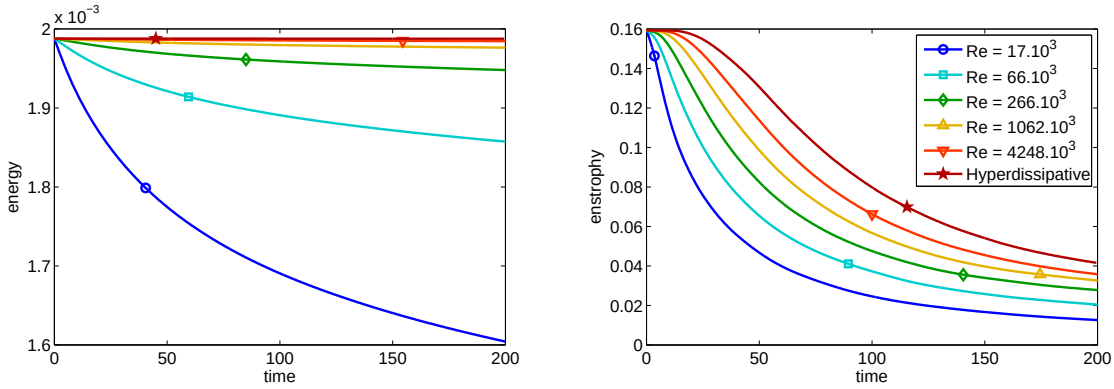


Figure V.3.1: Time evolution of energy (left) and enstrophy (right) for solutions of the 2D NSE equations at various Reynolds numbers.

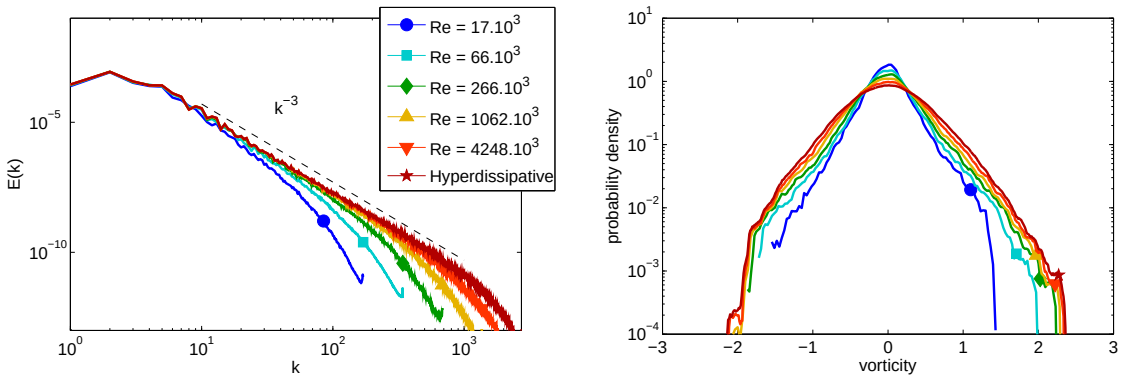


Figure V.3.2: Energy spectra (left) and vorticity pdfs (right) at $t = 50$ for solutions of the 2D NSE equations at various Reynolds numbers.

V.3.4 Statistical analysis

V.3.4.1 Classical statistics

For consistency with previous work, we check that our solutions behave as expected with respect to classical diagnostics. Energy decays in time with a rate that goes to zero like Re^{-1} (Fig. V.3.1, left), while the enstrophy dissipation rate has an apparent dependency on Re which decays much more slowly when $\mathrm{Re} \to \infty$ (Fig. V.3.1, right), in conformity with the results reported by (Dmitruk and Montgomery, 2005; Tran and Dritschel, 2006). We note that at $t = 50$ the spectrum has a k^{-3} power law decay range (Fig. V.3.2, left), so that we consider that a state of fully developed turbulence has been reached. In Fig. V.3.2 (right), we show also for $t = 50$ a histogram of the values taken by the vorticity field on the collocation grid, which is commonly referred to as probability distribution function (PDF). The PDFs have an exponential decay for large values of ω, and a relatively flat core region around $\omega = 0$. It is also noteworthy that both the energy spectrum and vorticity PDF appear to converge to a limit when $\mathrm{Re} \to \infty$, which is consistent with the fact that the solution converges.

V.3.4.2 Wavelet transform

In this section, we briefly introduce the wavelet representation, mostly for the sake of notation. Details may be found in textbooks, see e.g. (Mallat, 1999). Let ψ be a 1-periodic wavelet generating an orthogonal multiresolution analysis of $L^2(\mathbb{T})$, and φ be the associated scaling function. A function f in $L^2(\mathbb{T}^2)$ can be expanded as follows:

$$f = \overline{f} + \sum_{\lambda \in \Lambda} \widetilde{f}_\lambda \psi_\lambda \tag{V.3.6}$$

where $\overline{f}$ is the mean value of f on $\mathbb{T}^2$, $\widetilde{f}_\lambda = \langle f \mid \psi_\lambda \rangle$,

$$\Lambda = \left\{ \lambda = (j, \mathbf{i}, \mu) \mid j \in \mathbb{N}, \mathbf{i} \in \left\{0, \ldots, 2^j - 1\right\}^2, \mu \in \{1, 2, 3\} \right\}$$

and

$$\begin{aligned}
\psi_{(j,\mathbf{i},1)}(x_1, x_2) &= 2^j \psi(2^j x_1 - i_1) \varphi(2^j x_2 - i_2) \\
\psi_{(j,\mathbf{i},2)}(x_1, x_2) &= 2^j \varphi(2^j x_1 - i_1) \psi(2^j x_2 - i_2) \\
\psi_{(j,\mathbf{i},3)}(x_1, x_2) &= 2^j \psi(2^j x_1 - i_1) \psi(2^j x_2 - i_2).
\end{aligned}$$

The index j corresponds to the scale of the wavelet, with the convention that $j = 0$ is the coarsest scale and j increases from coarse to fine scales. The multi-index $\mathbf{i}$ corresponds to the position of the wavelet on the grid of size $2^j \times 2^j$ which is associated to its scale. Finally, the value of μ indicates the directions of oscillation of the wavelet, 1 for the x_1 direction, 2 for the x_2 direction, and 3 for both directions.

In order to approximate the wavelet coefficients of a function using its values at the nodes of a Cartesian grid of size $N \times N = 2^J \times 2^J$, the classical fast wavelet transform algorithm (Mallat, 1999) is utilized. The finest scale which is resolved in this manner is $J-1$. Following earlier work (Farge et al., 2001), we work with Coiflet wavelets corresponding to filters of length 12 (Daubechies, 1993).

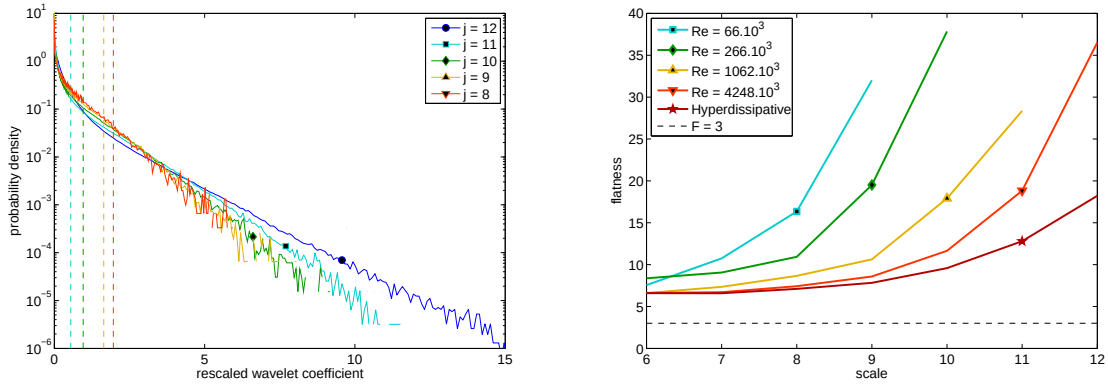


Figure V.3.3: Left: scale-wise PDFs for the HNSE reference solution for $t = 50$ and $\mu = 1$. The vertical dashed lines indicate the threshold determined by the scale-wise CVE algorithm, see Sec. V.3.5.1. Right: scale-wise flatness for varying Reynolds number for $t = 50$ and $\mu = 1$. The dashed line indicates the value 3 which is the flatness of a Gaussian distribution.

V.3.4.3 Scale-wise statistics

In 3D turbulent flows, energy containing regions have a smaller and smaller area when going from coarse to fine scales, a phenomenon known as intermittency. To quantify this phenomenon, scale dependent statistics have often been used, in particular since Sandborn introduced the scale dependent flatness (Sandborn, 1959). Wavelets offer a convenient way of defining scale dependent statistics and of computing them efficiently. Such statistics have been considered e.g. in (Bos et al., 2007) for analyzing fine scale intermittency in anisotropic turbulence. Their relationship with previously introduced statistics that are constructed from structure functions is well known (Farge et al., 1992a). Their computation for 2D flows instead of 3D flows does not pose any particular technical difficulty, but their behavior, as we recall below, is quite different. In the following, except otherwise noticed, by scale dependent statistics we mean quantities that depend on μ and j, that is, they are also direction dependent.

In an orthogonal wavelet representation, there is a separation between different scales of the flow, so that we may legitimately speak of scale by scale — or *scale-wise* — statistics. The main such object that we would like to focus on is the statistical distribution of the wavelet coefficients of the vorticity field at a given scale and in a given direction. It is a generic scale-wise statistical object, since many others can be recovered from it, for example all scale-wise moments of the vorticity field. One way of approximating the scale-wise distribution is to consider the PDF of wavelet coefficients within each scale and direction. Surprisingly, such PDFs have not been considered in previous work. They are akin to the well known PDFs of velocity increments (Kailasnath et al., 1992). However, the PDFs of velocity increments pose some problems when the spectrum is steeper than k^{-3} (Babiano et al., 1985; Biferale et al., 2003). Scale-wise PDFs do not suffer from the same shortcoming, provided that the analyzing wavelet has enough vanishing moments (Schneider et al., 2004). Since we are using a Coiflet with 4 vanishing moments, we would need to start worrying about seeing the effects of the spectrum of the wavelet itself only if the field we are analyzing had a spectrum steeper than k^{-9}.

The results of the scale-wise analysis for the solutions we have at our disposal are re-

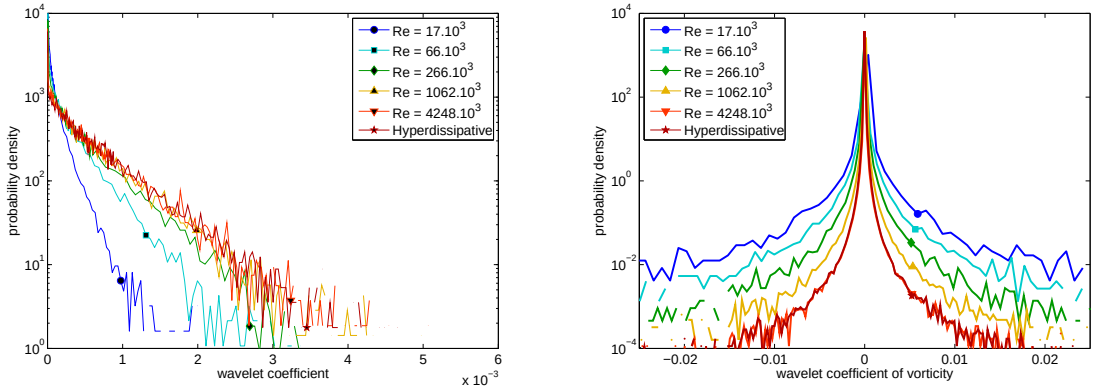


Figure V.3.4: Wavelet coefficients PDFs in the direction $\mu = 1$ for the reference solutions at $t = 50$, for varying Reynolds number. Left: scale-wise PDF at $j = 7$. Right: global PDF of wavelet coefficients.

grouped in Figs. V.3.3 and V.3.4. We focus on $t = 50$ and, since the setup is isotropic, we restrict ourselves to one direction, $\mu = 1$.

We first consider the scale-wise PDF for different scales at the maximum Reynolds number reached by our data (Fig. V.3.3, left). The wavelet coefficients have been rescaled by their standard deviation. All the PDFs have pronounced exponential tails. We notice that the curves for $j = 8, 9, 10$, corresponding to the inertial range, almost superimpose, while the PDFs corresponding to finer scales $j = 11$ and $j = 12$ have longer tails and a more concave shape. We may therefore conclude that the vorticity field is close to being self-similar in the inertial range of scales, while the behavior in the molecular dissipation range is distinct. This conclusion is supported by the behavior of the scale-wise flatness (Fig. V.3.3, right), which seems to approach a value lying between 6 and 7, independent on scale in the inertial range, as Re increases, while it reaches much higher values in the molecular dissipation range.

We now fix the scale to $j = 7$ and look at the scale-wise PDF of $\widetilde{\omega}$ for various Reynolds numbers (Fig. V.3.4, left). We observe convergence towards a limiting curve as Re increases, consistent with our working hypothesis that 2D turbulence has a definite behavior when $\mathrm{Re} \to \infty$. On the contrary, the global PDFs of all wavelet coefficients (Fig. V.3.4, right) do not have a limit when $\mathrm{Re} \to \infty$. To understand this, recall that the global PDF can be seen as a mixture of the scale-wise and direction-wise PDFs for all the active scales of motion, each one being weighted by its contribution to the total number of resolved wavelets. Since the finest scale $j = J - 1$ accounts for the majority of wavelet coefficients (75%), and since its behavior is Re-dependent, it is not surprising that the global PDF inherits this property.

To push this last point further, one can even consider the global and scale-wise PDFs of the same flow but seen at two different numerical resolutions. In Fig. V.3.5, vorticity wavelet coefficients PDFs are compared for the reference flow computed at $N = 4096$ and $\mathrm{Re} \simeq 10^6$, and for the same flow upsampled on a grid with twice the resolution. The upsampling is done in Fourier space, but it more or less boils down to the addition of a new scale $j = 12$ to the flow with nearly vanishing wavelet coefficients. The change in the scale-wise PDFs for $j \leq 11$ are too little to be noticed on the graphs, and the curves overlap (Fig. V.3.5, left). On the contrary, the change in the global PDF cannot be neglected (Fig. V.3.5, right).

The results that we have presented in this section are consistent with most earlier numer-

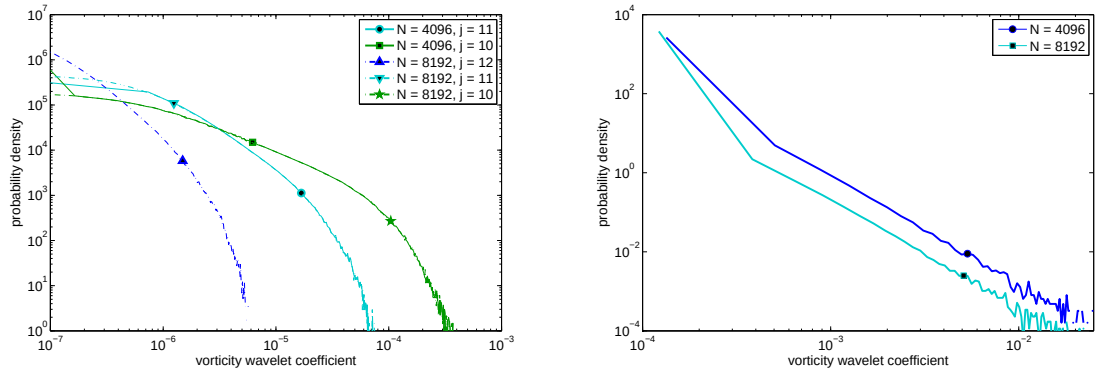


Figure V.3.5: Comparison of vorticity wavelet coefficients PDFs for the reference flow at $\mathrm{Re} \simeq 10^6$ and $t = 50$ (computed with $N = 4096$), and for the same flow upsampled on a grid of size 8192 in each direction. Left: scale-wise PDFs. Right: global PDFs.

ical studies and with experimental studies of flows in soap films, based mostly on structure functions and on vorticity increments statistics, see e.g. the review (Tabeling, 2002) and references therein. One may summarize the situation by stating that vorticity statistics are non Gaussian but nearly self-similar in the inertial range, while intermittency is found in the molecular dissipation range (Fig. V.3.3, left). The main advantage of wavelets over previously used tools is that they form an orthogonal basis: there is no mixing up of information between different scales, and the entire flow can be reconstructed from its wavelet coefficients. The perspectives are appealing, both for computation (see (Schneider and Vasilyev, 2010)), and for physical modelling on which we focus in the remaining sections.

Concerning theory, the situation is much less clear. In (Dritschel et al., 2007), a correction to the K67 energy spectrum was proposed, which consists in keeping the self-similar k^{-3} range but multiplying it with a logarithmic Re-dependent factor. This stands in contradiction with our numerical results, which indicate that the energy spectrum converges to a Re-independent curve. Moreover, a deviation from exact self-similarity is unavoidable, because the solution converges when $\mathrm{Re} \to \infty$ towards an analytic solution to the Euler equation, which must have an exponentially decaying spectrum for large k. This is a further indication that the important point is not to introduce Re-dependent effects, but instead, to change the way our statistical model is built.

V.3.5 Scale-wise coherent vorticity extraction

V.3.5.1 Extraction algorithm

The definition of coherent structures in turbulence has been a matter of debate for decades. In (Farge et al., 1999), it was proposed to define coherent structures as the part of the flow which is not a noise, an approach now known as coherent vorticity extraction (CVE). Such a minimal and negative definition of coherent structures was put forward in the hope that it would be consensual, since it does not rely on the a priori choice of a template for a coherent structure. It also falls neatly in line with the ideas of Kraichnan concerning conditional statistical modelling of turbulence that were reviewed in section V.3.2. The noise corresponds to information which has been dissipated, while coherent structures correspond

to the explicit flow. The price of the approach is that hypotheses need to be made on the noise. As a most simple guess, the noise was assumed to be stationary, additive, Gaussian and uncorrelated (Farge et al., 1999), and it was shown that it could be separated from the rest of the vorticity field using wavelet denoising techniques.

The idea behind the denoising algorithm used to single out the coherent part of the vorticity field is as follows. Once ω has been expanded in an orthogonal wavelet basis (Eq. V.3.6), the terms are split into two groups: coherent terms and incoherent terms. Each coefficient gets attributed to one of the two groups depending on whether its modulus is larger or smaller than a threshold, Θ. An iterative procedure was proposed in (Azzalini et al., 2004) in order to find an optimal value of Θ under the assumption that the incoherent part was uncorrelated, which implied in particular that Θ should not depend on scale. We shall refer to such a Θ as a global threshold, and to the associated CVE algorithm as *global CVE*. In previous work, global CVE has been applied to 2D (Farge et al., 1999; Schneider et al., 2006) and 3D (Farge et al., 2001; Okamoto et al., 2007) turbulence.

In the previous section, we have outlined the fact that the statistics of the wavelet coefficients of a turbulent vorticity field are scale dependent, and shown that over a certain range of scales the scale-wise statistics are resolution independent and have a definite limit when $\mathrm{Re} \to \infty$, while the global statistics enjoy neither of those two properties. Therefore we would like to propose a modified CVE algorithm, called *scale-wise CVE*, which is based on the idea that the threshold should be scale dependent. In the remaining part of this section, the global CVE and scale-wise CVE algorithms are both thoroughly described. Then, in the remaining sections of the paper, global CVE and scale-wise CVE will be compared from several angles.

For $\Theta > 0$, let $\mathbb{1}_\Theta$ be the indicator function of the interval $[-\Theta, \Theta]$, choose $\lambda = (j, \mathbf{i}, \mu)$, and denote by I_λ the set of all λ' which share with λ the same scale j and the same direction μ. Then define the two quantities:

$$N_\lambda(\Theta) = \sum_{\lambda' \in I_\lambda} \mathbb{1}_\Theta\left(\widetilde{\omega}_{\lambda'}\right) \tag{V.3.7}$$

$$\sigma_\lambda(\Theta)^2 = \frac{1}{N_\lambda(\Theta)} \sum_{\lambda' \in I_\lambda} \mathbb{1}_\Theta\left(\widetilde{\omega}_{\lambda'}\right) \widetilde{\omega}_{\lambda'}^2 \tag{V.3.8}$$

which in fact depend only on scale j and direction μ (and not on $\mathbf{i}$), but for which we keep the multi-index λ for convenience of notation. Remark that $N_\lambda(\Theta)$ is the number of wavelet coefficients at scale j and direction μ that are contained in the interval $[-\Theta, \Theta]$, while $\sigma_\lambda(\Theta)$ is their empirical standard deviation. Then construct by recurrence the sequence $(\Theta_{n,\lambda})_{n\in\mathbb{N}}$ such that

$$\begin{cases} \Theta_{0,\lambda} = \infty \\ \Theta_{n+1,\lambda} = q\sigma_\lambda\left(\Theta_{n,\lambda}\right) \end{cases} \tag{V.3.9}$$

where q is a dimensionless constant. Iterating forward in the sequence of thresholds $(\Theta_{n,\lambda})_{n\in\mathbb{N}}$ gradually makes the interval $[-\Theta, \Theta]$ tighten around the wavelet coefficients at scale j and direction μ that are close to zero, while expelling those that are far from zero (Azzalini et al., 2004). The latter are known in statistics as "outliers" .

The constant q controls how restrictive our definition of an outlier is. In the following we have taken $q = 2.8$ for scale-wise CVE, so that for a standard Gaussian random variable, the probability of falling outside the interval $[-q, q]$ is about 0.5%. q can thus be interpreted

as a quantile of the standard Gaussian distribution. For global CVE, we have taken a larger value, $q = 5$, in order to get a compression rate of the same order of magnitude as for scale-wise CVE. In previous publications on global CVE (Azzalini et al., 2004; Schneider et al., 2006; Okamoto et al., 2007), the value $q = \sqrt{2\ln(N^2)}$ was enforced, which is known to be asymptotically optimal for denoising a Gaussian white noise when $N \to \infty$ (Donoho and Jonhstone, 1994), but has the disadvantage of being resolution-dependent. For comparison, if $N = 2048$, $\sqrt{2\ln(N^2)} \simeq 5.52$. We shall come back on the touchy issue of the choice of q in the final discussion.

We should also mention that our definition (V.3.8) for $\sigma_\lambda(\Theta)$ slightly differs from the one in (Azzalini et al., 2004), because we use $N_\lambda(\Theta)$ as denominator (which is the correct denominator to use when computing a variance), while (Azzalini et al., 2004) used the total number of wavelet coefficients. The price of this slight difference is that the proof of convergence of the sequence $(\Theta_{n,\lambda})_{n\in\mathbb{N}}$ given in (Azzalini et al., 2004) does not apply to our version of the algorithm. Nevertheless we observe experimentally that above a certain value of n (less than 100), $\Theta_{n,\lambda}$ becomes constant with a value Θ_λ satisfying:

$$\Theta_\lambda = q\sigma_\lambda\left(\Theta_\lambda\right) \tag{V.3.10}$$

Since the number of distinct wavelet coefficients within scale j and direction μ is 2^{2j}, the very coarse scales of the flow contain too few wavelet coefficients for any statistical quantity to be meaningful, and applying the above procedure to these scales would be quite hazardous. Hence for $j \leq 4$ we prefer to impose $\Theta_\lambda = 0$ (everything is coherent).

Once Θ_λ has been obtained, wavelet coefficients whose modulus lies below Θ_λ, depending on scale j and direction μ, are defined as incoherent, while the remaining ones are defined as coherent. The wavelet coefficients index set Λ is thus split into a set of incoherent coefficients, Λ^I, and a set of coherent coefficients, Λ^C. Note that a similar procedure was independently proposed for data classification in (Hennig, 2003). To impose a regularizing effect of the thresholding operation on the Euler equations (Nguyen van yen et al., 2009), we always enforce that $\Theta_\lambda = \infty$ for the finest scale at the current resolution $j = J-1$. The global CVE algorithm is obtained if one changes (V.3.7-V.3.8) by extending the sums to all directions and to all scales except the finest one $j = J-1$, for which we keep $\Theta_\lambda = \infty$, for the reason already mentioned.

In either case we then define coherent vorticity ω^C via its wavelet coefficients:

$$\widetilde{\omega}_\lambda^C = \mathbb{1}\!\!\!\!\mathbb{1}_{\Theta_\lambda}\left(\widetilde{\omega}_\lambda\right)\widetilde{\omega}_\lambda = \begin{cases} \widetilde{\omega}_\lambda \text{ if } |\widetilde{\omega}_\lambda| \geq \Theta_\lambda \\ 0 \text{ otherwise} \end{cases} \tag{V.3.11}$$

and incoherent vorticity is given by the difference with total vorticity:

$$\omega^I = \omega - \omega^C \tag{V.3.12}$$

Since the wavelet basis is orthogonal, the respective enstrophies Z_C and Z_I of the coherent incoherent part are related to the total enstrophy Z by:

$$Z = Z_C + Z_I \tag{V.3.13}$$

From ω^C and ω^I, coherent and incoherent velocity fields $\mathbf{u}^C$ and $\mathbf{u}^I$ can be reconstructed, but they are in general not orthogonal, so that to define an analog split for the energy, a cross term needs to be taken into account:

$$E = \frac{1}{2}\|\mathbf{u}^C\|^2 + \frac{1}{2}\|\mathbf{u}^I\|^2 + \langle \mathbf{u}^C \mid \mathbf{u}^I \rangle = E_C + E_I + E_{CI} \tag{V.3.14}$$

V.3.5.2 Results

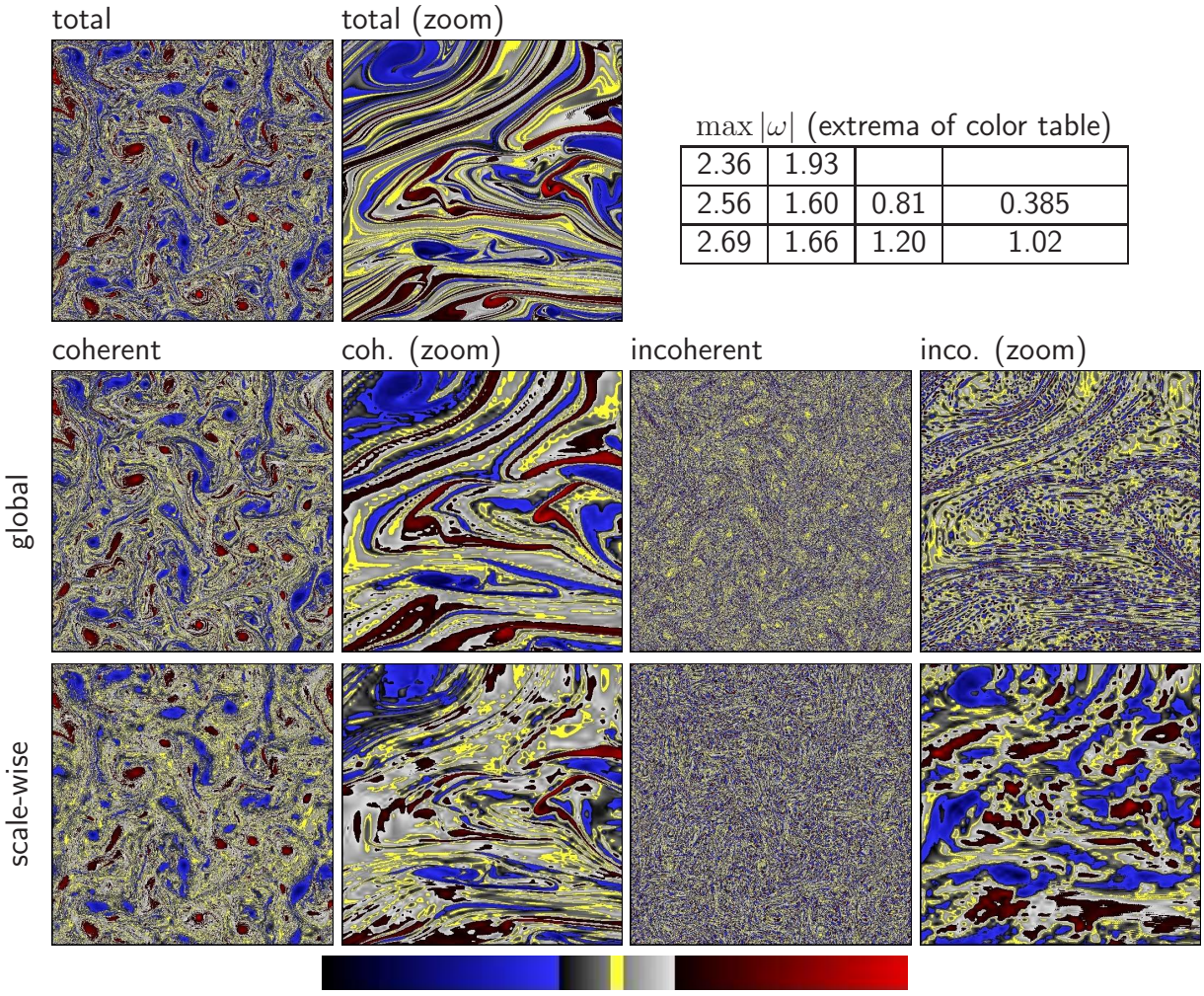
Now that the global CVE and scale-wise CVE algorithms have been introduced, we apply them to analyze the vorticity fields obtained after the computations described in section V.3.3, and whose parameters are summarized in Table V.3.1. This section is a follow up on previous results concerning 2D turbulence (Farge et al., 1999; Schneider et al., 2006), with a strong focus on the two following aspects:

- the dependency on Re in the limit $\mathrm{Re} \to \infty$,
- the interpretation in terms of turbulent dissipation.

In Fig. V.3.6, we show snapshots of the vorticity fields obtained by performing the split for the HNSE reference solution at $t = 50$. The main structures visible by eye in the total vorticity field are preserved in the coherent vorticity field for both methods (first column). The difference between global and scale-wise CVE is better seen by looking at the incoherent parts (third column): for scale-wise CVE it appears quite homogeneous, whereas for global CVE the remnants of structures can still be glimpsed. Now since the resolution of the field is $N = 8192$ in each direction, its snapshots lacks a lot of details, which may give us a false impression. To check this, we look at zooms on small squares of size $\frac{1}{16}$ located in the top-bottom corners of each pictures (Fig. V.3.6, second and fourth row). We notice that scale-wise CVE enhances most of the sharp features of the vorticity field, while the effect of global CVE is not very pronounced. In fact, as was pointed out already in (Schneider et al., 2006), for global CVE the coherent part looks very similar to the total vorticity field. On the contrary the distinction is immediate for scale-wise CVE. Perhaps the most striking difference is found on the zooms on the incoherent parts shown in the fourth column: for scale-wise CVE we see a smooth and disorganized field, while for global CVE we see a rough field containing the trace of fine scale filaments.

To increase the compression rate, it was proposed in (Schneider et al., 2006) to use intermediate thresholds belonging to the sequence $(\Theta_{n,\lambda})$ defined by (V.3.9), instead of the limit $n \to \infty$. Different result may be obtained using this technique, but since compression is not our central concern here, we have not considered it.

Now turning to more quantitative features of the coherent and incoherent flows, we consider their vorticity PDFs (Fig. V.3.7) and energy spectra (Fig. V.3.8, left). For both global and scale-wise CVE, the vorticity PDF is relatively well approximated by the coherent vorticity PDF, and the extrema of the vorticity field are captured by its coherent part. The incoherent vorticity PDFs are supported on a narrower interval and have a nearly Gaussian shape. Quantile-quantile plots (Fig. V.3.7, right) allow us to enhance the deviation with respect to Gaussiannity, which is seen to be more pronounced for global CVE than for scale-wise CVE. The distinction between global and scale-wise CVE becomes more evident when looking at the energy spectra (Fig. V.3.8, left). For global CVE, the coherent energy spectrum closely follows the total energy spectrum except in the molecular dissipation range, where incoherent energy becomes dominant, in agreement with earlier results (Schneider et al., 2006). However the k^{-1} scaling of the incoherent energy spectrum (corresponding to enstrophy equipartition) that was observed in (Schneider et al., 2006) can be seen here only over a restricted range of wavenumbers, while the overall shape of the incoherent energy spectrum seems better described by a k^0 scaling. For scale-wise CVE, the total, coherent, and incoherent energy spectra all have a k^{-3} scaling range. The incoherent spectrum falls



$\max|\omega|$ (extrema of color table)

2.36	1.93		
2.56	1.60	0.81	0.385
2.69	1.66	1.20	1.02

Figure V.3.6: Snapshots of the total vorticity field (first row), and of its coherent (first and second columns) and incoherent parts (third and fourth columns) for global CVE (middle row) and scale-wise CVE (bottom row). The second (resp. fourth) column shows a restriction of the coherent (resp. incoherent) part to the subdomain $[\frac{1}{16}, \frac{1}{16}]$ (corresponding to the lower left corner of the full image). The absolute maxima of the respective fields are given in the table on the top right, following the same arrangement as for the images of the fields. The color table, shown at the bottom of the figure, varies from $-\max|\omega|$ to $+\max|\omega|$ for each picture.

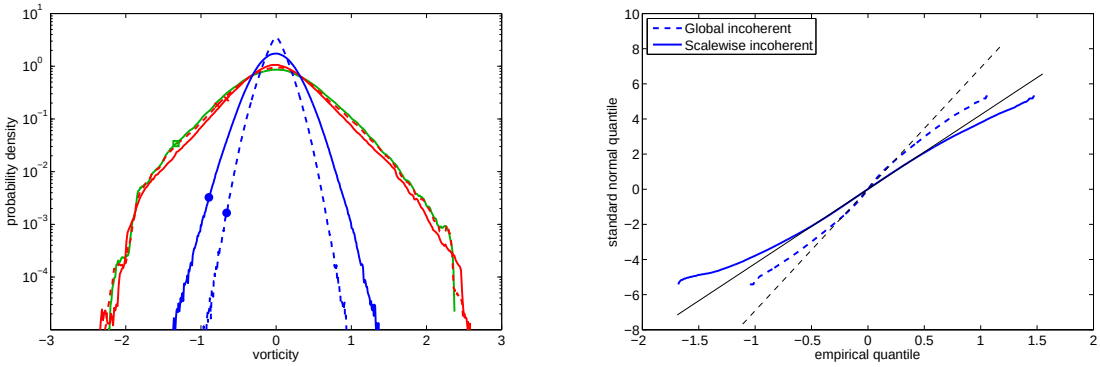


Figure V.3.7: Analysis of total, coherent, and incoherent flows, as defined by scale-wise CVE (full lines) and global CVE (dashed lines), for the HNSE reference solution at $t = 50$. Left: vorticity PDFs. Right: quantile-quantile plots of the incoherent vorticities versus the standard normal distribution. Linear fits corresponding to normal distributions with matching means and standard deviations are shown for comparison.

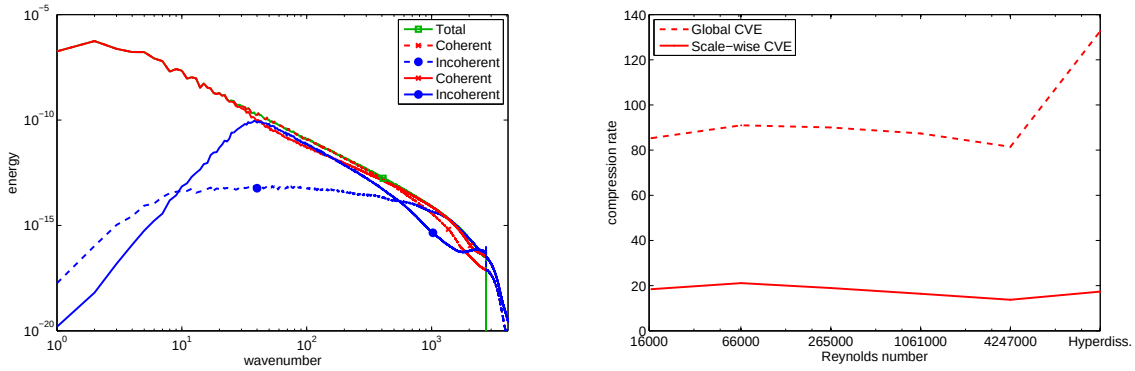


Figure V.3.8: Left: energy spectra of total, coherent, and incoherent flows, as defined by scale-wise CVE (full lines) and global CVE (dashed lines), for the HNSE reference solution at $t = 50$. Right: compression rate as a function of Reynolds number for global and scale-wise CVE.

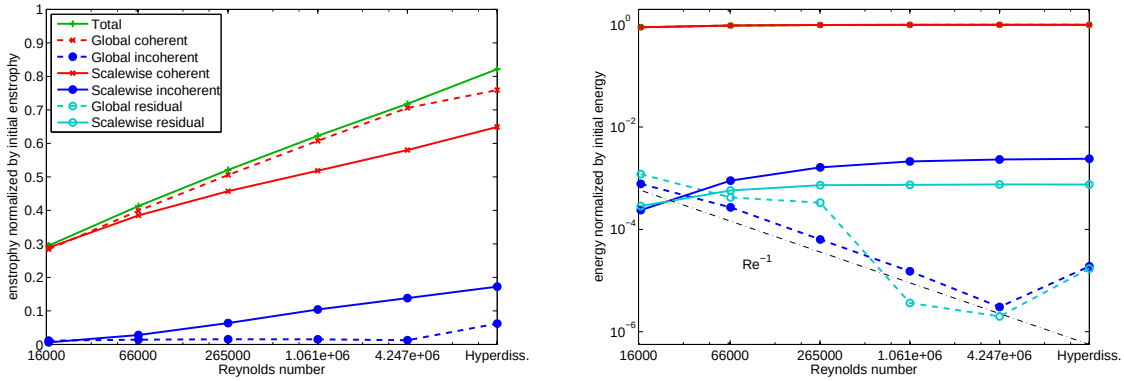


Figure V.3.9: Euler invariants of total, coherent and incoherent flows, as defined by scale-wise CVE (full lines) and global CVE (dashed lines), at $t = 50$, as functions of Reynolds number. Left: enstrophy, normalized by the initial enstrophy $Z(0)$. Right: energy, normalized by the initial energy $E(0)$, and plotted with a logarithmic vertical scale.

down more rapidly than the coherent one in the far inertial range, and therefore the coherent part becomes dominant in the dissipative range, while in the inertial range coherent and incoherent enstrophies are of the same order of magnitude. Recall that for scale-wise CVE all wavelet coefficients at scales $j \leq 4$ are automatically coherent. Therefore the behavior of the incoherent spectrum for $k \leq 16$ is just a byproduct of the spectrum of the analyzing wavelet.

Since one of our goals is to compare global and scale-wise CVE, it is also important to measure how much the amount of information contained in the flow is reduced if one keeps track explicitly only of coherent vorticity instead of total vorticity. The measurement of information is a difficult problem on which we shall come back in section V.3.7.1. Here, we limit ourselves to a well known diagnostic, the compression rate $\mathcal{R}$ defined by:

$$\mathcal{R} = \frac{\#\Lambda}{\#\Lambda_C}$$

where $\#\cdot$ denotes the number of elements in a set. $\mathcal{R}$ is plotted as a function of Reynolds number in Fig. V.3.8 (right). We observe that $\mathcal{R}$ is close to 20 for scale-wise CVE and close to 80 for global CVE. The higher value, close to 130, reached for global CVE with hyperdissipation can be explained by the fact that the global CVE is more sensitive to the behavior of the wavelet coefficients in the dissipation range, which is quite different for the hyperdissipative solution, as can be deduced from its scale-wise flatness (Fig. V.3.3, right). Overall, an improvement in the compression rate when Re increases is observed neither for global CVE nor for scale-wise CVE. Further study will be necessary to determine if this result also applies to wall-bounded 2D turbulence or to 3D turbulence (see (Kevlahan et al., 2007; Okamoto et al., 2007) for recent results along this line).

Based on the above results we would like to conjecture that for scale-wise CVE, in the limit $\mathrm{Re} \to \infty$, the incoherent enstrophy and energy both converge to a nonzero limit.To test this conjecture, consider the energy and enstrophy of the flow at $t = 50$, normalized by their initial values, as functions of Reynolds number (Fig. V.3.9). In agreement with the literature (Dmitruk and Montgomery, 2005; Tran and Dritschel, 2006), the molecular enstrophy dissipation vanishes $\mathrm{Re} \to \infty$, so that the ratio $\frac{Z(t=50)}{Z(t=0)}$ approaches 1. For global CVE, the amount of incoherent enstrophy is small and nearly Re-independent, while for

scale-wise CVE, it increases with Re. Unfortunately, a saturation regime for incoherent enstrophy is not reached in the range of Re that we have been able to consider. However, such a regime is reached for incoherent energy (Fig. V.3.9, right), albeit at a relatively low fraction of about $2 \cdot 10^{-3}$ of the initial energy. Hence the conjecture is valid at least for energy, and is likely to be valid also for enstrophy. For global CVE, coherent energy goes to zero like Re^{-1}. Since the main difference between the HNSE and the NSE solutions is the behavior in the dissipation range, the fact that the last points stand out on the curves corresponding to global CVE suggests that the latter is most sensitive to this range of scales, a property that we have already encountered earlier.

Now that we have established the statistical properties of the coherent and incoherent parts, we would like to assess the practical relevance of the split as regards the 2D Euler dynamics. As a first step in that direction, we consider the enstrophy transfers, both between the coherent and incoherent parts, and between different scales of motion.

V.3.6 Interscale enstrophy transfers and production of incoherent enstrophy

V.3.6.1 Transfers in Fourier space

In the regime that we are considering, enstrophy is transferred on average from low wavenumber modes to large wavenumber modes. To quantify this process, it is convenient to introduce the orthogonal projector P_k on modes with wavenumbers whose modulus is smaller than k:

$$P_k(f) = \sum_{|\mathbf{k}| \leq k} \widehat{f}[\mathbf{k}] \exp(2\iota\pi \mathbf{k} \cdot \mathbf{x})$$

and the vorticity field can then be split as follows:

$$\omega = P_k\omega + (1 - P_k)\omega$$

where the two terms are orthogonal to each other. Thanks to the Pythagore identity the enstrophy can in turn be split into two terms: $Z = Z_{\leq} + Z_{>} = \frac{1}{2}\|P_k\omega\|^2 + \frac{1}{2}\|(1 - P_k)\omega\|^2$, and the goal is to determine the transfer from $Z_{\leq}$ to $Z_{>}$, or interscale enstrophy transfer.

The procedure is classic (see e.g. (Kraichnan, 1967)) but we would like to recall it in detail here since it will serve as an introduction to the next paragraph where wavelet transfers are to be considered. One first writes down the evolution equation for $Z_{\leq}$ by bracketing the NSE with $P_k(\omega)$:

$$\frac{\mathrm{d}Z_{\leq}}{\mathrm{d}t} + \langle \mathbf{u} \cdot \boldsymbol{\nabla}\omega \mid P_k\omega \rangle + \nu \langle \Delta\omega \mid P_k\omega \rangle = 0 \tag{V.3.15}$$

Then one may define a trilinear form a by

$$a(\omega_1, \mathbf{u}, \omega_2) = \langle \mathbf{u} \cdot \boldsymbol{\nabla}\omega_1 \mid \omega_2 \rangle \tag{V.3.16}$$

where $\mathbf{u}$ is a divergence free vector field, and ω_1, ω_2 are scalar fields. The essential property of a is that it is antisymmetric with respect to its first and last variables (Foiaš et al., 2001): $a(\omega_1, \mathbf{u}, \omega_2) = -a(\omega_2, \mathbf{u}, \omega_1)$, and in particular for any $\mathbf{u}$ and ω_1 we have $a(\omega_1, \mathbf{u}, \omega_1) = 0$. Using that property along with (V.3.15), one finally obtains the system

$$\begin{cases} \frac{\mathrm{d}Z_{\leq}}{\mathrm{d}t} + a((1 - P_k)\omega, \mathbf{u}, P_k\omega) + \nu\|\boldsymbol{\nabla} P_k\omega\|^2 = 0 \\ \frac{\mathrm{d}Z_{>}}{\mathrm{d}t} - a((1 - P_k)\omega, \mathbf{u}, P_k\omega) + \nu\|\boldsymbol{\nabla}(1 - P_k)\omega\|^2 = 0 \end{cases} \tag{V.3.17}$$

where $Z_{\leq}$ and $Z_{>}$ are now ostensibly coupled by the transfer term

$$\Pi_k = a((1-P_k)\omega, \mathbf{u}, P_k\omega) \tag{V.3.18}$$

Following (Okamoto et al., 2007), we would like to discriminate between the coherent and incoherent contributions to Π_k. Since it is trilinear, eight contributions can be pulled out:

$$\Pi^{\alpha\beta\gamma}[k] = a((1-P_k)\omega^{\alpha}, \mathbf{u}^{\beta}, P_k\omega^{\gamma}) \tag{V.3.19}$$

where $(\alpha, \beta, \gamma) \in \{I, C\}^3$. To make the distinction with the next section clear, let us insist on the fact that for any choice of α, β and γ, $\Pi^{\alpha\beta\gamma}$ is just part of the transfer from $Z_{\leq}$ to $Z_{>}$, but contains no information about the production of incoherent enstrophy.

$\Pi^{\alpha\beta\gamma}[k]$ is plotted as a function of k in Fig. V.3.10, for the HNSE reference simulation at $t = 50$. For global CVE (Fig. V.3.10, left), we find that the term Π_{CCC} dominates the transfers in the inertial range, while some other terms become non negligible only in the molecular dissipation range. This is not surprising given what we have learned in the previous section, namely that the fraction of incoherent enstrophy is very low in the inertial range. The four transfer terms associated to the incoherent velocity field, namely $\Pi^{\alpha I\gamma}$ (Fig. V.3.10, dashed lines), are two to three orders of magnitude smaller than those associated to the coherent velocity field (Fig. V.3.10, full lines). For scale-wise CVE, all terms of type $\Pi^{\alpha C\gamma}$ are non-negligible throughout the inertial range, which means that both the coherent and incoherent parts participate in the nonlinear transfer of enstrophy from low wavenumber modes to large wavenumber modes. The term Π^{CIC} also participates to the transfer with a share of up to 10%.

V.3.6.2 Transfers in wavelet space

The above procedure may be followed again, but this time starting from an orthogonal wavelet basis instead of the Fourier basis (Meneveau, 1991). The projector P_k is replaced by $\widetilde{P}_j$, the orthogonal projector on the subspace generated by wavelets whose scale is coarser than j. One ends up with a wavelet interscale transfer term $\widetilde{\Pi}_j$ which is the exact analog of Π_k defined by (V.3.18). Two main disadvantages have to be put up with when considering transfers in wavelet space as opposed to Fourier space:

- in addition to the effect of the nonlinear term, there can be an enstrophy transfer due to the molecular dissipation term. Indeed, the nice decoupling observed in (V.3.17) was possible because the projector P_k commutes with the Laplace operator, whereas $\widetilde{P}_j$ does not,
- the wavelets are not as localized in Fourier space as the Fourier modes themselves. This is a price to pay for the space localization of the wavelets.

The bright side is that wavelets allow us to refine the transfers analysis by looking separately at the coherent and incoherent parts, as we now proceed to explain. Let us assume that the set Λ^C of coherent wavelet coefficients is chosen by applying one of the two algorithms of the previous section, either global CVE, or scale-wise CVE. There are three types of contributions to the enstrophy transfers between the coherent and the incoherent parts:

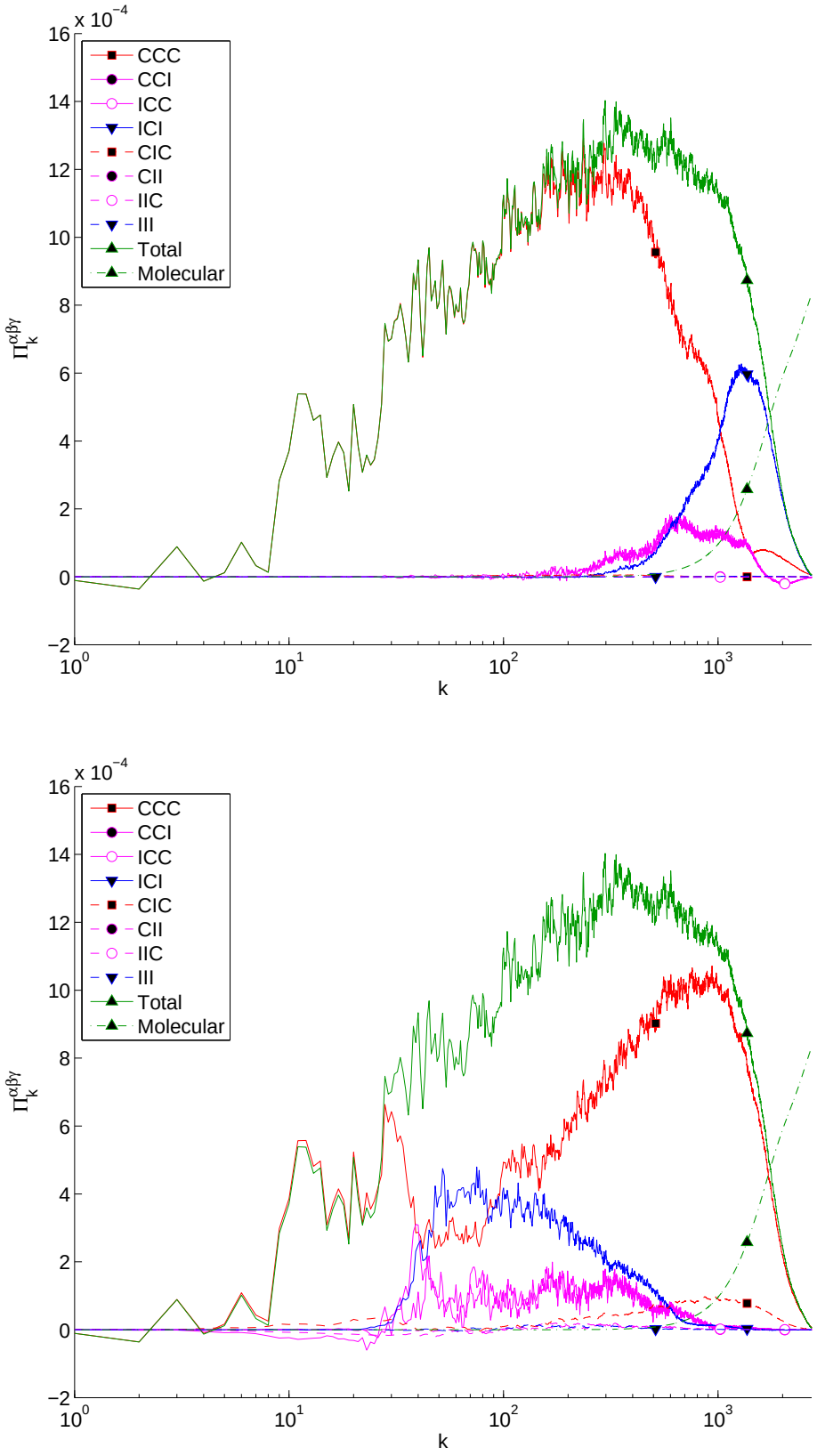


Figure V.3.10: Contributions to enstrophy transfer between Fourier modes. Left: global CVE. Right: scale-wise CVE.

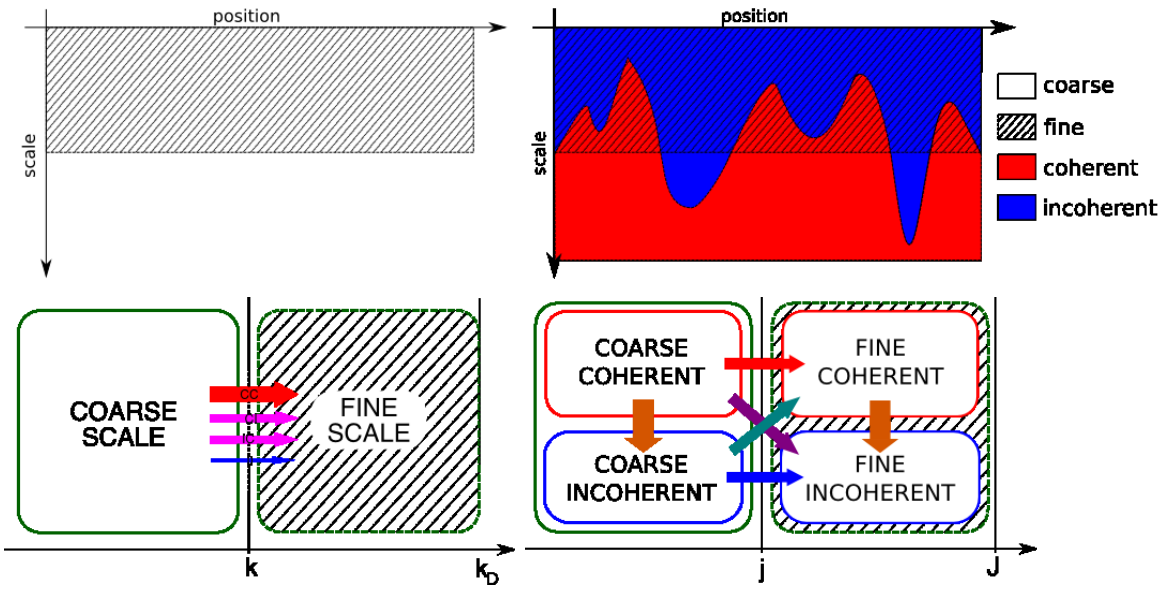


Figure V.3.11: Top row: schematic view of the segmentation used in the wavelet transfer analysis (Sec. V.3.6.2). On the left, only the distinction between coarse and fine scales is made, while on the right coherent and incoherent coefficients are considered separately. Bottom row: diagram showing the possible enstrophy transfer paths. Left: transfers in Fourier space. Right: transfers in wavelet space.

- the coupling by the nonlinear term,
- the coupling by the molecular dissipation term,
- the change in time of the sets Λ^C and Λ^I.

Here we shall leave out the second contribution, since we are mostly interested in nonlinear transfers, and also the third contribution, by assuming that Λ^C and Λ^I are fixed.

Denoting by $\widetilde{P}^C$ the projector on wavelets ψ_λ such that $\lambda \in \Lambda^C$, and $\widetilde{P}^I = 1 - \widetilde{P}^C$, we may thus split the enstrophy as follows:

$$\begin{aligned} Z &= \frac{1}{2}\|\widetilde{P}^C \widetilde{P}_j \omega\|^2 + \frac{1}{2}\|\widetilde{P}^I \widetilde{P}_j \omega\|^2 && \text{(V.3.20)} \\ &\quad + \frac{1}{2}\|\widetilde{P}^C \left(1 - \widetilde{P}_j\right) \omega\|^2 + \frac{1}{2}\|\widetilde{P}^I \left(1 - \widetilde{P}_j\right) \omega\|^2 && \text{(V.3.21)} \\ &= Z^C_{\leq} + Z^I_{\leq} + Z^C_{>} + Z^I_{>}, && \text{(V.3.22)} \end{aligned}$$

Given that $\widetilde{P}^C$ and $\widetilde{P}_j$ commute and are both orthogonal, a system of four equations similar to (V.3.17) can be derived to describe the time evolution of the four terms. The difference between the Fourier and wavelet viewpoints is summarized by the diagram in Fig. V.3.11. A similar kind of diagram was introduced in (Goldstein and Vasilyev, 2004) to describe the stochastic coherent adaptive large eddy simulation (SCALES) computational approach. However the enstrophy transfers between the various components have not been measured before.

The wavelet transfers are shown in Fig. V.3.12, for the HNSE reference solution at $t = 50$. While reading the following discussion, keep in mind that we are analyzing a single

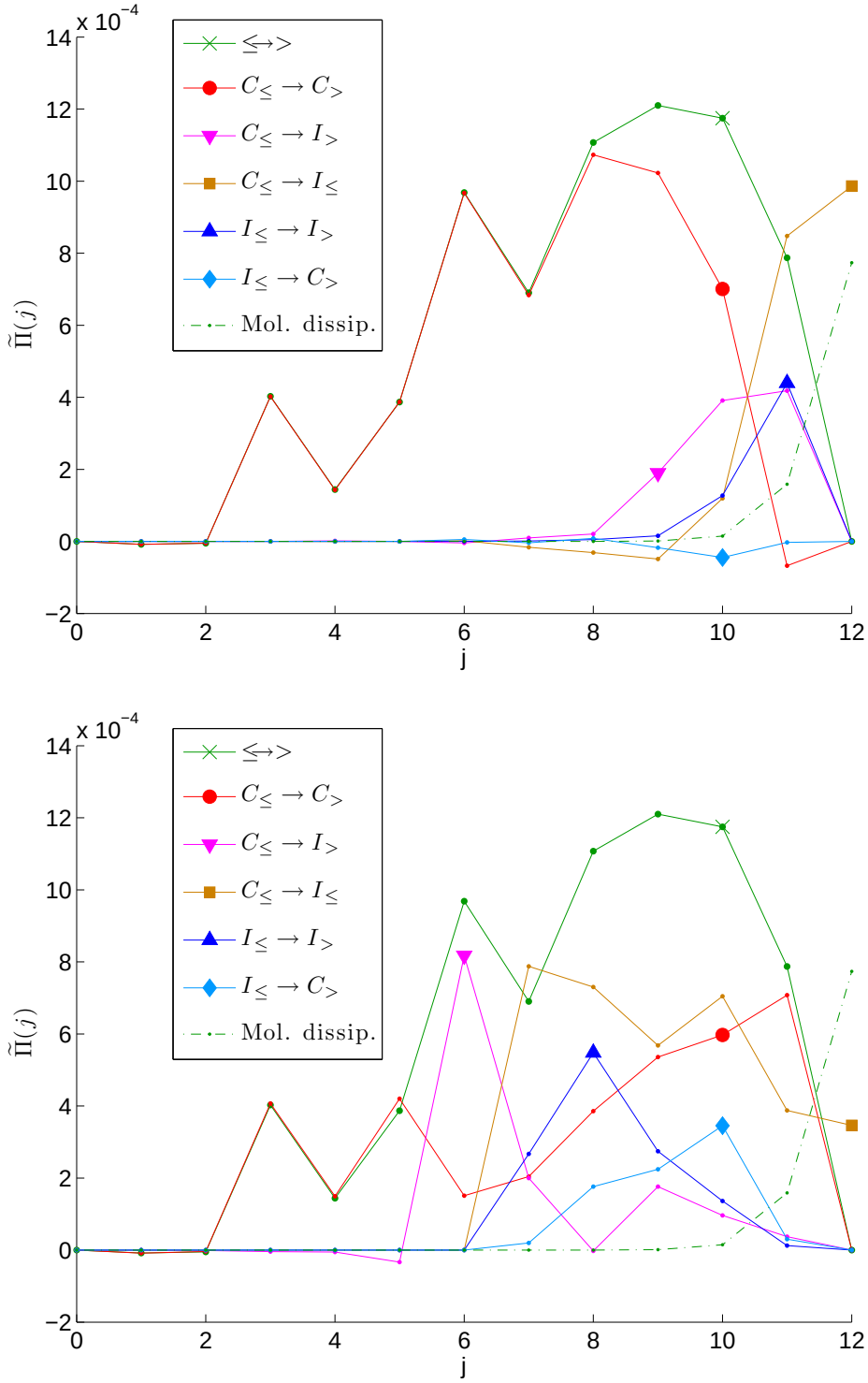


Figure V.3.12: Wavelet transfers. Left: global CVE. Right: scale-wise CVE.

	Global	Scale-wise
$C_{\leq} \to C_{>}$	92.3%	48.4%
$C_{\leq} \to I_{>}$	17.1%	15.9%
$C_{>} \to I_{\leq}$	-1.58%	-20.2%
$I_{\leq} \to I_{>}$	1.39%	24.8%
$\leq \to >$	100%	100%
$C_{\leq} \to I_{\leq}$	-4.4%	51.3%

Table V.3.2: Wavelet enstrophy transfers across $j = 9$ for global and scale-wise CVE applied to the reference HNSE solution at $t = 50$. The numbers are expressed as percentages relative to the total interscale transfer ($\leq \to >$).

realization at a single instant, and that we are not making any stationarity hypothesis. We are not attempting to extract scaling laws, but rather we are focusing on understanding CVE in relation to the time dependent Navier-Stokes dynamics.

The total enstrophy transfer from coarse to fine scales ($\leq \to >$) is in agreement with the result obtained in the previous section using Fourier analysis. So is the molecular dissipation effect. Key differences between global and scale-wise CVE are revealed by looking at the other curves. For global CVE (Fig. V.3.12, left), the production of incoherent enstrophy ($C_{\leq} \to I_{\leq}$ and $C_{\leq} \to I_{>}$) occurs only at scales where molecular dissipation cannot be neglected. At coarser scales, all the enstrophy transfers are dominated by $C_{\leq} \to C_{>}$, that is, the transfer of coherent enstrophy from coarse to fine scales, similarly to what was observed for energy in 3D turbulence (Okamoto et al., 2007).

Since in the case of scale-wise CVE the dependency with j is quite complex, let us first consider a single abscissa, $j = 9$. The transfers, expressed as percentages of the total nonlinear transfer, are shown in Table V.3.2. The situation as a result of global CVE is also shown for comparison purposes. The first important thing that we notice is that for scale-wise CVE, and contrary to global CVE, the production of incoherent enstrophy ($C_{\leq} \to I_{\leq}$) is of the same order of magnitude as the interscale transfer $C_{\leq} \to C_{>}$. Recall the comment of Kraichnan (Kraichnan, 1974): the processes which mix enstrophy at scales $j \leq 9$ act on the same time scale as the processes which send enstrophy from $C_{\leq}$ to $C_{>}$. Next, we should mention the transfer $C_{>} \to I_{\leq}$, which measures how much of the fine scale coherent enstrophy is transferred to coarse scale incoherent enstrophy. From the point of view that we have adopted, any transfer $C \to I$ is seen as a dissipation of enstrophy. Hence the fact that the transfer $C_{>} \to I_{\leq}$ is negative and non negligible (-23%) raises the issue of *negative dissipation*. Notice that the same property holds for global CVE, albeit with a smaller amplitude. But we should not be scandalized by observing negative dissipation in turbulent flows. Indeed, it is well known that organized structures do spontaneously emerge out of initially random flows. The CVE approach has the merit of quantifying this phenomenon.

Now let us turn to Fig. V.3.12 (right) which shows the j-dependency of the transfers in the case of scale-wise CVE. We consider only scales $j \geq 5$ since below $j = 4$ everything was assumed to be coherent in the definition of the algorithm. The first thing to note is that for $6 \leq j \leq 9$ the dependency of the transfers on j is relatively benign, and the various terms keep the same orders of magnitude. We thus recover the almost self-similar behavior that could be deduced from the scale-wise statistics. But looking more closely, it appears that the

transfer due to the coherent part tends to increase when going to finer scales. We interpret this as a time-dependent effect: the initial condition consists only of coarse scale motion, and it takes some time for incoherent enstrophy to build up and to start transferring energy to fine scales. The behavior for $j \geq 10$ is also worthy of some notice, especially since it is highly counter-intuitive. Indeed, when molecular dissipation dominates, one may expect the transfers to be mostly incoherent, as was the case for global CVE. But the situation is quite the opposite: for $j = 10$ and $j = 11$, the transfer is actually dominated by the $C_{\leq} \to C_{>}$ term.

We now come back to the overall picture, as it may be understood from the wavelet transfer analysis. It appears that global CVE behaves in a manner which is similar to molecular dissipation: indeed, the production of incoherent part occurs mostly in the dissipative range of scales, while the nonlinear transfer in the inertial range is associated almost exclusively to the coherent part. Scale-wise CVE offers a more radical view point on the inertial range in 2D turbulence, by splitting the interscale transfer of enstrophy into two parallel channels, one associated to the coherent part, and one associated to the incoherent part. The enstrophy transfers through both channels are of the same order of magnitude, although only a few percent of the wavelet coefficients of the coherent vorticity field are nonzero. Since the scale-wise PDFs do not depend much on j in the inertial range, there is a small fraction of coherent coefficients within each scale. From the two previous sentences we may infer that the transfer of enstrophy through the coherent channel is due to localized events, involving few wavelet coefficients, while the transfer through the incoherent channel is much more homogeneous in space. The dominant exchange between the two channels is the conversion of coherent into incoherent enstrophy, $C_{\leq} \to I_{\leq}$, but there is also a non-negligible backwards conversion $I_{\leq} \to C_{>}$, which is maximal in the far inertial range, i.e. just before the dissipation range. This negative dissipation feeds the fine scale coherent part, and may act as a source of flow intermittency in the dissipation range (Farge et al., 1992b). As a result, in the dissipation range, the flow is intermittent, and the dominant channel is the coherent one.

From the above discussion, we conjecture that simulation using global CVE does not need a turbulence model, and can thus be considered as DNS. In contrast, a turbulence model would be necessary for simulation based on scale-wise CVE in order to take into account the retroaction of the incoherent part onto the coherent part. This will be the topic of the next section.

V.3.7 Dynamical influence of the incoherent part

V.3.7.1 Randomization as dissipation

As pointed out in the introduction, it is said that a quantity is dissipated when the corresponding degrees of freedom are replaced by random variables. The remaining explicitly computed degrees of freedom are then perturbed stochastically by the dissipated (e.g. incoherent) ones. We would like to find out more about this perturbation in the two specific cases of global and scale-wise CVE.

When a wavelet coefficient indexed by λ is replaced by a random variable, we say that it has been randomized, and we denote it $\widetilde{W_\lambda}$. We have already defined in section V.3.5.1 the index set of incoherent wavelet coefficients according to global CVE and to scale-wise

CVE. There remains to choose a probability distribution for the random vector $(\widetilde{W}_\lambda)_{\lambda\in\Lambda^I}$. A natural constraint that we would like to impose is that applying the randomization operator twice in a row is equivalent to applying it only once, that is, that the operator is idempotent. For this condition to hold exactly, the incoherent coefficients must stay below the threshold

$$\widetilde{W}_\lambda \in [-\Theta_\lambda, \Theta_\lambda] \tag{V.3.23}$$

and the value of the threshold itself should stay invariant under randomization, implying that

$$\mathbb{E}(\widetilde{W}_\lambda) = 0 \tag{V.3.24}$$

and also, because of Eq. V.3.10:

$$\frac{1}{N_\lambda(\Theta_\lambda)} \sum_{\lambda'\in\Lambda^I\cap I_\lambda} \widetilde{W}^2_{\lambda'} = \left(\frac{\Theta_\lambda}{q}\right)^2. \tag{V.3.25a}$$

Under these three constraints, the most unbiased choice that we may make is the distribution which maximizes the Shannon entropy, that is, the uniform distribution on the manifold defined by Eqs. (V.3.23) and (V.3.25a) in $\mathbb{R}^{\#\Lambda^I}$, that is, the intersection of a hypersphere and of a hypercube. Unfortunately, we have not found an efficient way of generating pseudo-random numbers distributed accordingly. Hence for practical reasons, we propose to replace (V.3.25a) with the more tractable constraint that

$$\mathbb{E}\left(\widetilde{W}^2_\lambda\right) = \left(\frac{\Theta_\lambda}{q}\right)^2 \tag{V.3.25b}$$

which ensures that enstrophy is conserved by the randomization operator in the ensemble average sense. Note the analogy of (V.3.25a) with the conservation of internal energy characterizing the microcanonical ensemble, and of (V.3.25b) with the temperature constraint of the canonical ensemble.

We admit that the solution of the entropy maximization problem under the three constraints (V.3.23-V.3.24-V.3.25b) can be factorized into a product of univariate distributions f_λ. To maximize the entropy of f_λ it is convenient to first rescale it by its standard deviation:

$$f_\lambda(w) = \frac{1}{\sigma_\lambda} f_0\left(\frac{w}{\sigma_\lambda}\right)$$

and then solve the equivalent problem of maximizing the entropy of the rescaled distribution

$$S(f_0) = \int_{\mathbb{R}} f_0 \ln(f_0)$$

under the constraints that f_0 has variance 1 and is supported on $[-q, q]$. The solution turns out to be a truncated Gaussian distribution, given by

$$f_0(w) = \begin{cases} \mathcal{Z} \exp\left(-\frac{w^2}{2s_0^2}\right) & \text{if } |w| \le q \\ 0 \text{ otherwise} \end{cases}$$

where $\mathcal{Z}$ is a normalization factor and the value of s_0 has to be chosen so that

$$\int_{-q}^{q} \mathrm{d}w w^2 f_0(w) = \frac{\int_{-q}^{q} w^2 \exp\left(-\frac{w^2}{2s_0^2}\right) \mathrm{d}w}{\int_{-q}^{q} \exp\left(-\frac{w^2}{2s_0^2}\right) \mathrm{d}w} = 1 \tag{V.3.26}$$

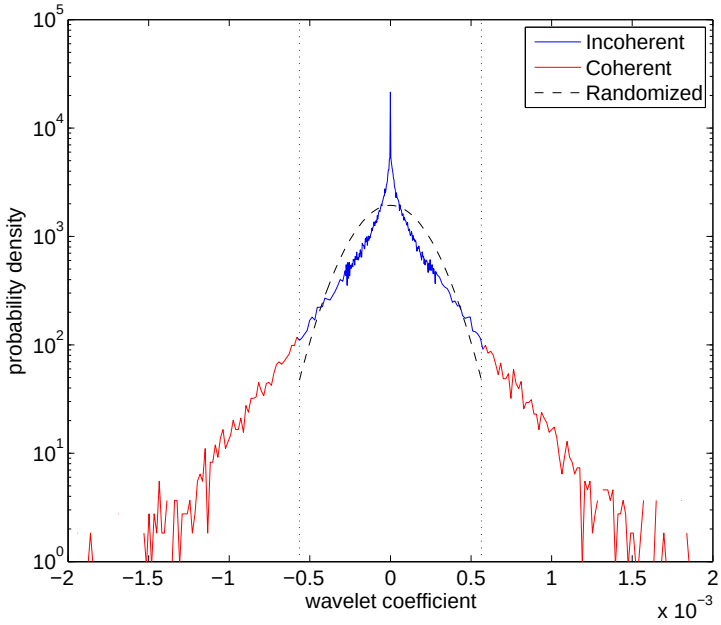


Figure V.3.13: For the HNSE reference solution at $t = 50$, comparison between the empirical PDFs of the coherent and incoherent vorticity wavelet coefficients at scale $j = 8$ (full red and blue lines) and the analytical PDF (Eq. V.3.26) obtained after maximizing the entropy (black dashed line). The vertical dotted lines indicate the interval $[-\Theta, \Theta]$, where Θ is the threshold.

which is imposed by Eq. V.3.10. s_0 can be understood as a dilation factor which compensates for the truncation of the wings of the Gaussian distribution by slightly dilating it in order to preserve its variance. Provided that $q > \sqrt{3}$, (V.3.26) admits a unique real solution which we approximate using a numerical solver. The entropy of f_λ is then:

$$S(f_\lambda) = S(f_0) + \ln(\sigma_\lambda) \tag{V.3.27}$$

and the total entropy S of the flow is then by definition the entropy of the tensor product of the f_λ for $\lambda \in \Lambda^I$, which is simply obtained from (V.3.27) by summation. In this framework, the increase of incoherent enstrophy at a given scale implies an increase of entropy, which is consistent with our interpretation in terms of dissipation.

In Fig. V.3.13, the truncated Gaussian PDF resulting from entropy maximization is shown alongside the empirical PDF of vorticity wavelet coefficients at scale $j = 8$ and for $t = 50$, as obtained from the HNSE reference solution. Notice that the truncated Gaussian PDF does not approach the empirical PDF well, since it is much less peaked around zero.

V.3.7.2 Method

We would like to find out what are the consequences of the loss of information implied by the statistical model that we have just built regarding our ability to predict the explicit flow. To do this we adopt a Monte-Carlo approach. We start from the NSE reference solution at $t = 50$ for $N = 2048$ and $\mathrm{Re} \simeq 2.66 \cdot 10^5$. Ten different realizations of the randomization

Type of CVE	Fate of incoherent part	Shorthand
Global	Discarded	GD
Global	Randomized	GR
Scale-wise	Discarded	SD
Scale-wise	Randomized	SR

Table V.3.3: Four dissipation mechanisms considered in section V.3.7.

operators are applied to the vorticity field, and the NSE are then integrated separately up to $t = 100$ for each one of them. In practice, normally distributed random numbers with standard deviation s_λ are generated using the Mersenne twister algorithm, and the $\tilde{W}_\lambda$ are obtained by retaining only those that fall within the interval $[-\Theta_\lambda, \Theta_\lambda]$. For comparison, we also consider the case where the incoherent part is completely discarded instead of being randomized. The four cases that we are going to compare are summarized in Table V.3.3. There remains to decide under what terms the comparison is to be performed.

Defining a meaningful way of comparing several turbulent flows is a research topic in itself, closely connected to the chosen statistical framework. In the setting of complete statistical modelling, only quantities characterizing the attractor of the dynamical system corresponding to the turbulent flow may be meaningfully compared, since all other properties are considered to be random. On the contrary, in a completely deterministic framework, one wishes to compare two solutions of the Navier-Stokes equations point-wise in space and in time, but that may yield little relevant information since the dynamics are known to be chaotic, i. e. unstable to perturbations. Two solutions initially very close to each other will always end up far away, even though it is not yet completely clear what influences the rate of separation. As an intermediate between these two extreme approaches, we propose to compare only the explicit flows, as defined within the framework of the conditional statistical model we are working with.

Even within a given statistical framework, there remains the question of what features of the flow are to be compared. On the one hand, we shall consider integral quantities, namely enstrophy and palinstrophy, which we know to be predictable because they have been stabilized by spatial averaging. But it is also important to assess to what extent local properties of the flow can be predicted as well. Therefore, we shall consider the measure

$$\delta_X(t) = \frac{\|\mathbf{u}_X^C(t) - \mathbf{u}^C(t)\|}{\|\mathbf{u}^C(t)\|} \tag{V.3.28}$$

where X stands for either one of the four dissipation mechanisms recalled in Table V.3.3, $\mathbf{u}_X^C(t)$ is the coherent velocity field of the perturbed solution at time t, and $\mathbf{u}^C(t)$ is the coherent velocity field of the reference solution at time t.

V.3.7.3 Results

We observe that the enstrophy of the randomized flow (Fig. V.3.14, left) decays on average faster than the one of the reference flow, while remaining close to it. The palinstrophy (Fig. V.3.14, right) gets on average larger for the randomized vorticity field at at $t = 50$, but comes back to the reference value after $t = 75$. For enstrophy, at $t = 50$, the fluctuation

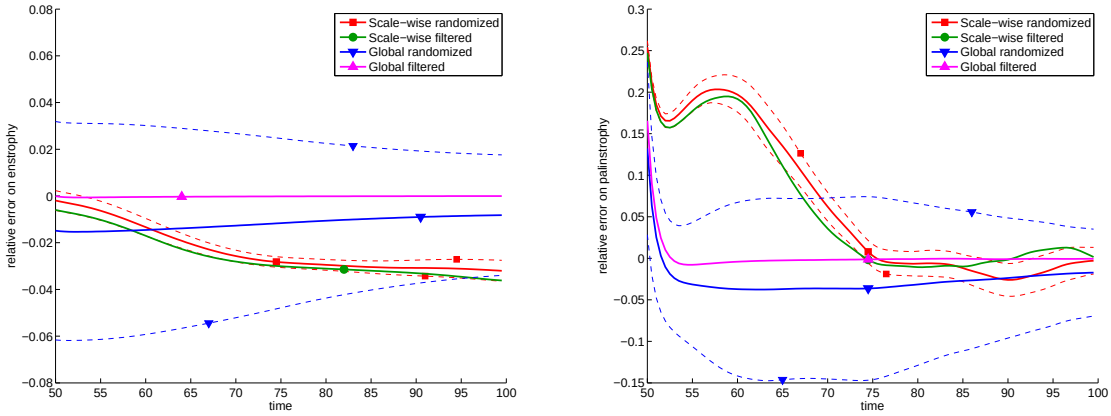


Figure V.3.14: Relative error on integral quantities as a function of time between the filtered and randomized flows and the reference flow at $N = 2048$. The error is averaged over 10 realization of the randomization operator. The dashed lines are one standard deviation away from the average. Left: enstrophy. Right: palinstrophy.

X	GD	GR	SD	SR
$\langle \delta_X(0) \rangle$	$1.5 \cdot 10^{-21}$	$2.5 \cdot 10^{-13}$	$1.1 \cdot 10^{-19}$	$2.5 \cdot 10^{-7}$

Table V.3.4: Initial value of δ_X for the four dissipation mechanisms, averaged over 10 realizations.

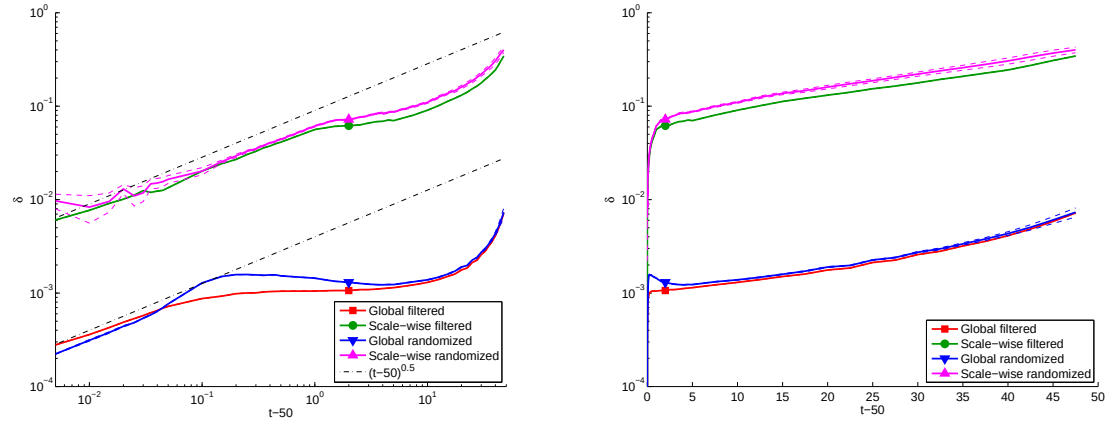


Figure V.3.15: Left: time evolution of the average error between the reference and perturbed solutions, as defined by Eq. V.3.28, on log-log scales. The dashed lines delimit an interval of one standard deviation above and below the average error, as can be estimated from 10 realizations. The dashed-dotted lines indicate the scaling $\delta(t) \propto (t - t_0)^{\frac{1}{2}}$. Right: same data, but represented using a linear scale for the horizontal axis.

level among the ensemble of randomized solutions is a few percents, and for palinstrophy it is goes up to 25%.

Let us now turn to the evolution in time of the difference between the perturbed and reference solutions, as defined by Eq. V.3.28. As expected, the initial values of δ are close to the round-off precision for SD and GD (Table V.3.4). They are also quite small for SR and GR, thanks to the constraints imposed on the randomization algorithm as described in the previous section. To lower $\delta(0)$ for SR and GR even more, we would need a microcanonical randomization operator, which we have not achieved in the present work. The evolution in time of δ (Fig. V.3.15) goes through two distinct phases which are the same for the four operators:

(i) a power-law behavior for short times, which can be roughly described by the scaling $\delta(t) \propto (t-50)^{\frac{1}{2}}$,

(ii) an exponential growth for large times.

The ratio between the errors for global CVE and for scale-wise CVE keeps the same order of magnitude 10^{-2} independent on time. Although scale-wise CVE initially attributes a larger proportion of the energy and enstrophy to the dissipated flow than global CVE, and thereby entails a bigger loss of information, the dissipated flows associated to global and to scale-wise CVE retroact on the explicit flow in proportion to their initial amplitude, and can be seen as equivalent in this respect. The singularity of $\delta(t)$ in $t = 0$ is also a property common to both approaches, and is probably linked to the lack of regularity of their respective thresholding operators.

V.3.8 Conclusion

In this paper, we have undertaken an extensive statistical analysis of 2D turbulence in the enstrophy cascade regime, by numerical study of solutions of the Navier-Stokes equations in the vanishing viscosity limit, and using the wavelet representation of the vorticity field as the essential mathematical tool. We have shown that the classical enstrophy cascade phenomenology could be recovered in the wavelet representation, by considering the scale-wise statistics, and interscale enstrophy transfers. By computing the scale-wise PDFs of the wavelet coefficients, we have been able to extract more information than the one contained in the more classical energy spectrum and structure functions. From a practical point of view, wavelet statistics are appealing because they can be obtained efficiently without computing the Fourier transform of the vorticity field, and can be generalized to non-periodic boundary conditions and adaptive grids. Another advantage which may become more and more important is the scaling efficiency of the parallel wavelet transform, which is in theory better than the one of the parallel Fourier transform.

We have introduced scale-wise coherent vorticity extraction (CVE) as a way to separate extreme events and very probable events within each scale of the flow. Formally a simple extension of the existing global CVE approach, scale-wise CVE has the advantages of being much more robust in the limit $\mathrm{Re} \to \infty$, and of being independent on numerical discretization effects. Scale-wise CVE yields an incoherent part which contains a non negligible fraction of the total enstrophy of the flow, and also a small fraction of the energy. There are good

indications that both fractions may converge to a nonzero limit when $\mathrm{Re} \to \infty$. We propose that the productions of incoherent vorticity and energy are Re-independent measures of dissipation for 2D turbulence. This dissipation should be understood in a "subjective" sense (Richardson and Gaunt, 1930), as a transfer of enstrophy between degrees of freedom that we choose to compute explicitly and degrees of freedom that we choose to model statistically. Yet once it has been defined, this dissipation can be quantified in a fully rigorous way by considering the transfers between various regions of wavelet space, as we have done in section V.3.6.2. We have shown that, contrary to global CVE for which the intrascale transfer of enstrophy in the inertial range is dominated by its coherent component, scale-wise CVE entails a non-negligible intrascale transfer of incoherent enstrophy in the inertial range.

These findings shed a new light on earlier results that have shown how the 2D and 3D Navier-Stokes (Schneider et al., 2006; Okamoto et al., 2007) and 2D Euler (Nguyen van yen et al., 2009) dynamics were well preserved by discarding after each timestep the incoherent part determined from global CVE. Our explanation is that global CVE acts mostly in the dissipation range, and does not influence the inertial range. Thanks to its nonlinear character, global CVE may stand as a more judicious choice than a classical, linear dissipation operator, since it allows for a more economical representation of the flow in the dissipation range by greatly reducing the number of necessary degrees of freedom. As a counterpart, we expect the compression rate that can be attained by global CVE to be limited by the size of the dissipation range. For example, taking the three finest scales as a rough estimate of the dissipation range in a standard, well resolved turbulent flow, the compression rate for global CVE can not raise much above 64 in 2D and 512 in 3D. Anyway, for 2D homogeneous turbulence, we have shown here that neither global CVE nor scale-wise CVE induces the pronounced increase of the compression rate with Re that would be necessary to make the approach computationally competitive. We expect that the situation will be different for 3D flows, or even for 2D flows with boundaries, which we propose to study in detail in future work.

Turbulent dissipation entails a loss of information which affects our ability to predict the explicit flow exactly. To estimate the error on the explicit flow, we have introduced a statistical model for the dissipated flow. We have first defined an ensemble of flows over which we have defined a probability measure by maximizing the entropy under minimal realistic constraints. The elements of this ensemble are all the realizations of the total flow that are compatible with the observed explicit flow, or in other words they are conditioned by the explicit flow. Adopting a Monte-Carlo approach, we have picked from this ensemble ten different perturbations of the reference flow, and we have integrated them in time. By monitoring the time evolution of enstrophy and palinstrophy, we have checked that the global properties of the ten flow realizations are close to those of the reference solution. This finding is not very surprising, since integral quantities like enstrophy and palinstrophy are generally thought to be quite stable to perturbations. To benefit from a more discriminating test, we have then considered the time evolution of the error between the perturbed and reference solutions. To measure the error, we have projected the perturbed solution using the same filter as the one used to define the perturbation. We have shown that the L^2 error on the coherent flow velocity first undergoes a power-law growth, with an exponent close to 0.5, and then a relatively slow exponential growth. The explicit flow is thus sufficient to partially predict the time evolution of the total flow for more than 10 eddy turnover times, both for global and scale-wise CVE. The performance is especially good for global CVE, for which the relative L^2 error remains below 1% on that same period.

An important limitation of the present work is that our splitting algorithms, both global CVE and scale-wise CVE, depend on the choice of a constant q (see Eq. V.3.9). We have not studied the dependency of our results on this constant, and we have not been able to provide a rigorous justification for the choice of q. We conjecture that q can be used as a control parameter to enhance or degrade the faithfulness of the representation of the flow by its coherent part. This remains an important subject for future research. More generally, one should define an objective way of deciding when a split is better than another split, although the choice of how accurate one wants the final results to be is likely to remain subjective.

An interesting perspective is to implement the split that we have proposed as a turbulence model in a numerical code. Much work has already been done on adaptive codes in wavelet bases (Schneider et al., 2006; Keetels et al., 2007; Schneider and Vasilyev, 2010), and the novelty would be to include the stochastic terms modelling the incoherent part. Our randomization technique could also benefit ensemble forecasting, whose performance depends heavily on the way the ensemble of initial conditions is constructed.

Acknowledgments

The authors would like to thank Claude Bardos, Mathias Holschneider, Rupert Klein and Edriss Titi for inspiring discussions, and Greg Hammett for an important remark concerning wavelet filtering. The three authors thank the Wissenschaftskolleg zu Berlin for its hospitality, as well as the SMAI and the organizers of CEMRACS 2010 for supporting their stay at CIRM (Marseille, France) while writing this paper, and IDRIS-CNRS for providing computing time. We acknowledge financial support from the Euratom-CEA association and the French Federation for Fusion Studies.

V.4 Analysis of 3D turbulent boundary layers

The contents of this section are reproduced from an article submitted for publication in the Proceedings of the Center for Turbulence Research, see (Khujadze et al., 2010).

V.4.1 Introduction

The motivation for this study is the importance of turbulent boundary layers in many fields of applied physics, for example, technical devices like flows around airplanes, cars or golf balls, where determining the drag coefficient is directly related to this thin layer around the obstacle. In geophysical flows, the atmospheric boundary layer also plays a prevailing role. For a review on the subject we refer to the classical text book by (Schlichting, 1979). Model free numerical simulations of turbulent boundary layers are still a tough problem and constitutes a major challenge in computational fluid dynamics for both the numerical discretization schemes and the computer resources. The stiffness is due to the very high resolution near the wall which is required to resolve all dynamically active scales of the flow. (Spalart, 1988) did the first numerical simulations of turbulent boundary layers. Over the past years a number of simulations of such flows for higher Reynolds numbers became available, cf. (Skote, 2001; Khujadze and Oberlack, 2004, 2007; Simens et al., 2009; Schlatter et al., 2009). One important research subject is the identification and extraction of coherent structures in turbulent boundary layers. This is inspired by the existence of horseshoe vortices first observed by (Theodorsen, 1952). The observation of forests of horseshoe vortices in experimental data by (Adrian et al., 2000) and in direct numerical simulations (DNS) recently performed by (Wu and Moin, 2009) gave a second breath to this topic.

Wavelet techniques have been developed for more than 20 years (see, e.g., (Farge, 1992) for an early review) to analyze, model and compute turbulent flows. The multiscale representation obtained by wavelet decompositions is lucrative to understand the physics of turbulent flows as locality in both space and scale is preserved. Thus localized features of turbulent flows, like coherent structures and intermittency can be extracted and analyzed. Coherent Vorticity Extraction (CVE) has been introduced for two- and three-dimensional turbulent flows in (Farge et al., 1999) and (Farge et al., 2001), respectively. The underlying idea is that coherent structures are defined as what remains after denoising and hence only a hypothesis on the noise has to be made. In the present study we suppose the noise to be Gaussian and white. Preliminary results of CVE applied to wall bounded flows, for a channel flow, have been reported in (Weller et al., 2006). Scale dependent and directional statistics in wavelet space have been presented in (Bos et al., 2007) to quantify the intermittency of anisotropic flows. Sheared and rotating flows have been analyzed recently in (Jacobitz et al., 2010). An up to date review on wavelet techniques in computational fluid dynamics can be found in (Schneider and Vasilyev, 2010).

In the present paper we apply orthogonal wavelet analysis for the first time to DNS data of turbulent boundary layers. Additional difficulties are encountered due to the non equidistant grid in the wall normal direction. The aim of the paper is to extract coherent structures out of high resolution DNS of zero pressure gradient turbulent boundary layer flow at $Re_\theta \approx 1500$. The total flow is decomposed into coherent and incoherent vorticities and scale dependent statistics, i.e., variance, flatness and probability distribution functions, are computed at different wall normal positions. The performed analyses are a first step as they

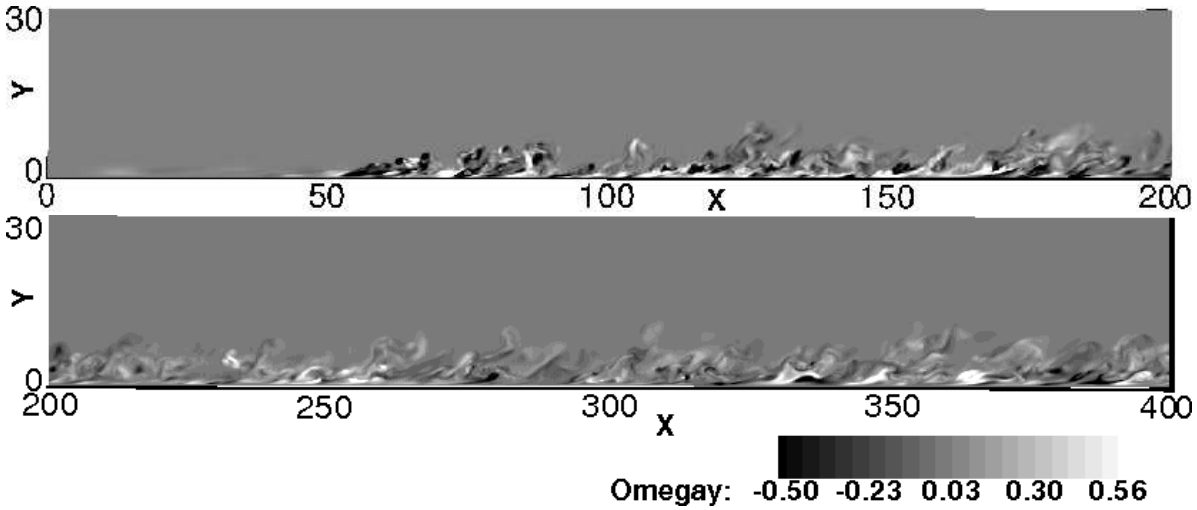


Figure V.4.1: Two dimensional slices of ω_y at $z = 0$; $x \in [0, 200]$ (top) and for $x \in [200, 400]$ (bottom).

are in the present work limited to flow snapshots. A detailed investigation of the dynamics of the coherent and incoherent flow contributions is left for future work.

The paper is organized as follows. Section 2 presents the flow configuration and the computational approach. Some visualization and analyzes of the DNS data are also given. CVE methodology is described in section 3, mentioning technical details like the required interpolation on a dyadic adapted grid, the adaptive anisotropic wavelet transform and the wavelet based statistics which are applied in the numerical results section 4. The latter discusses the total, coherent and incoherent flows using both flow visualization and statistical analyzes. The efficiency of CVE is also assessed. Finally, conclusions are drawn in section 5 and some perspectives for future investigations are given.

V.4.2 Flow configuration and parameters

The DNS code for solving the incompressible Navier–Stokes equations was developed at KTH, Stockholm, we refer to (Lundbladh et al., 1999) for details. A spectral method with a Fourier decomposition is used in the horizontal directions while a Chebyshev discretization is applied in the wall normal direction. The time integration is performed using a third order Runge-Kutta scheme for the advective and forcing terms and a Crank-Nicolson scheme for the viscous terms. Since the boundary layer is developing in downstream direction the fringe region (where the outflow is forced by a volume force to the laminar inflow Blasius profile) has to be added to the physical domain to satisfy periodic boundary condition. A wall-normal trip force is used to trigger the transition to turbulence.

The spectral method as a computational tool has evolved vigorously in the early 1970's. This strategy was also applied successfully for simulating zero pressure-gradient (ZPG) turbulent boundary layers by (Spalart, 1988) where the growth of the boundary layer thickness in the streamwise direction is small. Extensive studies of turbulent boundary layer flows were performed by (Skote, 2001). Here we give some details about the simulations used in our study. DNS of ZPG turbulent boundary layer flow was performed for two different number

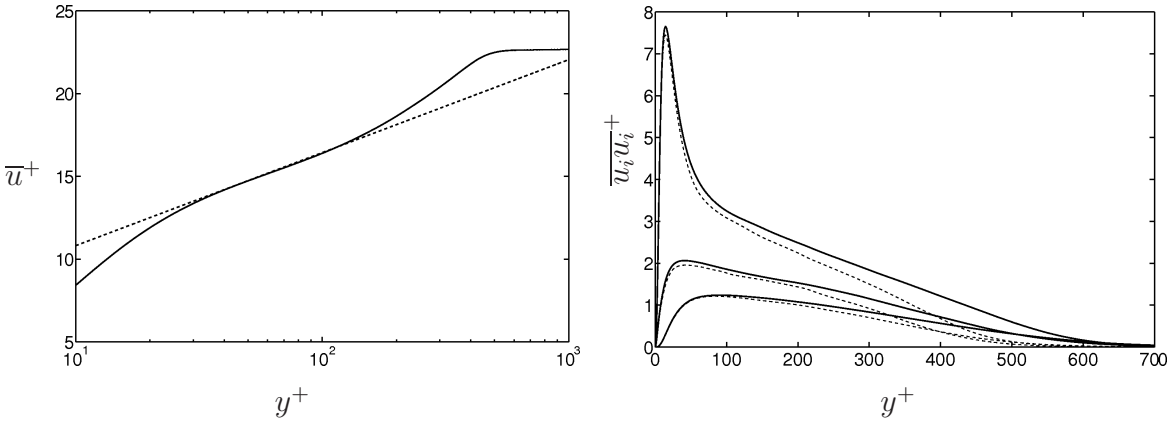


Figure V.4.2: *Left plot*: Mean velocity profile in lin-log scaling at $Re_\theta = 1500$. --- $1/\kappa\ log\ y^+ + B$ with $\kappa = 0.41$ and $B = 5.2$; *Right plot*: Diagonal components of Reynolds stress tensor for the present DNS (---) and from (Simens et al., 2009) (——) at $Re \approx 1500$.

of grid points $N_x \times N_y \times Nz = 2048 \times 513 \times 256$ at starting laminar Reynolds number $Re_{\delta^*}|_{x=0} \equiv \frac{u_\infty \delta^*|_{x=0}}{\nu} = 600$. All quantities were non-dimensionalized by the free-stream velocity u_∞ and the displacement thickness δ^* at $x = 0$ where the flow is laminar. The size of the computational box was $L_x \times L_y \times L_z = 1000\delta^*|_{x=0} \times 30\delta^*|_{x=0} \times 34\delta^*|_{x=0}$ (Fig. V.4.1 represents the part of computational box. x, y, z are axes correspondingly in streamwise, wall-normal and spanwise directions). The simulations were run for a total of 11500 time units ($\delta^*|_{x=0}/u_\infty$). The turbulent domain was $Re_\theta \approx 500 - 1500$ (Reynolds number $Re_\theta = \frac{u_\infty \theta}{\nu}$ where θ is the the momentum loss thickness). The grid resolution in viscous or plus units ($\Delta x^+ \equiv \Delta x/u_\tau \nu$, where u_τ is the friction velocity) was $\Delta x^+ \times \Delta y^+_{max} \times \Delta z^+ = 12.8 \times 5 \times 3.5$.

Fig. V.4.2 (left) shows the mean velocity profile in lin-log scaling. The dashed line corresponds to the fit of the classical log-law. The solid line represents the present DNS at $Re_\theta = 1500$. The right plot shows the comparison of Reynolds diagonal stresses from our and (Simens et al., 2009) data.

V.4.3 Orthogonal wavelet decomposition of the turbulent boundary layer flow

In the following we introduce a new anisotropic wavelet decomposition with an adaptive grid in the wall normal direction which allows for the analysis of the DNS data. Then, the coherent vorticity extraction is presented and different scale dependent wavelet based statistics are described.

V.4.3.1 Adaptive anisotropic wavelet decomposition

From the velocity field $\boldsymbol{u} = (u_1, u_2, u_3)$ we compute the vorticity field $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3) = \nabla \times \boldsymbol{u}$. Both fields are given on discrete grid points (x_i, y_n, z_k) for $i = 1, ..., N_x, n = 1, ..., N_y$ and $k = 1, ..., N_z$. The grid is equidistant in the wall parallel directions x and z, while in the wall normal direction y a Chebychev grid is used, i.e., $y_n = \cos(\pi\theta_n)$ with

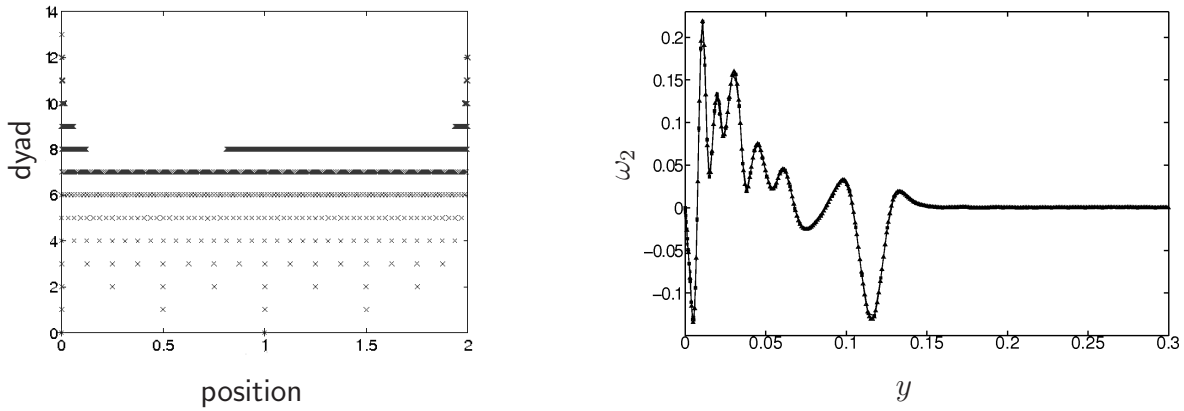


Figure V.4.3: *Left plot:* Adapted dyadic grid by the position of the corresponding wavelets. *Right plot:* The result of the interpolation in wall-normal directions. ⊟ Original/Chebyshev grid, △ dyadic grid, × reinterpolated on Chebyshev grid.

$\theta_n = (n-1)/N_y - 1$ for $n = 1, \ldots, N_y$.

Each component ω_ℓ of the vorticity vector is then decomposed into a two-dimensional orthogonal wavelet series in the wall parallel directions x-z using a two-dimensional multiresolution analysis. The number of scales J_{xz} is defined as the maximum integer such that $N_x = k2^{J_{xz}}$ and $N_z = k'2^{J_{xz}}$ where k and k' are any integers. For a fixed wall normal position y_n we thus obtain for $\ell = 1, 2, 3$,

$$\omega_\ell(x, y_n, z) = \sum_{j_{xz}=0}^{J_{xz}-1} \sum_{i_x=0}^{2^{J_{xz}}-1} \sum_{i_z=0}^{2^{J_{xz}}-1} \sum_{\mu=1}^{3} \langle \omega_\ell(y_n), \psi^\mu_{j_{xz},i_x,i_z} \rangle_{xz} \, \psi^\mu_{j_{xz},i_x,i_z}(x, z) \qquad \text{(V.4.1)}$$

with the wavelet

$$\psi^\mu_{j_{xz},i_x,i_z}(x, z) = \begin{cases} \psi_{j_{xz},i_x}(x)\phi_{j_{xz},i_z}(z) & \text{for } \mu = 1 \\ \phi_{j_{xz},i_x}(x)\psi_{j_{xz},i_z}(z) & \text{for } \mu = 2 \\ \psi_{j_{xz},i_x}(x)\psi_{j_{xz},i_z}(z) & \text{for } \mu = 3 \end{cases} \qquad \text{(V.4.2)}$$

where ϕ and ψ are the one-dimensional scaling function and wavelet, respectively and $\mu = 1, 2$ and 3 corresponds to the direction of wavelets in the x, y and xy direction, respectively. The scalar product is defined in the x-z plane, $\langle f, g \rangle_{xz} = \int f(x, z) g(x, z) dx dz$. Here we use Coiflet 12 (see, e.g. (Farge, 1992)) wavelets and the scaling coefficients on the finest scale are identified with the grid point values.

Before performing a one-dimensional wavelet transform in the y-direction (while fixing the x-z direction), the vorticity components ω_ℓ have to be interpolated from the Chebychev grid onto a locally refined dyadic grid. For that a Lagrange interpolation of 4-th order is used and a Haar wavelet transform is applied to the Chebychev grid $\arccos(y_n)$ to define the locally refined dyadic grid $\tilde{y}_n = i_y/2^{j_y}$ for $j_y = 0, ..., J_y - 1$ and $i_y = 0, ..., 2^{j_y} - 1$ using nonlinear approximation. The number of grid points in the y-direction is fixed, here to $\tilde{N}_y = 1024$. The maximal scale in y-direction, J_y, is then determined from the Haar wavelet analysis retaining the N_y strongest coefficients. In the present case we obtain $J_y = 13$. The resulting dyadic grid is shown in Fig. V.4.3, left which yields the best approximation of the Chebychev grid using a dyadic grid with $\tilde{N}_y = 1024$ grid points. The one-dimensional vorticity cuts in the y-direction in Fig. V.4.3 show the original data on the Chebychev grid, the

data interpolated onto the refined dyadic grid and after reinterpolation onto the Chebychev grid. The agreement between the curves is satisfactory and thus we can conclude that the interpolation between the different grids can be performed with little loss of information.

A wavelet decomposition using Daubechies 4 wavelets, (Farge, 1992) is then applied to the data on the adaptive dyadic grid and the scaling coefficients on the finest scales are computed using a quadrature rule. Thereafter an adaptive wavelet transform is performed on the adaptive dyadic grid and we obtain a full wavelet decompositon in all three space directions,

$$\omega_\ell(x,y,z) = \sum_{j_{xz}=0}^{J_{xz}-1} \sum_{j_y=0}^{J_y-1} \sum_{i_x=0}^{2^{J_{xz}}-1} \sum_{i_y=0}^{2^{J_y}-1} \sum_{i_z=0}^{2^{J_{xz}}-1} \sum_{\mu=1}^{3} \widetilde{\omega}^{\ell,\mu}_{j_{xz},j_y,i_x,i_y,i_z}\, \psi^{\mu}_{j_{xz},i_x,i_z}(x,z)\, \psi_{j_y,i_y}(y) \tag{V.4.3}$$

for $\ell = 1,2,3$. Note that the wavelet coefficients

$$\widetilde{\omega}^{\ell,\mu}_{j_{xz},j_y,i_x,i_y,i_z} = \int\int\int \omega_\ell(x,y,z)\psi^{\mu}_{j_{xz},i_x,i_z}(x,z)\,dxdz\,\psi_{j_y,i_y}(y)dy$$

contain different scales in the wall parallel (x-z) and the wall normal (y) direction. This property allows to take into account the anisotropy of the structures observed in the DNS data.

V.4.3.2 Coherent vorticity extraction

The starting point of the coherent vorticity extraction is the wavelet representation of vorticity in eq. (V.4.3). The underlying idea is to perform denoising of vorticity in wavelet coefficient space.

Thresholding the wavelet coefficients then determines which coefficients belong to the coherent and to the incoherent contributions. The latter are supposed to be noise like. First we compute $\Omega = \left(\sum_{\ell=1}^{3} \left(\widetilde{\omega}^{\ell,\mu}_{j_{xz},j_y,i_x,i_y,i_z} \right)^2 \right)^{1/2}$ and then we reconstruct the coherent vorticity $\boldsymbol{\omega}_c$ from those wavelet coefficients for which $\Omega > \epsilon$ using eq. (V.4.3). The incoherent vorticity $\boldsymbol{\omega}_i$ is obtained from the remaining weak wavelet coefficients. In the first iteration the threshold ϵ is determined from the total enstrophy $Z = \frac{1}{2}\langle \boldsymbol{\omega}\cdot\boldsymbol{\omega}\rangle_{xyz}$ and the total number of grid points $N = N_x \tilde{N}_y N_z$, i.e., $\epsilon = \sqrt{4Z\ln N}$. Subsequently a new threshold is determined using instead of the total enstrophy the incoherent enstrophy computed from the weak wavelet coefficients. Then the thresholding is applied again and improved estimators of the coherent and incoherent vorticities are obtained. For more details on the iterative procedure we refer to (Farge et al., 1999). We note that thanks to the orthogonality of the decomposition the enstrophy and thus the threshold can be directly computed in coefficient space using Parseval's relation. Only at the end of the iterative procedure the coherent and incoherent vorticies are reconstructed by inverse wavelet transform in physical space, in x-z direction on a regular grid and in y direction on the locally refined dyadic grid. After the vorticity fields are reinterpolated in y-direction onto the Chebychev grid.

Finally, we thus obtain $\boldsymbol{\omega} = \boldsymbol{\omega}_c + \boldsymbol{\omega}_i$ and by construction we also have $Z = Z_c + Z_i$. For future work we anticipate that the corresponding velocity fields can also be reconstructed by applying Biot-Savart's kernel which necessitates the solution of three Poisson equations.

V.4.3.3 Scale dependent statistics

The wavelet based scale dependent statistics are built on the two-dimensional wavelet representation (V.4.1) for a fixed position in the wall normal direction y_n. First we define scale dependent p-th order moments of the three vorticity components ω_ℓ in wavelet coefficient space by,

$$M^{p,\ell}_{j_{xz}}(y_n) \;=\; \sum_{i_x=0}^{2^{j_{xz}}-1} \sum_{i_z=0}^{2^{j_{xz}}-1} \left(\langle \omega_\ell(y_n), \psi^{\mu}_{j_{xz},i_x,i_z} \rangle_{xz}\right)^p \tag{V.4.4}$$

The scale dependent variance of the vorticity components corresponds to $p = 2$. Dividing it by two the scalogram of enstrophy is obtained, which yields a scale distribution of enstrophy for a given position y_n.

Scale dependent flatness of each vorticity component ω_ℓ can be defined as

$$\mathcal{F}^{\ell}_{j_{xz}}(y_n) \;=\; \frac{M^{4,\ell}_{j_{xz}}(y_n)}{\left(M^{2,\ell}_{j_{xz}}(y_n)\right)^2} \tag{V.4.5}$$

Note that for Gaussian statistics the flatness equals three on all scales. The flatness (V.4.5) quantifies the intermittency of the flow and is directionly related to spatial fluctuations in the x-z plane of the enstrophy, as shown in (Bos et al., 2007). Hence increasing flatness values for finer scale are an indicator for intermittency.

Finally, we also consider the probability distribution functions (pdfs) of the wavelet coefficients for a given scale j_{xz} at a given position y_n estimated by histograms using 128 bins. As the number of wavelet coefficients decreases at each larger scale by a factor 4 we only consider the last three scales $j_{xz} = 6, 7, 8$ in order to have sufficient statistics.

V.4.4 Numerical results

Visualizations of two-dimensional cuts of the wall-normal vorticity component ω_y are shown in Fig. V.4.4 for the total (top), coherent (middle) and incoherent vorticity (bottom). It can be observed that the coherent vorticities present in the total field are well preserved in the coherent field using only 0.84 % of the total number of wavelet coefficients, which retain 99.61 % of the total enstrophy of the flow. In contrast the incoherent one has weaker amplitude and is almost structureless.

The statistics of the different flow contributions are quantified in Fig. V.4.5 by considering the second order moments and the flatness as a function of scale of the vorticity component ω_z for the total, coherent and incoherent flows at two different wall distances, at $y^+ = 30$ which is at the beginning of the log-layer and at $y^+ = 153$ which is inside the log-layer. The variance illustrates the good agreement between the total and coherent vorticity, while the variance of the incoherent one is more than 3 orders of magnitude smaller. The latter also only weakly depends on scales which indicates an equipartition of enstrophy and thus confirms the incoherent part to be close to white noise. The flatness for both the total and coherent vorticity increases with scale which is a signature of intermittency. The flatness of the incoherent part features values around 3, which is characteristic for Gaussian noise. The probability distribution functions of the wavelet coefficients at scales $j_{xz} = 6$ and 7 are plotted in Fig. V.4.6 at two different wall distances. Close to the wall, for $y^+ = 30$

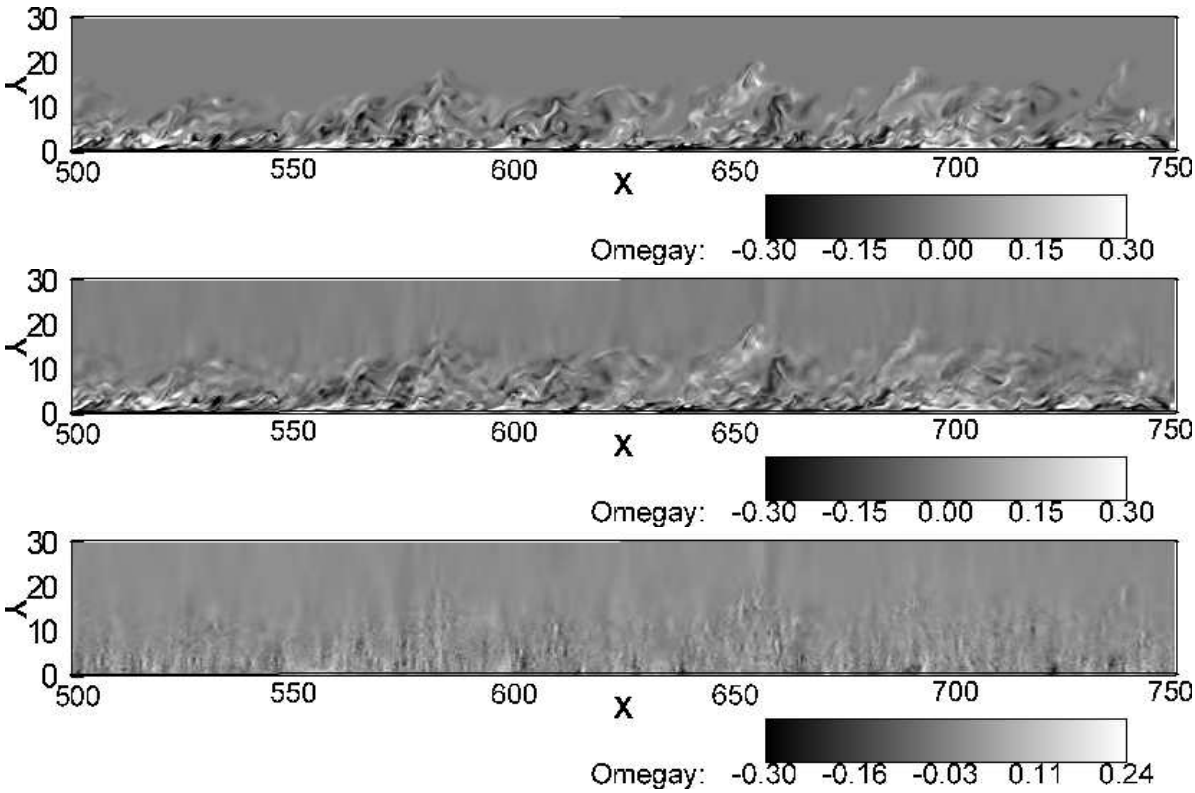


Figure V.4.4: Two dimensional slices of wall-normal vorticity (ω_y) at $z = 0$: total (top), coherent (meadle) and incoherent (bottom) parts.

we observe an algebraic decay of the pdf tails with slope -2, which is close to a Cauchy distribution and corresponds to strong intermittency. For distances further away from the wall ($y^+ = 153$) the tails of the pdf become exponential which shows that the flow becomes less intermittent. We can also observe that the pdfs do not differ much for the two scales considered here.

V.4.5 Conclusions and perspectives

Zero pressure gradient turbulent boundary layer flows have been studied by means of high resolution DNS. The flow data were shown to be in agreement with previous DNS results by (Simens et al., 2009) and visualizations confirmed the existence of horseshoe vortices.

A new adaptive wavelet transform has been developed which accounts for the flow anisotropy by using different scales in the wall normal and wall parallel directions. Coherent vorticity extraction has been applied and the obtained results showed that few ($< 1\%$) wavelet coefficients are sufficient to retain the coherent flow structures, while the large majority of the coefficients corresponds to the incoherent background flow which is unstructured and noise like. Scale dependent statistics quantified the total, coherent and incoherent flows for different wall normal positions and showed that the statistics of the total and coherent flows are in good agreement. The scale dependent flatness allowed to quantify the flow intermittency at different wall distances.

The current work is limited to snapshots of vorticty. The reconstruction of the velocity

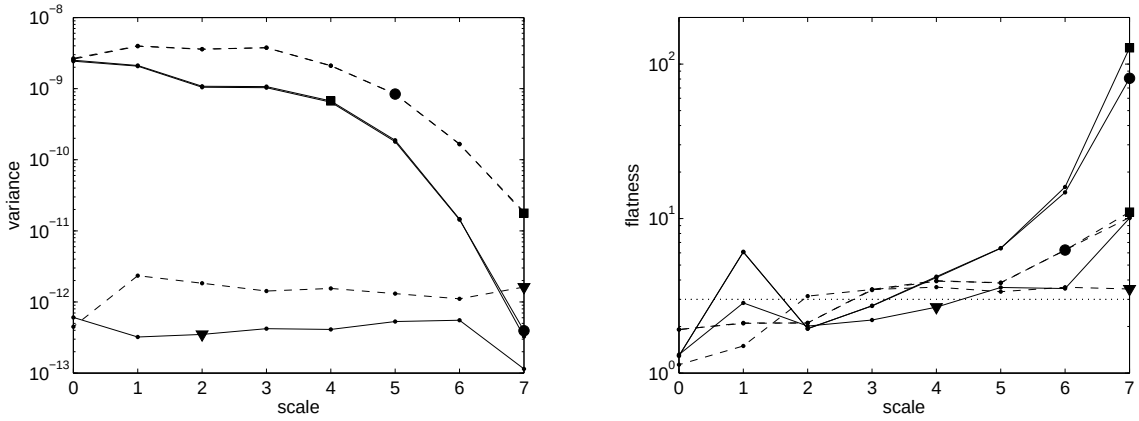


Figure V.4.5: Scale-dependent 2nd order moments (left) and scale-dependent flatness (right) of the vorticity component ω_z for the total (circles), coherent (squares) and incoherent (triangles) flows at two different wall distances, $y^+ = 30$ (——) and $y^+ = 153$ (---).

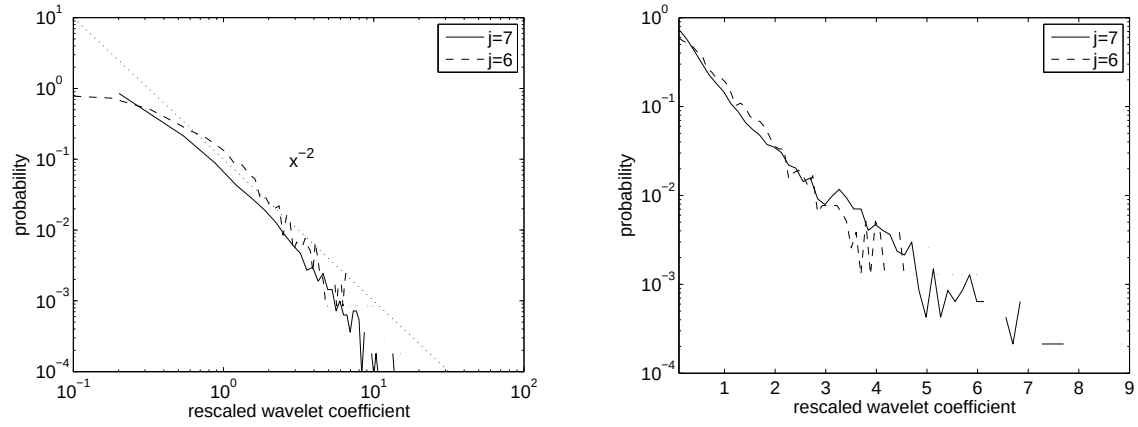


Figure V.4.6: Probability distribution functions of the wavelet coefficients of ω_z at scales $j_{xz} = 6$ and 7 estimated by histograms using 50 bins at two different wall distances, $y^+ = 30$ (left) and $y^+ = 153$ (right).

fields from the total, coherent and incoherent vorticity is a prerequisite to perform dynamical analyses of the flow, like determining the energy transfer between the different flow contributions. In future work we plan to rerun simulations initialized either with the total, coherent or incoherent flow. At longer term perspective we also envisage to perform Coherent Vorticity Simulation of turbulent boundary layer flows.

GKh, MO and KS acknowledge financial support and hospitality from the Center for Turbulence Research, Stanford University/NASA Ames, during the summer program 2010. MF and KS thankfully acknowledge financial support from the ANR, project M2TFP and MF thanks the Wissenschaftskolleg zu Berlin.

Conclusion

In this thesis, we have presented numerical results on the topic of dissipation in fluid and plasma flows, for the regimes where the microscopic coupling coefficients (respectively viscosity and collisionality) tend to zero. In Chapter 1, we have outlined the many common points between the two systems. However, some obstinate differences in the present formulation of the problems have forced us to separate the results concerning plasma flows (Chapter 3) and those concerning fluid flows (Chapters 4 and 5). To introduce necessary concepts and tools, we have preceded the presentation with a technically oriented chapter on wavelets (Chapter 2). The more unified version of the story that we aim at remains to be written, and we do not have the elements to make a convincing case here. Instead, we would like to rephrase the main results using a slightly subjective classification, between standalone results that deserve to be extracted and become independent of the tools and approaches that were used to obtain them, and result which remain too much tied to specific technical tools and make sense only in the course of the ongoing research program.

In the category of standalone results, a first theme is the regularization of conservative schemes for fluid flows. The behavior in the vanishing viscosity limit of solutions to the viscous Burgers equation can be recovered by wavelet regularization of a Fourier-Galerkin scheme applied to the inviscid Burgers equation (Sec. IV.1). Moreover, the solutions to the regularized equations converge strongly to the exact solutions in the limit of infinite resolution. The regularization can be achieved with real-valued orthogonal wavelets or with complex-valued Kingsbury wavelets (Sec. IV.2.3.1), but the latter behave better thanks to their translation invariance. These results fit nicely among the existings works concerning regularization of the Burgers equation by various methods. Interestingly, analogous results hold in the context of the 2D incompressible Euler equations (Sec. IV.2.3.2). However, in that case the solution to the conservative Galerkin-truncated equations appear to converge towards solutions to the full equations, in sharp contrast to what happens for the Burgers equation. Very recently, Bardos and Tadmor (2010) have announced that they have obtained a mathematical proof of this convergence, which sheds more light on the specific properties of the incompressible Euler equations with respect to the Burgers equation.

The second theme is dissipation at vanishing viscosity. For the dipole-wall collision, a generic event observed in turbulent two-dimensional incompressible flows in contact with walls (Sec. V.2), there is numerical evidence that the molecular dissipation does not go to zero in the vanishing viscosity limit, due to the production of a boundary layer with a thickness scaling like Re^{-1}, that we have called the Burgers-Kato layer. Moreover, energy dissipation continues inside spiral like structures that detach from the wall. This gives some hints on how to construct weak solutions to the Euler equations that dissipate energy, whose existence was conjectured long ago by Onsager. In the very different, but also idealized, test case of homogeneous isotropic 2D decaying turbulence (Sec. V.3), another kind of dissipation is observed. Indeed, using wavelets, the flow can be split within each scale between coherent and incoherent degrees of freedom, which correspond respectively to the tails and to the core of the probability density functions of the wavelet coefficients of the vorticity field (scale-wise PDFs). In the vanishing viscosity limit, the nonlinear interactions induce the transfer of a non-zero fraction of the energy and of the enstrophy of the flow from the coherent to the the incoherent part, a phenomenon which we have interpreted as a macroscopic dissipative process. Moreover, the statistics of the vorticity field are well captured by its scale-wise PDFs, which are nearly self-similar in the inertial range with a flatness close to 7, intermittent in

the molecular dissipation range, and have pronounced exponential tails at all scales. These precise features suggest that it is possible to construct statistical models of 2D homogeneous turbulence based on the wavelet coefficients of the vorticity field.

We consider that the remaining results are still in the intermediate stage. This includes in particular what we have obtained concerning plasma flows, namely the regularization of the electric field in particle solvers for the Vlasov-Poisson equations, using the wavelet-based density estimation method we have developed (Sec. III.1). Indeed, we have proven that particles subject to the regularized electric field do not exert forces on themselves (Sec. III.2), and we have demonstrated the convergence of the solution by direct comparison with numerical solutions obtained using well established schemes (particle-in-cell and semi-Lagrangian), but we have not yet adressed the key questions of entropy production and macroscopic mixing in the vanishing collisionality limit of plasma flows, which would lead to the plasmas counterparts of the results we have presented for fluid in Chapter 5. Concerning the inversion of tokamak edge plasma light emission (Sec. II.4), we have presented a proof of concept which seems encouraging, but which needs to be validated against alternative edge plasma diagnostics before it can be used in practice. The method could then be implemented to analyze the movies recorded by the fast cameras that have recently been installed on several tokamaks in the world, in particular Tore Supra (CEA, Cadarache). In the present configuration, the code can process movies where each frame contains up to about 10^5 pixels, so that meaningful comparisons with other experimental diagnostics can be envisaged in the near future.

At the end of Sec. III.1, we have proposed to model turbulent dissipation in 2D homogeneous fluid flow turbulence by replacing the incoherent part by a random process whose probability distribution can be obtained by maximizing an entropy functional. Using this model, the L^2 error on the coherent flow velocity with respect to the exact flow first undergoes a power-law growth, with an exponent close to 0.5, and then a relatively slow exponential growth. This model needs to be explored further, both for the specific system of 2D homogeneous turbulence, and in more general contexts, for example 3D homogeneous turbulence. It would also be natural to extend it to plasma flows, but this remains a middle term perspective. Finally, we have obtained very preliminary results towards an extension of the results presented in Chapter 5 to the 3D case, by analyzing a 3D turbulent boundary layer flow defined on an irregular grid (Sec. V.4). The completion of this project will however require much more work, given the computational difficulty of the problem.

***

To arrive at these results we had to develop some new tools which we now summarize. The parallel wavelet transform library, described in Chapter 2, was perhaps the most essential, and moreover its range of applicability is quite wide. Indeed, it can be applied in any space dimension, and with many wavelet families, including complex-valued Kingsbury wavelets. Its portability and good parallel scalability ensure that it will benefit from future increase in computer power. The adaptive version can be used to analyze flows defined on irregular grids, even though at present the tool requires that the grid be irregular in at most one direction, and homogeneous in all other directions. We have also derived an algorithm to actively refine or coarsen the grid in order to optimize the representation of a given function. This tool could be used in the future to develop adaptive solvers, which could have a positive impact on various fields such as atmosphere modeling. At the moment the

main obstacles on this path are the relatively poor performance, which could be tackled by a more extensive optimization, and the extension of the adaptivity to several dimensions, which requires advanced algorithms.

We have also brought some contributions to three existing wavelet thresholding algorithms, namely wavelet-vaguelette decomposition (WVD), wavelet-based density estimation (WBDE) and coherent vorticity extraction (CVE). WVD was implemented and tested in the specific case of the helical Abel transform, a generalization of the Abel transform that we have introduced to deal with the tokamak geometry (Sec. II.4). WBDE was extended to several space dimensions, and we have provided specific values of the parameters by testing it in the plasma context (Sec. III.1). CVE was parallelized and generalized into scale-wise CVE in order to cope with the case where the incoherent noisy part is correlated instead of white (Sec. V.3.5.1). Other, more specific tools were needed in some places. For example, in order to solve the Navier-Stokes equations in the presence of walls at high Reynolds number, we have modified the volume penalization method by using a mollified mask function, which allows for a better control of energy dissipation by the discrete solution (Sec. V.1).

All these tools are interfaced with a C++ platform we have developed for solving the 1D Burgers and 2D Navier-Stokes equations using parallel spectral methods, and the Vlasov-Poisson equations using particle methods. Thanks to this unified platform, the data is compatible, and many tools can be used in several places to avoid code redundancy. The C++ platform is available to the community under an open source license at the following address

http://justpmf.com/romain

and can be used in particular to reproduce all the results we have obtained. In the future the documentation will be completed and the interface will be made more user-friendly.

***

Practical applications in broader contexts of the results we have obtained are possible but will require non-trivial developments. Experimental tests for the existence of dissipative structures in 2D flows with walls would be of very much interest, especially since their outcome is highly uncertain given that the 2D approximation with no-slip boundary conditions may not be accurate to describe realistic flows. One approach could be to follow tracer particles in soap films. Another, more hazardous approach, could be to analyze the trajectories of drifting buoys in the oceans, but those are affected by the Earth's rotation and gravity waves, which were not taken into account in our study. Another range of application could be ensemble weather forecast, for which the dissipation model we have proposed in Sec. V.3 could serve to generate initial conditions. More generally, it could be integrated in existing approaches that aim to study the predictability of various types of turbulent flows.

Before the range of applications can be extended further, the remaining questions belonging to the themes we have worked on should be adressed. With the tools we have developed, it is already possible to apply scale-wise CVE to 3D homogeneous turbulence, and also to analyze in detail its scale-wise statistics. This will provide insight into the macroscopic dissipative processes playing an essential role in 3D turbulence. We know at a qualitative level that in the 3D case, there is an interplay between these macroscopic processes and molecular

dissipation which has been shown experimentally to remain finite at vanishing viscosity. The wavelet analysis, using the tools that we have developed during this thesis, could allow for a more quantitative description of this interplay. In the 2D case, we could also bring together the two ingredients that were studied separately up to now, namely the influence of walls and the homogeneous turbulence. Finally, the matter of dissipation in the vanishing collisionnality limit of plasma flows is very interesting. We expect that it is possible to isolate dissipative structures and macroscopic entropy production, much like in the Navier-Stokes case.

***

References

H. ABIDI and R. DANCHIN. Optimal bounds for the inviscid limit of Navier-Stokes equations. *Asymptotic Analysis*, **38**(1):35–46, 2004.

P. ABRY and D. VEITCH. Wavelet analysis of long-range-dependent traffic. *IEEE J. Information Theory*, **44**(1):2–15, 1998.

P. ABRY, P. GONÇALVES, and P. FLANDRIN. Wavelet-based spectral analysis of $\frac{1}{f}$ processes. In P. GONCALVES, editor, *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing ICASSP-93*, volume 3, pages 237–240 vol.3, 1993.

R. J. ADRIAN, C. D. MEINHART, and C. D. TOMKINS. Vortex organization in the outer region of the turbulent layer. *J. Fluid Mech*, **422**:1–54, 2000.

P. ANGOT, C.-H. BRUNEAU, and P. FABRIE. A penalization method to take into account obstacles in incompressible viscous flows. *Num. Math.*, **81**:497–520, 1999.

F. ANSELMET, Y. GAGNE, E. J. HOPFINGER, and R. A. ANTONIA. High-order velocity structure functions in turbulent shear flows. *J. Fluid Mech.*, **140**:63–89, 1984.

E. ARQUIS and J. CALGAGIRONE. Sur les conditions hydrodynamiques au voisinage d'une interface milieu fuide - milieu poreux: application a la convection naturelle. *C. R. Acad. Sci. Paris, Série II*, **299**(1):1–4, 1984.

S. P. AUERBACH. Collisional damping of Langmuir waves in the collisionless limit. *Phys. Fluids*, **20**:1836–1844, 1977.

M. AVELLANEDA and E. WEINAN. Statistical properties of shocks in Burgers turbulence. *Comm. Math. Phys.*, **172**:13–38, 1995.

A. Y. AYDEMIR. A unified monte carlo interpretation of particle simulations and applications to non-neutral plasmas. *Phys. Plasmas*, **1**(4):822–831, 1994.

A. AZZALINI, M. FARGE, and K. SCHNEIDER. Nonlinear wavelet thresholding: a recursive method to determine the optimal denoising threshold. *Appl. Comput. Harmon. Anal.*, **18** (2):177–185, 2004.

A. BABIANO, C. BASDEVANT, and R. SADOURNY. Structure functions and dispersion laws in two-dimensional turbulence. *J. Atmos. Sci.*, **42**(9):941–949, 1985.

R. BALIAN. *From Microphysics to Macrophysics*. Springer, 2006.

C. BARDOS and E. TADMOR. Private communication, 2010.

C. BARDOS and E. TITI. Euler equations for incompressible ideal fluids. *Russian Math. Surveys*, **62**(3):409–451, 2007.

C. BARDOS, J. S. LINSHIZ, and E. S. TITI. Global regularity for a Birkhoff-Rott-α approximation of the dynamics of vortex sheets of the 2D Euler equations. *Physica D*, **237**(14–17):1905 – 1911, 2008.

C. BASDEVANT and R. SADOURNY. Ergodic properties of inviscid truncated models of two-dimensional incompressible flows. *J. Fluid Mech.*, **69**(04):673–688, 1975.

C. Basdevant, B. Legras, R. Sadourny, and M. Béland. A study of barotropic model flows: intermittency, waves and predictability. *J. Atmos. Sci.*, **38**:2305–2326, 1981.

D. Batchelor, M. Beck, A. Becoulet, R. Budny, C. Chang, P. H. Diamond, J. Dong, G. Y. Fu, A. Fukuyama, T. S. Hahm, D. E. Keyes, Y. Kishimoto, S. Klasky, L. Lao, K. Li, Z. Lin, B. Ludaescher, J. Manickam, N. Nakajima, T. Ozeki, N. Podhorszki, W. M. Tang, M. A. Vouk, R. E. Waltz, S. J. Wang, H. Wilson, X. Xu, M. Yagi, and F. Zonca. Simulation of fusion plasmas: Current status and future direction. *Plasma Science and Technology*, **9**(3): 312–387, 2007.

G. K. Batchelor. *The Theory of Homogeneous Turbulence (Cambridge Science Classics)*. Cambridge University Press, 1953.

G. K. Batchelor. Computation of the energy spectrum in homogeneous two-dimensional turbulence. *Phys. Fluids*, **12**(12):(II)233–(II)239, 1969.

G. K. Batchelor and A. A. Townsend. Decay of isotropic turbulence in the initial period. *Proc. R. Soc. London, Ser. A*, **193**(1035):539–558, 1948.

J. T. Beale and A. Majda. Rates of convergence for viscous splitting of the navier-stokes equations. *Math. Comp.*, **37**(156):243–259, 1981.

J. T. Beale, T. Kato, and A. Majda. Remarks on the breakdown of smooth solutions for the 3-d Euler equations. *Comm. Math. Phys.*, **94**:61–66, 1984.

R. Benzi, G. Paladin, A. Vulpiani, and G. Parisi. On the multifractal nature of fully developed turbulence and chaotic systems. *Journal of Physics A Mathematical General*, **17**:3521–3531, 1984.

G. Beylkin. On the representation of operators in bases of compactly supported wavelets. *SIAM J. Numer. Anal.*, **29**(6):1716–1740, 1992.

L. Biferale, M. Cencini, A. S. Lanotte, and D. Vergni. Inverse velocity statistics in two-dimensional turbulence. *Phys. Fluids*, **15**:1012–1020, 2003.

C. K. Birdsall and A. B. Langdon. *Plasma Physics via Computer Simulation*. McGraw-Hill, New-York, 1985.

L. Boltzmann. Weitere studien über das wärmegleichgewicht unter gasmolekülen. *Sitzungsberichte der Akademie der Wissenschaften*, **66**:275–370, 1872.

W. J. T. Bos and J.-P. Bertoglio. Dynamics of spectrally truncated inviscid turbulence. *Phys. Fluids*, 18(7):071701, 2006.

W. J. T. Bos, L. Liechtenstein, and K. Schneider. Small-scale intermittency in anisotropic turbulence. *Phys. Rev. E*, **76**(4):046310, 2007.

F. Bouchut. Existence and uniqueness of a global smooth solution for the vlasov-poisson-fokker-planck system in three dimensions. *J. Funct. Anal.*, **111**(1):239 – 258, 1993.

F. Bouchut, F. Golse, and M. Pulvirenti. *Kinetic Equations and Asymptotic Theory*. Gauthier-Villars, 2000.

P. BRADSHAW. Turbulence: the chief outstanding difficulty of our subject. *Experiments in Fluids*, **16**:203–216, 1994.

W. BRAUN and K. HEPP. The Vlasov dynamics and its fluctuations in the 1/ N limit of interacting classical particles. *Comm. Math. Phys.*, **56**:101–113, 1977.

J. BRICMONT. Science of chaos or chaos in science? In P. R. GROSS, N. LEVITT, and M. W. LEWIS, editors, *The Flight from Science and Reason*, volume 775, pages 131–175. The New York Academy of Sciences, 1996.

A. J. BRIZARD and T. S. HAHM. Foundations of nonlinear gyrokinetic theory. *Rev. Mod. Phys.*, **79**:421–468, 2007.

S. BRUNNER, E. VALEO, and J. A. KROMMES. Collisional delta-f scheme with evolving background for transport time scale simulations. *Phys. Plasmas*, **6**:4504–4521, 1999.

S. G. BRUSH. *The kind of motion we call heat*. North-Holland, 1976.

J. M. BURGERS. On the resistance experienced by a fluid in turbulent motion. *Proc. KNAW*, **26**:582–604, 1923.

J. M. BURGERS. Experiments on the fluctuations of the velocity in a current of air. *Proc. KNAW*, **29**:547–558, 1926.

J. M. BURGERS. On the application of statistical mechanics to the theory of turbulent fluid motion. *Proc. KNAW*, **32**:414–425, 643–657, 818–833, 1929.

J. M. BURGERS. Mathematical examples illustrating relations occurring in the theory of turbulent fluid motion. *Verh. KNAW, Afd. Natuurkunde, XVII*, **2**:1–53, 1939.

J. M. BURGERS. A mathematical model illustrating the theory of turbulence. volume 1 of *Advances in Applied Mechanics*, pages 171 – 199. Elsevier, 1948.

J. M. BURGERS. On the coalescence of wavelike solutions of a simple non-linear partial differential equation. *Proc. KNAW B, LVII*, **1**:45–72, 1954.

J. M. BURGERS. *The nonlinear diffusion equation: asymptotic solutions and statistical problems*. D. Reidel Pub. Co., 1974.

L. CAFFARELLI, R. KOHN, and L. NIRENBERG. Partial regularity of suitable weak solutions of the Navier-Stokes equations. *Comm. Pure Appl. Math.*, **35**:771–831, 1982.

C. CANUTO, A. QUARTERONI, M. Y. HUSSAINI, and T. A. ZANG. *Spectral methods in fluid dynamics*. Springer-Verlag, 1988.

G. CARBOU and P. FABRIE. Boundary layer for a penalization method for viscous incompressible flow. *Adv. Differential Equations*, **8**:1453, 2003.

A. L. CAUCHY. Recherches sur l'équilibre et le mouvement intérieur des corps solides ou fluides, élastiques ou non élastiques. In *Oeuvres complètes, série II*, volume 2, pages 300–304. Gauthier-Villars et fils (Paris), 1823.

J. R. CHASNOV. On the decay of two-dimensional homogeneous turbulence. *Phys. Fluids*, **9**:171–180, 1997.

D. Chaver, M. Prieto, L. Pinuel, and F. Tirado. Parallel wavelet transform for large scale image processing. In M. Prieto, editor, *Parallel and Distributed Processing Symposium., International, IPDPS 2002*, pages 4–9, 2002.

Y. Chen and S. E. Parker. Coarse-graining phase space in delta f particle-in-cell simulations. *Phys. Plasmas*, **14**(8):082301, 2007.

A. Cheskidov, P. Constantin, S. Friedlander, and R. Shvydkoy. Energy conservation and Onsager's conjecture for the Euler equations. *Nonlinearity*, **21**(6):1233, 2008.

S.-T. Chiu. Bandwidth selection for kernel density estimation. *Ann. Statist.*, **19**(4):1883–1905, 1991.

C. Cichowlas, P. Bonaïti, F. Debbasch, and M. Brachet. Effective dissipation and turbulence in spectrally truncated Euler flows. *Phys. Rev. Lett.*, **95**(26):264502, 2005.

H. Clercx and C.-H. Bruneau. The normal and oblique collision of a dipole with a no-slip boundary. *Comput. Fluids*, **35**(3):245 – 279, 2006.

H. Clercx and G. J. van Heijst. Dissipation of kinetic energy in two-dimensional bounded flows. *Phys. Rev. E*, **65**:066305, 2002.

A. Cohen, I. Daubechies, and P. Vial. Wavelet bases on the interval and fast algorithms. *Appl. Comput. Harm. Anal.*, **1**:54–81, 1993.

R. R. Coifman and D. L. Donoho. Translation-invariant de-noising. In A. Antoniadis and G. Oppenheim, editors, *Wavelets and Statistics*, pages 125–150. Springer-Verlag, New York, 1995.

J. Cole. On a quasi-linear parabolic equation occurring in aerodynamics. *Q. appl. Math.*, **9**:225, 1954.

G. Comte-Bellot and S. Corrsin. The use of a contraction to improve the isotropy of grid-generated turbulence. *J. Fluid Mech.*, **25**(04):657–682, 1966.

P. Constantin. Note on loss of regularity for solutions of the 3–D incompressible Euler and related equations. *Comm. Math. Phys.*, **104**(2):311–326, 1986.

P. Constantin. On the Euler equations of incompressible fluids. *Bulletin of the American Mathematical Society*, **44**:603–621, 2007.

P. Constantin and J. Wu. Inviscid limit for vortex patches. *Nonlinearity*, **8**:735–742, 1995.

P. Constantin and J. Wu. The inviscid limit for non-smooth vorticity. *Indiana University Mathematics Journal*, **45**:67–81, 1996.

P. Constantin, E. Weinan, and E. S. Titi. Onsager's conjecture on the energy conservation for solutions of Euler's equation. *Comm. Math. Phys.*, **165**:207–209, 1994.

E. Cormier-Michel, B. A. Shadwick, C. G. R. Geddes, E. Esarey, C. B. Schroeder, and W. P. Leemans. Unphysical kinetic effects in particle-in-cell modeling of laser wakefield accelerators. *Phys. Rev. E*, **78**(1):016404, 2008.

F. CORON. Derivation of slip boundary conditions for the Navier–Stokes system from the Boltzmann equation. *J. Stat. Phys.*, **54**(3):829–857, 1989.

G. COTTET and P. D. KOUMOUTSAKOS. *Vortex Methods*. Cambridge University Press, 2000.

M. M. COUETTE. Études sur le frottement des liquides. *Annales de Chimie et de Physique, série 6*, **23**:433–510, 1890.

C. COULOMB. Expériences destinées à déterminer la cohérence des fluides et les lois de leur résistance dans les mouvements très lents. *Mémoires de l'Institut National des Sciences et Arts*, **3**:246–306, 1800.

R. COURANT, K. FRIEDRICHS, and H. LEWY. Über die partiellen Differenzengleichungen der mathematischen Physik. *Mathematische Annalen*, **100**:32–74, 1928. 10.1007/BF01448839.

A. DAHAN DALMEDICO, J.-L. CHABERT, and K. CHEMLA, editors. *Chaos et déterminisme*. Seuil, 1992.

M. H. DARCY. Recherches expérimentales relatives au mouvement de l'eau dans les tuyaux. In *Mémoires présentés par divers savants à l'Académie royale des sciences de l'Institut de France*. Institut de France, 1858.

O. DARRIGOL. *Worlds of flow*. Oxford University Press, 2005.

I. DAUBECHIES. Orthonormal bases of compactly supported wavelets. *Comm. Pure Appl. Math.*, **41**(7):909–996, 1988.

I. DAUBECHIES. *Ten Lectures on Wavelets*. SIAM, 1992.

I. DAUBECHIES. Orthonormal bases of compactly supported wavelets II. Variations on a theme. *J. Math. Anal.*, **24**(2):499–519, 1993.

I. DAUBECHIES and J. C. LAGARIAS. Two-scale difference equations. I. Existence and global regularity of solutions. *SIAM Journal on Mathematical Analysis*, **22**(5):1388–1410, 1991.

I. DAUBECHIES and J. C. LAGARIAS. Two-scale difference equations II. Local regularity, infinite products of matrices and fractals. *SIAM Journal on Mathematical Analysis*, **23**(4):1031–1079, 1992.

I. DAUBECHIES, A. GROSSMANN, and Y. MEYER. Painless nonorthogonal expansions. *J. Math. Phys.*, **27**(5), 1986.

P. A. DAVIDSON. Cascades and fluxes in two-dimensional turbulence. *Phys. Fluids*, **20**(2):025106, 2008.

C. DE LELLIS and L. SZÉKELYHIDI. The Euler equations as a differential inclusion. *Ann. Math. (2)*, **170**(3):1417–1436, 2009.

C. DE LELLIS and L. SZÉKELYHIDI. On admissibility criteria for weak solutions of the Euler equations. *Arch. Rat. Mech. Anal.*, **195**(1):225–260, 2010.

J. W. Deardorff. A numerical study of three-dimensional turbulent channel flow at large Reynolds numbers. *J. Fluid Mech.*, **41**:453–480, 1970.

P. Debye and E. Hückel. Zur Theorie der Elektrolyte. *Phys. Zeitsch.*, **24**:185, 1923.

J. C. del Álamo, J. Jiménez, P. Zandonade, and R. D. Moser. Scaling of the energy spectra of turbulent channels. *J. Fluid Mech.*, **500**(-1):135–144, 2004.

D. del Castillo-Negrete. Nonlinear evolution of perturbations in marginally stable plasmas. *Phys. Lett. A*, **241**(1–2):99–104, 1998.

D. del Castillo-Negrete, D. Spong, and S. Hirshman. Proper orthogonal decomposition methods for noise reduction in particle-based transport calculations. *Phys. Plasmas*, **15**:092308, 2008.

J.-M. Delort. Existence de nappes de tourbillon en dimension deux. *J. Am. Math. Soc.*, **4**(3):553–586, 1991.

J. Denavit. Numerical simulation of plasmas with periodic smoothing in phase space. *J. Comp. Phys.*, **9**:75–98, 1972.

L. Desvillettes and C. Villani. On the trend to global equilibrium in spatially inhomogeneous entropy-dissipating systems: The linear Fokker-Planck equation. *Comm. Pure Appl. Math.*, **54**:1–42, 2001.

L. Desvillettes and C. Villani. On the trend to global equilibrium for spatially inhomogeneous kinetic systems: The Boltzmann equation. *Inventiones Mathematicae*, **159**:245–316, 2004.

P. Dmitruk and D. C. Montgomery. Numerical study of the decay of enstrophy in a two-dimensional Navier–Stokes fluid in the limit of very small viscosities. *Phys. Fluids*, **17**:035114, 2005.

D. Donoho and I. Jonhstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, **81**:425–455, 1994.

D. L. Donoho. Nonlinear Solution of Linear Inverse Problem by Wavelet-Vaguelette Decomposition, 1992.

D. L. Donoho, I. M. Johnstone, G. Keryacharian, and D. Picard. Density estimation by wavelet thresholding. *Ann. Statist.*, **24**(2):508–539, 1996.

D. G. Dritschel. Contour dynamics and contour surgery: Numerical algorithms for extended, high-resolution modelling of vortex dynamics in two-dimensional, inviscid, incompressible flows. *Computer Physics Reports*, **10**(3):77 – 146, 1989.

D. G. Dritschel, C. V. Tran, and R. K. Scott. Revisiting Batchelor's theory of two-dimensional turbulence. *J. Fluid Mech.*, **591**:379–391, 2007.

H. L. Dryden and A. M. Kuethe. The measurement of fluctuations of air speed by the hot-wire anemometer. Technical Report 320, NACA, 1929.

T. Dubois, F. Jauberteau, and R. Temam. Solution of the incompressible Navier-Stokes equations by the nonlinear Galerkin method. *J. Sci. Comp.*, **8**:167–194, 1993.

W. EHM, T. GNEITING, and D. RICHARDS. Convolution roots of radial positive definite functions with compact support. *Trans. Am. Math. Soc.*, **356**:4655–4685, 2004.

A. EISNER and B. TURKINGTON. Nonequilibrium statistical behavior of nonlinear Schrödinger equations. *Physica D*, **213**:85–97, 2006.

R. S. ELLIS, K. HAVEN, and B. TURKINGTON. Large deviation principles and complete equivalence and nonequivalence results for pure and mixed ensembles. *J. Stat. Phys.*, **101**: 999–1064, 2000.

G. L. EYINK. Dissipative anomalies in singular Euler flows. *Physica D*, **237**:1956–1968, 2008.

G. L. EYINK and K. R. SREENIVASAN. Onsager and the theory of hydrodynamic turbulence. *Rev. Mod. Phys.*, 78(1):87, 2006.

M. FARGE. L'approche numérique : simulation ou simulacre des phenomènes ? In J. PETITOT, editor, *Logos et Theorie des Catastrophes*. Patino, 1988b.

M. FARGE. Wavelet transforms and their applications to turbulence. *Annu. Rev. Fluid Mech.*, **24**:395, 1992.

M. FARGE and G. RABREAU. Transformée en ondelettes pour détecter et analyser les structures cohérentes dans les écoulements turbulents bidimensionnels. *C. R. Acad. Sci. Paris Série II b*, **307**:433–440, 1988a.

M. FARGE and K. SCHNEIDER. Coherent vortex simulation (CVS), a semi-deterministic turbulence model using wavelets. *Flow, Turbulence and Combustion*, **66**(4):393–426, 2001.

M. FARGE, E. GOIRAND, Y. MEYER, F. PASCAL, and M. V. WICKERHAUSER. Improved predictability of two-dimensional turbulent flows using wavelet packet compression. *Fluid Dynam. Res.*, **10**(4-6):229 – 250, 1992a.

M. FARGE, M. HOLSCHNEIDER, and T. PHILIPOVITCH. Formation et stabilite des structures coherentes quasi-singulières en turbulence bidimensionnele. *C. R. Acad. Sci. Paris, series II*, **315**:1585–1592, 1992b.

M. FARGE, K. SCHNEIDER, and N. KEVLAHAN. Non-Gaussianity and coherent vortex simulation for two-dimensional turbulence using an adaptative orthogonal wavelet basis. *Phys. Fluids*, **11**(8):2187–2201, 1999.

M. FARGE, G. PELLEGRINO, and K. SCHNEIDER. Coherent vortex extraction in 3D turbulent flows using orthogonal wavelets. *Phys. Rev. Lett.*, **87**(5):054501, 2001.

M. FARGE, K. SCHNEIDER, and P. DEVYNCK. Extraction of coherent bursts from turbulent edge plasma in magnetic fusion devices using orthogonal wavelets. *Phys. Plasmas*, **13**:042304, 2006.

C. L. FEFFERMAN. Existence and smoothness of the Navier-Stokes equation. In J. CARLSON, A. JAFFE, and A. WILES, editors, *The millenium prize problems*. Clay Mathematics Institute, 2000.

J. H. Ferziger and M. Perić. *Computational methods for fluid dynamics*. Springer, 2002.

J. Flor and G. J. van Heijst. An experimental study of dipolar vortex structures in a stratified fluid. *J. Fluid Mech.*, **279**:101–133, 1994.

C. Foiaš and G. Prodi. Sur le comportement global des solutions non-stationnaires des équations de Navier-Stokes en dimension 2. *Rend. Sem. Mat. Univ. Padova*, **39**:1–34, 1967.

C. Foiaš, O. P. Manley, R. Temam, and Y. M. Treve. Asymptotic analysis of the Navier-Stokes equations. *Physica D*, **9**:157–188, 1983.

C. Foiaš, O. Manley, R. Rosa, and R. Temam. *Navier-Stokes equations and turbulence*. CUP, 2001.

M. Frigo and S. G. Johnson. The design and implementation of FFTW3. *Proceedings of the IEEE*, **93**(2):216–231, 2005. Special issue on "Program Generation, Optimization, and Platform Adaptation".

U. Frisch. *Turbulence: the legacy of A.N. Kolmogorov*. Cambridge University Press, 1995.

U. Frisch and J. Bec. Burgulence. In *New trends in turbulence*, Les Houches, pages 341–383. Springer, 2001.

U. Frisch and G. Parisi. On the singularity structure of fully developed turbulence. In *International School of Physics "Enrico Fermi" Course 88*, pages 84–88. North Holland, 1985.

U. Frisch, S. Kurien, R. Pandit, W. Pauls, S. Sankar Ray, A. Wirth, and J.-Z. Zhu. Hyperviscosity, Galerkin truncation and bottlenecks in turbulence. *Phys. Rev. Lett.*, **101**:144501, 2008.

G. Gallavotti. Ergodicity, ensembles, irreversibility in Boltzmann and beyond. *J. Stat. Phys.*, **78**(5):1571–1589, 1995.

S. Gassama, E. Sonnendrücker, K. Schneider, M. Farge, and M. O. Domingues. Wavelet denoising for postprocessing of a 2d particle-in-cell code. *ESAIM : Proceedings*, **16**:196–210, 2007.

L. Genovese, A. Neelov, S. Goedecker, T. Deutsch, S. A. Ghasemi, A. Willand, D. Caliste, O. Zilberberg, M. Rayson, A. Bergman, and R. Schneider. Daubechies wavelets as a basis set for density functional pseudopotential calculations. *J. Comp. Phys.*, **129**(1):014109–+, 2008.

D. Gerard-Varet and E. Dormy. On the ill-posedness of the Prandtl equation. *ArXiv e-prints*, , 2009.

P. Germain and R. Bader. Unicité des écoulements avec chocs dans la mécanique de Burgers. Technical Report 1/1711-1, ONERA, 1953.

P. Ghendrih, Y. Sarazin, G. Attuel, S. Benkadda, P. Beyer, G. Darmet, G. Falchetto, C. Figarella, X. Garbet, V. Grandgirard, and M. Ottaviani. Statistical analysis of turbulent front propagation in the scrape-off-layer. *J. Nucl. Mater.*, **337-339**:347–351, 2005. PSI-16.

J. D. Gibbon. The three-dimensional Euler equations: Where do we stand? *Physica D*, **237**:1894–1904, 2008.

E. Goirand, V. Wickerhauser, and M. Farge. A parallel two-dimensional wavelet packet transform and some applications in computing and compression analysis. In B. Joseph and R. Motard, editors, *Applications of Wavelets to Chemical Engineering*, pages 275–319. Kluwer, 1994.

D. E. Goldstein and O. V. Vasilyev. Stochastic coherent adaptive large eddy simulation method. *Phys. Fluids*, **16**:2497–2513, 2004.

S. Goldstein, editor. *Modern developments in fluid dynamics*. Dover, New York, 1965.

K. K. Golovkin. Vanishing viscosity in the Cauchy problem for the equations of hydromechanics. *Trudy Mat. Inst. Steklov*, **92**:31–49, 1966.

G. H. Golub and C. F. van Loan. *Matrix computations*. Johns Hopkins University Press, 1996.

P. Gonzales, J. C. Cabaleiro, and T. F. Pena. Parallel computation of wavelet transforms using the lifting scheme. *J. Supercomputing*, **18**:141–152, 2001.

D. Gottlieb and J. S. Hesthaven. Spectral methods for hyperbolic problems. *J. Comput. Appl. Math.*, **128**(1–2):83 – 131, 2001.

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. A drift-kinetic Semi-Lagrangian 4D code for ion turbulence simulation. *J. Comp. Phys.*, **217**:395–423, 2006.

A. Grossmann and J. Morlet. Decomposition of Hardy functions into square integrable wavelets of constant shape. *J. Math. Anal.*, **15**(4):723–736, 1984.

S. N. Gurbatov, S. I. Simdyankin, E. Aurell, U. Frisch, and G. Tóth. On the decay of Burgers turbulence. *J. Fluid Mech.*, **344**:339–374, 1997.

O. D. Gürcan, X. Garbet, P. Hennequin, P. H. Diamond, A. Casati, and G. L. Falchetto. Wave-number spectrum of drift-wave turbulence. *Phys. Rev. Lett.*, **102**(25):255002, 2009.

O. D. Gürcan, P. Hennequin, L. Vermare, X. Garbet, and P. H. Diamond. Shell models and the possibility of application to fusion plasmas. *Plasma Phys. Control. Fusion*, **52**(4):045002, 2010.

J. Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, **13**:49–52, 1902.

J. M. Hammersley, D. C. Handscomb, and G. Weiss. *Monte Carlo Methods*. Taylor & Francis, 1975.

A. Harten. High resolution schemes for hyperbolic conservation laws. *J. Comp. Phys.*, **49**(3):357 – 393, 1983.

M. Hauray and P. Jabin. N-particles Approximation of the Vlasov Equations with Singular Potential. *Arch. Rational Mech. Anal.*, **183**:489–524, 2007.

R. D. Hazeltine. Transport theory in the collisionless limit. *Phys. Plasmas*, **5**:3282–3286, 1998.

J. Heikkinen, S. Janhunen, T. Kiviniemi, and F. Ogando. Full f gyrokinetic method for particle simulation of tokamak transport. *J. Comp. Phys.*, **227**(11):5582 – 5609, 2008.

H. Helmholtz. Über Integrale der hydrodynamischen Gleichungen, welche den Wirbelbewegungen entsprechen. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, **1858**(55):25–55, 1858.

C. Hennig. Clusters, outliers, and regression: Fixed point clusters. *Journal of Multivariate Analysis*, **86**:183–212, 2003.

R. W. Hockney. Computer experiment of anomalous diffusion. *Phys. Fluids*, **9**(9): 1826–1835, 1966.

R. W. Hockney and J. W. Eastwood. *Computer Simulation using Particles*. IOP, Bristol, Philadelphia, 1988.

D. D. Holm, J. E. Marsden, and T. S. Ratiu. Euler-Poincaré models of ideal fluids with nonlinear dispersion. *Phys. Rev. Lett.*, **80**(19):4173–4176, 1998.

M. Holmström. Parallelizing the fast wavelet transform. *Parallel Computing*, **21**:1837–1848, 1995.

E. Hopf. The partial differential equation $u_t + uu_x = mu_{xx}$. *Comm. Pure Appl. Math.*, **3** (3):201–230, 1950.

E. Hopf. Über die Anfangswertaufgabe für die hydrodynamischen Grundgleichungen. *Math. Nachr.*, **4**:213–231, 1951.

S. Hoyas and J. Jiménez. Scaling of the velocity fluctuations in turbulent channels up to $\mathrm{Re}_\tau = 2003$. *Phys. Fluids*, 18(1):011702, 2006.

Y. Idomura, M. Ida, S. Tokuda, and L. Villard. New conservative gyrokinetic full$-f$ Vlasov code and its comparison to gyrokinetic δ-f particle-in-cell code. *J. Comp. Phys.*, **226**(1):244 – 262, 2007.

A. Ihler. Kde toolbox for matlab, http://www.ics.uci.edu/ ˜ihler/code/kde.html, 2003.

T. Ishihara, T. Gotoh, and Y. Kaneda. Study of high Reynolds number isotropic turbulence by direct numerical simulation. *Annu. Rev. Fluid Mech.*, **41**:165–180, 2009.

F. Jacobitz, K. Schneider, W. J. T. Bos, and M. Farge. On the structure and dynamics of sheared and rotating turbulence: Anisotropy properties and geometrical scale-dependent statistics. *Phys. Fluids*, **22**:085101, 2010.

S. JAFFARD. Pointwise smoothness, two-microlocalization and wavelet coefficients. *Publicaciones Matemàtiques*, **35**:155–168, 1991.

S. JAFFARD. Wavelet methods for fast resolution of elliptic problems. *SIAM J. Numer. Anal.*, **29**(4):965–986, 1992.

E. T. JAYNES. Information theory and statistical mechanics. *Phys. Rev.*, **106:620–630 and 108**:171–190, 1957.

E. T. JAYNES. Gibbs vs Boltzmann entropies. *Am. J. Phys.*, **33**(5):391–398, 1965.

E. T. JAYNES. Where do we go from here? In C. RAY SMITH and W. T. GRANDY JR., editors, *Maximum-Entropy and Bayesian Methods in Inverse Problems*. D. Reidel Publishing Company, 1985.

J. JIMÉNEZ and R. D. MOSER. What are we learning from simulating wall turbulence? *Philos. Trans. Roy. Soc. London, Series A*, **365**:715–732, 2007.

B. R. JOHNSON, J. P. MODISETTE, P. J. NORDLANDER, and J. L. KINSEY. Quadrature integration for orthogonal wavelet systems. *J. Chem. Phys.*, **110**:8309–8317, 1999.

I. JOHNSTONE and B. SILVERMAN. Wavelet treshold estimators for data with correlated noise. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, **59**:319–351, 1997.

S. JOLLIET, A. BOTTINO, P. ANGELINO, R. HATZKY, T. TRAN, B. MCMILLAN, O. SAUTER, K. APPERT, Y. IDOMURA, and L. VILLARD. A global collisionless PIC code in magnetic coordinates. *Comput. Phys. Commun.*, **177**(5):409 – 425, 2007.

D. A. JONES and E. TITI. Upper bounds on the number of determining modes, nodes, and volume elements for the Navier–Stokes equations. *Indiana University Mathematics Journal*, **42**:1–12, 1993.

P. KAILASNATH, K. R. SREENIVASAN, and G. STOLOVITZKY. Probability density of velocity increments in turbulent flows. *Phys. Rev. Lett.*, **68**:2766–2769, 1992.

Y. KANEDA, T. ISHIHARA, M. YOKOKAWA, K. ITAKURA, and A. UNO. Energy dissipation rate and energy spectrum in high resolution direct numerical simulations of turbulence in a periodic box. *Phys. Fluids*, **15**(2):L21–L25, 2003.

L. M. KAPLAN and C.-C. J. KUO. An improved method for 2-D self-similar image synthesis. *IEEE T. Image. Process.*, **5**(5):754, 1996.

T. KATO. On classical solutions of the two-dimensional non-stationary Euler equation. *Arch. Rat. Mech. Anal.*, **25**(3):188–200, 1967.

T. KATO. Nonstationary flows of viscous and ideal fluids in $\mathbb{R}^3$. *J. Funct. Anal.*, **9**:296–305, 1972.

T. KATO. Remarks on zero viscosity limit for nonstationary Navier-Stokes flows with boundary. In *Seminar on nonlinear partial differential equations*, pages 85–98. MSRI, Berkeley, 1984.

G. Keetels, U. D'Ortona, W. Kramer, H. Clercx, K. Schneider, and G. J. van Heijst. Fourier spectral and wavelet solvers for the incompressible Navier–Stokes equations with volume-penalization: Convergence of a dipole–wall collision. *J. Comp. Phys.*, **227**:919–945, 2007.

N. Kevlahan and M. Farge. Vorticity filaments in two-dimensional turbulence: creation, stability and effect. *J. Fluid Mech.*, **346**:49–76, 1997.

N. Kevlahan, J. Alam, and O. V. Vasilyev. Scaling of space time modes with Reynolds number in two-dimensional turbulence. *J. Fluid Mech.*, **570**:217–226, 2007.

M. Kholmyansky and A. Tsinober. On an alternative explanation of anomalous scaling and how well-defined is the concept of inertial range. *Phys. Lett. A*, **373**(27-28):2364–2367, 2009.

B. Khouider and E. S. Titi. An inviscid regularization for the surface quasi-geostrophic equation. *Comm. Pure Appl. Math.*, **61**(10):1331–1346, 2008.

G. Khujadze and M. Oberlack. DNS and scaling laws from new symmetry groups of ZPG turbulent boundary layer flow. *Theor. Comput. Fluid Dynam.*, **18**:391–411, 2004.

G. Khujadze and M. Oberlack. New scaling laws in ZPG turbulent boundary layer flow. In *Proceedings of fifth international symposium on turbulence and shear flow phenomena*, pages 443–448, 2007.

G. Khujadze, R. Nguyen van yen, K. Schneider, M. Oberlack, and M. Farge. Coherent vorticity extraction in 3d turbulent boundary layers using orthogonal wavelets. Preprint submitted to Proceedings of the CTR, 2010.

S. Kida. Asymptotic properties of Burgers turbulence. *J. Fluid Mech.*, **93**(2):337–377, 1979.

J. Kim, P. Moin, and R. Moser. Turbulence statistics in fully developed channel flow at low Reynolds number. *J. Fluid Mech.*, **177**(-1):133–166, 1987.

N. Kingsbury. The dual-tree complex wavelet transform : A new technique for shift invariance and directional filters. In *Proc. 8th IEEE DSP Workshop*, 1998.

N. Kingsbury. Complex wavelets and shift invariance. *IEEE Seminar on time-scale and time-frequency analysis and applications*, **19**, 2000.

N. Kingsbury. Complex wavelets for shift invariant analysis and filtering of signals. *Appl. Comput. Harm. Anal.*, **10**(3):234–253, 2001.

N. Kingsbury. Shift invariant properties of the Dual-Tree Complex Wavelet Transform. 1999.

A. N. Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large Reynolds' numbers. *Dokl. Akad. Nauk SSSR*, **30**:301–305, 1941.

R. H. Kraichnan. Irreversible statistical mechanics of incompressible hydromagnetic turbulence. *Phys. Rev.*, **109**:1407–1422, 1958.

R. H. KRAICHNAN. Inertial ranges in two-dimensional turbulence. *Phys. Fluids*, **10**(7): 1417–1423, 1967.

R. H. KRAICHNAN. On Kolmogorov's inertial-range theories. *J. Fluid Mech.*, **62**:305–330, 1974.

R. H. KRAICHNAN. Eddy viscosity in two and three dimensions. *J. Atmos. Sci.*, **33**: 1521–1536, 1976.

R. H. KRAICHNAN. Reduced descriptions of hydrodynamic turbulence. *J. Stat. Phys.*, **51** (5):949–963, 1988.

R. H. KRAICHNAN and S. CHEN. Is there a statistical mechanics of turbulence? *Physica D*, **37**(1-3):160 – 172, 1989.

W. KRAMER, H. CLERCX, and G. J. VAN HEIJST. Vorticity dynamics of a dipole colliding with a no-slip wall. *Phys. Fluids*, **19**:126603, 2007.

J. A. KROMMES. Dielectric response and thermal fluctuations in gyrokinetic plasma. *Phys. Fluids B*, **5**:1066–1100, 1993.

J. A. KROMMES. Nonquilibrium gyrokinetic fluctuation theory and sampling noise in gyrokinetic particle-in-cell simulations. *Phys. Plasmas*, **14**:090501, 2007.

J. A. KROMMES and G. HU. The role of dissipation in the theory and simulations of homogeneous plasma turbulence, and resolution of the entropy paradox. *Phys. Plasmas*, **1**:3211–3238, 1994.

S. N. KRUZHKOV. First order quasi-linear systems in several independent variables. *Math. USSR Sb.*, **10**:217–243, 1970. (*Amer. Math. Transl.*, Series 2, **26**, 95-172).

O. A. LADYZHENSKAYA. Solution "in the large" of the nonstationary boundary value problem for the Navier-Stokes system with two space variables. *Comm. Pure Appl. Math.*, **12**(3):427–433, 1959.

O. A. LADYZHENSKAYA. *The Mathematical Theory of Viscous Incompressible Flow*. Gordon and Breach, 1963.

L. LANDAU. On the vibration of the electronic plasma. *J. Phys. USSR*, **10**:25, 1946. (Collected works, D. ter Haar ed., Pergamon Press 1965, pp. 445–460).

A. B. LANGDON and C. K. BIRDSALL. Theory of plasma simulation using finite-size particles. *Phys. Fluids*, **13**(8):2115, 1970.

I. LANGMUIR. Oscillation in ionized gases. *Proc. Nat. Acad. Sci. US*, **14**:627, 1928.

P. D. LAX. Weak solutions of nonlinear hyperbolic equations and their numerical computation. *Comm. Pure Appl. Math.*, **7**(1):159–193, 1954.

P. D. LAX. *Hyperbolic systems of conservation laws and the mathematical theory of shock waves*. SIAM, 1973.

J. LE ROND D'ALEMBERT. Paradoxe proposé aux géomètres sur la résistance des fluides. In *Opuscules Mathématiques*, volume 5, chapter XXXIV, pages 132–138. Briasson, 1768.

C. E. Leith. Diffusion Approximation for Two-Dimensional Turbulence. *Phys. Fluids*, **11**: 671–672, 1968.

A. Lenard and I. B. Bernstein. Plasma oscillations with diffusion in velocity space. *Phys. Rev.*, **112**(5):1456–1459, 1958.

J. Leray. Essai sur le mouvement d'un liquide visqueux emplissant l'espace. *Acta Mathematica*, **63**:193–248, 1934.

M. Lesieur, O. Métais, and P. Comte. *Large-eddy simulations of turbulence*. Cambridge University Press, 2005.

J. L. V. Lewandowski. Low-noise collision operators for particle-in-cell simulations. *Phys. Plasmas*, **12**(5):052322–+, 2005.

D. K. Lilly. The representation of small scale turbulence in numerical simulation experiments. In *IBM Scientific Computing Symposium on environmental sciences*, pages 195–210, Yorktown heights, 1967.

J.-L. Lions and G. Prodi. Un théoreme d'existence et d'unicité dans les équations de Navier-Stokes en dimension 2. *C. R. Acad. Sci. Paris*, **248**:3519–3521, 1959.

P.-L. Lions. *Mathematical topics in fluid mechanics*, volume I. Clarendon Press, Oxford, 1996.

P.-L. Lions and B. Perthame. Propagation of moments and regularity for the 3-dimensional Vlasov-Poisson system. *Inventiones Mathematicae*, **105**:415–+, 1991.

M. C. Lopes Filho, A. Mazzucato, and H. J. Nussenzveig Lopes. Weak solutions, renormalized solutions and enstrophy defects in 2d turbulence. *Arch. Rational Mech. Anal.*, **179**:353–387, 2006.

M. C. Lopes Filho, A. Mazzucato, H. J. Nussenzveig Lopes, and M. Taylor. Vanishing viscosity limits and boundary layers for circularly symmetric 2d flows. *Bull. Brazilian Math. Soc.*, **39**(4):471–513, 2008.

M. C. Lopes Filho, A. L. Mazzucato, and H. J. Nussenzveig Lopes. Vanishing viscosity limit for incompressible flow inside a rotating circle. *Physica D*, **237**:1324–1333, 2008.

E. N. Lorenz. Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**(2):130–141, 1963.

E. N. Lorenz. The predictability of a flow which possesses many scales of motion. *Tellus*, **21**(3):289–307, 1969a.

E. N. Lorenz. Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.*, **26**:636–+, 1969b.

E. N. Lorenz. Atmospheric predicability experiments with a large numerical model. *Tellus*, **34**:505–+, 1982.

E. N. Lorenz. Predictability: A problem partly solved. In *Proc. ECMWF Seminar on Predictability*, volume 1, pages 1–18. ECMWF, Reading, UK, 1996.

A. LUNDBLADH, S. BERLIN, M. SKOTE, C. HILDINGS, J. CHOI, J. KIM, and D. S. HENNINGSON. An efficient spectral method for simulation of incompressible flow over a flat plate. Technical report, KTH, 1999.

S. MALLAT. *A wavelet tour of signal processing*. Academic Press, 1999.

J. H. MALMBERG and C. B. WHARTON. Collisionless damping of electrostatic plasma waves. *Phys. Rev. Lett.*, **13**(6):184–186, 1964.

I. MARUSIC, B. J. MCKEON, P. A. MONKEWITZ, H. M. NAGIB, A. J. SMITS, and K. R. SREENIVASAN. Wall-bounded turbulent flows at high Reynolds numbers: Recent advances and key issues. *Phys. Fluids*, **22**(6):065103, 2010.

N. MASMOUDI. Examples of singular limits in hydrodynamics. In C. DAFERMOS and E. FEIREISL, editors, *Handbook of Differential Equations: Evolutionary Equations*, volume 3, chapter 3, pages 195 – 275. North-Holland, 2007.

N. MASMOUDI. Remarks about the inviscid limit of the Navier-Stokes system. *Comm. Math. Phys.*, **270**:777–788, 2007.

M. MATSUMOTO and T. NISHIMURA. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Trans. Model. Comput. Simul.*, **8**(1):3–30, 1998.

J. C. MAXWELL. On the dynamical theory of gases. *Philos. Trans. Royal Soc. London, Series I*, **157**:49–88, 1867.

B. J. MCKEON and A. S. SHARMA. A critical-layer framework for turbulent pipe flow. *J. Fluid Mech.*, **658**(-1):336–382, 2010.

B. F. MCMILLAN, S. JOLLIET, T. M. TRAN, L. VILLARD, A. BOTTINO, and P. ANGELINO. Long global gyrokinetic simulations: Source terms and particle noise control. *Phys. Plasmas*, 15(5):052308, 2008.

C. MENEVEAU. Dual spectra and mixed energy cascade of turbulence in the wavelet representation. *Phys. Rev. Lett.*, **66**(11):1450–1453, 1991.

C. MENEVEAU and J. KATZ. Scale-invariance and turbulence models for large-eddy simulation. *Annu. Rev. Fluid Mech.*, **32**:1–32, 2000.

C. MENEVEAU and K. R. SREENIVASAN. Simple multifractal cascade model for fully developed turbulence. *Phys. Rev. Lett.*, **59**:1424–1427, 1987.

R. MOSER and P. MOIN. Direct numerical simulation of curved turbulent channel flow. Technical Report 85974, NASA, 1984.

C. MOUHOT and C. VILLANI. Landau damping. *J. Math. Phys.*, **51**(1):015204–+, 2010.

C. NAVIER. Mémoire sur les lois du mouvement des fluides. *Mém. Acad. Sci. Inst. France*, **6**:389–440, 1823.

A. I. NEELOV and S. GOEDECKER. An efficient numerical quadrature for the calculation of the potential energy of wavefunctions expressed in the Daubechies wavelet basis. *J. Comp. Phys.*, **217**:312–339, 2006.

W. M. Nevins, G. W. Hammett, A. M. Dimits, W. Dorland, and D. Shumaker. Discrete particle noise in particle-in-cell simulations of plasma microturbulence. *Phys. Plasmas*, **12**:122305, 2005.

C. S. Ng, A. Bhattacharjee, and F. Skiff. Weakly collisional Landau damping and three-dimensional Bernstein-Greene-Kruskal modes: New results on old problems. *Phys. Plasmas*, **13**(5):055903–+, 2006.

R. Nguyen van yen, M. Farge, D. Kolomenskiy, K. Schneider, and N. Kingsbury. Wavelets meet Burgulence: CVS-filtered Burgers equation. *Physica D*, **237**(14–17):2151–2155, 2008.

R. Nguyen van yen, M. Farge, and K. Schneider. Wavelet regularization of a Fourier-Galerkin method for solving the 2D incompressible Euler equations. *ESAIM: Proc.*, **29**:89–107, 2009.

R. Nguyen van yen, D. del Castillo-Negrete, K. Schneider, M. Farge, and G. Chen. Wavelet based density estimation for noise reduction in plasma simulation using particles. *J. Comp. Phys.*, **229**(8):2821–2839, 2010a.

R. Nguyen van yen, M. Farge, and K. Schneider. Energy dissipating structures in the vanishing viscosity limit of 2D incompressible flows with solid boundaries. Preprint submitted to Phys. Rev. Lett., 2010b.

R. Nguyen van yen, M. Farge, and K. Schneider. Scale-wise coherent vorticity extraction for conditional statistical modelling of homogeneous isotropic 2d turbulence. Preprint submitted to Physica D, 2010c.

R. Nguyen van yen, E. Sonnendrücker, K. Schneider, and M. Farge. Particle-in-wavelets scheme for the 1D Vlasov-Poisson equations. Preprint submitted to ESAIM:Proc, 2010d.

O. M. Nielsen and M. Hegland. Parallel performance of fast wavelet transforms. *Int. J. High Speed Comp.*, **11**:55–73, 2000.

A. M. Obukhov. On the distribution of energy in the spectrum of a turbulent flow. *Dokl. Akad. Sci. Nauk SSSR*, **32A**:22–24, 1941.

N. Okamoto, K. Yoshimatsu, K. Schneider, M. Farge, and Y. Kaneda. Coherent vortices in high resolution direct numerical simulation of homogeneous isotropic turbulence: A wavelet viewpoint. *Phys. Fluids*, **19**(11):115109, 2007.

H. Okuda, A. T. Lin, C. C. Lin, and J. M. Dawson. Splines and high order interpolations in plasma simulations. *Comput. Phys. Commun.*, **17**(3):227 – 231, 1979.

O. A. Oleinik. Discontinuous solutions of nonlinear differential equations. *Usp. Mat. Nauk*, **12**(3):3–73, 1957. (*Amer. Math. Transl.*, Series 2, **26**, 95-172).

L. Onsager. Statistical hydrodynamics. *Il Nuovo Cimento*, **6**(0):279–287, 1949.

P. Orlandi. Vortex dipole rebound from a wall. *Phys. Fluids A*, **2**:1429–1436, 1990.

P. Orlandi. *Fluid Flow Phenomena: A Numerical Toolkit*. Springer, 2000.

S. A. ORSZAG. Analytical theories of turbulence. *J. Fluid Mech.*, **41**:363–386, 1970.

S. OSHER and F. SOLOMON. Upwind difference schemes for hyperbolic systems of conservation laws. *Math. Comp.*, **38**(158):339–374, 1982.

S. E. PARKER and W. W. LEE. A fully nonlinear characteristic method for gyrokinetic simulation. *Phys. Fluids B*, **5**(1):77, 1993.

E. PARZEN. On estimation of a probability density function and mode. *Ann. Math. Stat.*, **33**(3):1065–1076, 1962.

J. PEDLOSKY. *Geophysical fluid dynamics*. Springer, 1987.

K. PFAFFELMOSER. Global classical solutions of the Vlasov-Poisson system in three dimensions for general initial data. *J. Diff. Eq.*, **95**(2):281 – 303, 1992.

A. POUQUET, M. LESIEUR, J. C. ANDRÉ, and C. BASDEVANT. Evolution of high Reynolds number two-dimensional turbulence. *J. Fluid Mech.*, **72**(02):305–319, 1975.

L. PRANDTL. Über Flüssigkeitsbewegung bei sehr kleiner Reibung. In *Proc. 3rd Inter. Math. Congr. Heidelberg*, pages 484–491, 1904.

L. PRANDTL. Über die ausgebildete Turbulenz. *ZS ang. Math. Mech.*, **5**:136–139, 1925.

L. PRANDTL. Zur turbulenten strömung in rohren und längs platten. *Ergeb. Aerodyn. Versuch.*, **4**:18–29, 1932.

T. H. PULLIAM. Artificial dissipation models for the Euler equations. *AIAA Journal*, **24**: 1931–1940, 1986.

L. RAYLEIGH. On the stability, or instability, of certain fluid motions. *Proc. London Math. Soc.*, **s1-11**(1):57–72, 1879.

O. REYNOLDS. An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous, and of the law of resistance in parallel channels. *Philos. Trans. Roy. Soc. London, Series A*, **174**:935–982, 1883.

O. REYNOLDS. On the dynamical theory of incompressible viscous fluids and the determination of the criterion. *Proc. Roy. Soc. London Ser. A*, **451**(1941):5–47, 1995.

L. F. RICHARDSON. *Weather Prediction by Numerical Process*. Cambridge University Press, 1922.

L. F. RICHARDSON and J. A. GAUNT. Diffusion regarded as a compensation for smoothing. *Memoirs of the Royal Meteorological Society*, **3**:171–175, 1930.

R. ROBERT and C. ROSIER. Long range predictability of atmospheric flows. *Nonlinear Processes Geophys.*, **8**:55–67, 2001.

R. ROBERT and J. SOMMERIA. Relaxation towards a statistical equilibrium in two-dimensional perfect fluid dynamics. *Phys. Rev. Lett.*, **69**:2776–2779, 1992.

D. RUELLE, editor. *Turbulence, Strange attractors, and Chaos*. World Scientific, 1995.

D. Ruelle and F. Takens. On the nature of turbulence. *Comm. Math. Phys.*, **20**: 167–192, 1971.

A.-J.-C. B. Saint-Venant. Note à joindre au mémoire sur la dynamique des fluides présenté le 14 avril 1834. *C. R. Acad. Sci. Paris*, **17**:1240–1243, 1843.

M. Sammartino and R. E. Caflisch. Zero viscosity limit for analytic solutions of the Navier-Stokes equation on a half-space I. Existence for Euler and Prandtl equations. *Comm. Math. Phys.*, **192**:433–461, 1998a.

M. Sammartino and R. E. Caflisch. Zero viscosity limit for analytic solutions of the Navier-Stokes equation on a half-space. II. Construction of the Navier-Stokes solution. *Comm. Math. Phys.*, **192**:463–491, 1998b.

V. A. Sandborn. Measurements of intermittency of turbulent motion in a boundary layer. *J. Fluid Mech.*, **6**:221–240, 1959.

A. Sarthou, S. Vincent, J. P. Caltagirone, and P. Angot. Eulerian Lagrangian grid coupling and penalty methods for the simulation of multiphase flows interacting with complex objects. *Internat. J. Numer. Methods Fluids*, **56**:1093–1099, 2008.

V. Scheffer. Partial regularity of solutions to the Navier-Stokes equations. *Pacific J. Math.*, **66**(2):540–551, 1976.

A. A. Schekochihin, S. C. Cowley, W. Dorland, G. W. Hammett, G. G. Howes, G. G. Plunk, E. Quataert, and T. Tatsuno. Gyrokinetic turbulence: a nonlinear route to dissipation through phase space. *Plasma Phys. Control. Fusion*, **50** (12):124024–+, 2008.

R. Schiestel. *Modélisation et simulation des écoulements turbulents*. Hermès, Paris, 1993.

P. Schlatter, R. Örlü, Q. Li, G. Brethouwer, J. H. M. Fransson, A. V. Johansson, P. H. Alfredsson, and D. S. Henningson. Turbulent boundary layers up to $\mathrm{Re}_\theta = 2500$ studied through simulation and experiment. *Phys. Fluids*, 21(5): 051702, 2009.

H. Schlichting. *Boundary layer theory*. McGraw-Hill, New York, 1979.

K. Schneider and M. Farge. Numerical simulation of the transient flow behaviour in tube bundles using a volume penalization method. *J. Fluids Struct.*, **20**:555–566, 2005.

K. Schneider and O. Vasilyev. Wavelet methods in computational fluid dynamics. *Annu. Rev. Fluid Mech.*, **42**, 2010.

K. Schneider, M. Farge, and N. Kevlahan. Spatial intermitency in two-dimensional turbulence: a wavelet approach. In N. Tongring and R. Penner, editors, *Perspectives in Mathematics and Physics*, volume 34, pages 302–328. World Scientific, 2004.

K. Schneider, M. Farge, A. Azzalini, and J. Ziuber. Coherent vortex extraction and simulation of 2D isotropic turbulence. *J. Turbulence*, **7**:N44, 2006.

I. W. Selesnick. The design of approximate Hilbert transform pairs of wavelet bases. *IEEE T. Signal. Proces.*, **50**(5):1144–1152, 2002.

C. E. SHANNON. A mathematical theory of communication. *Bell System Technical Journal*, **27**:379–423 and 623–656, 1948.

Z.-S. SHE, E. AURELL, and U. FRISCH. The inviscid Burgers equation with initial data of Brownian type. *Comm. Math. Phys.*, **148**:623–641, 1992.

A. SHNIRELMAN. On the nonuniqueness of weak solution of the Euler equation. *Comm. Pure Appl. Math.*, **50**(12):1261 – 1286, 1998.

R. SHVYDKOY. On the energy of inviscid singular flows. *J. Math. Anal. Appl.*, **349**(2): 583–595, 2009.

R. SHVYDKOY. Lectures on the Onsager conjecture. *Discrete Contin. Dyn. Sys., Series S*, **3**:473–496, 2010.

B. W. SILVERMAN. Wavelets in statistics: beyond the standard assumptions. *Philos. Trans. Roy. Soc. London, Series A*, , 1999.

B. W. SILVERMAN. *Density estimation for statistics and data analysis*. Chapman and Hall, 1986.

M. P. SIMENS, J. JIMENEZ, S. HOYAS, and Y. MIZUNO. A high-resolution code for turbulent boundary layers. *J. Comp. Phys.*, **228**:4218–4231, 2009.

M. SKOTE. *Studies of turbulent boundary layer flow through direct numerical simulation*. PhD thesis, Royal Institute of Technology (KTH), Stockholm, Sweden, 2001.

J. SMAGORINSKY. General circulation experiments with the primitive equations. *Mon. Wea. Rev.*, **91**(3):99–164, 1963.

S. SMALE. Differentiable dynamical systems. *Bull. Am. Math. Soc.*, **73**:747–808, 1967.

E. SONNENDRÜCKER, J. ROCHE, P. BERTRAND, and A. GHIZZO. The semi-Lagrangian method for the numerical resolution of the Vlasov equation. *J. Comp. Phys.*, **149**:201–220, 1999.

P. R. SPALART. Direct simulation of a turbulent boundary layer up to $\mathrm{Re}_\theta = 1410$. *J. Fluid Mech*, **187**:61–98, 1988.

C. G. SPEZIALE. Analytical methods for the development of Reynolds-stress closures in turbulence. *Annu. Rev. Fluid Mech.*, **23**:107–157, 1991.

K. R. SREENIVASAN. On the scaling of the turbulence energy dissipation rate. *Phys. Fluids*, **27**:1048–1051, 1984.

G. STOKES. On the theories of internal friction of fluids in motion, and of the equilibrium and motion of elastic solids. *Trans. Cambridge Phil. Soc.*, **8**:287–305, 1845.

H. S. G. SWANN. The convergence with vanishing viscosity of nonstationary Navier-Stokes flow to ideal flow in $\mathbb{R}^3$. *Trans. Am. Math. Soc.*, **157**:373–397, 1971.

P. K. SWEBY. High resolution schemes using flux limiters for hyperbolic conservation laws. *SIAM J. Numer. Anal.*, **21**(5):995–1011, 1984.

W. Sweldens and R. Piessens. Quadrature formulae and asymptotic error expansions for wavelet approximations of smooth functions. *SIAM J. Numer. Anal.*, **31**(4):1240–1264, 1994.

P. Tabeling. Two-dimensional turbulence: a physicist approach. *Phys. Rep.*, **362**:1–62, 2002.

E. Tadmor. Convergence of spectral methods for nonlinear conservation laws. *SIAM J. Numer. Anal.*, **26**(1):30–44, 1989.

A. H. Taub, editor. *The collected works of John von Neumann*. Oxford: Pergamon, 1963.

G. I. Taylor. Statistical theory of turbulence. *Proc. R. Soc. London, Ser. A*, **151**(873): 421–444, 1935.

G. I. Taylor. Fluid friction between rotating cylinders. I. Torque measurements. *Proc. R. Soc. London, Ser. A*, **157**(892):546–564, 1936.

C. Tenllado, J. Setoain, M. Prieto, L. Pinuel, and F. Tirado. Parallel implementation of the 2d discrete wavelet transform on graphics processing units: Filter bank versus lifting. *IEEE T. Parall. Distr.*, **19**:299–310, 2008.

H. Tennekes and J. L. Lumley. *A first course in turbulence*. MIT Press, Cambridge, MA, 1972.

B. Terzić, I. V. Pogorelov, and C. L. Bohn. Particle-in-cell beam dynamics simulations with a wavelet-based Poisson solver. *Phys. Rev. ST Accel. Beams*, **10**:034201, 2007.

A. H. Tewfik and M. Kim. Correlation structure of the discrete wavelet coefficients of fractional Brownian motion. *IEEE T. Inform. Theory.*, **38**(2):904–910, 1992.

T. Theodorsen. Mechanism of turbulence. In *Proceedings of the second Midwestern conference on fluid mechanics*, Columbus, OH, 1952. Ohio state university.

W. Thomson. On vortex motion. *Trans. R. S. Edin.*, **25**, 1869.

C. V. Tran and D. G. Dritschel. Vanishing enstrophy dissipation in two-dimensional Navier-Stokes turbulence in the inviscid limit. *J. Fluid Mech.*, **559**:107–116, 2006.

L. Trefethen. *Spectral Methods in Matlab*. SIAM, 2000.

C. Truesdell. *Essays in the history of mechanics*. Springer Verlag, 1968.

C. Truesdell. *The tragicomical history of thermodynamics, 1822-1854*. Springer, 1980.

B. Turkington, A. Majda, K. Haven, and M. Dibattista. Statistical equilibrium predictions of jets and spots on Jupiter. *Proceedings of the National Academy of Science*, **98**:12346–12350, 2001.

G. K. Vallis. Remarks on the predictability properties of two- and three-dimensional flow. *Quart. J. Roy. Meteor. Soc.*, **111**:1039–1047, 1985.

M. Vannucci and B. Vidakovic. Preventing the Dirac disaster: Wavelet based density estimation. *J. Italian Stat. Soc.*, **6**(DP 95-24):15–19, 1998.

M. VERGASSOLA, B. DUBRULLE, U. FRISCH, and A. NOULLEZ. Burgers equation, devils staircases and the mass distribution for large-scale structures. *Astronomy and Astrophysics*, **289**:325–356, 1994.

H. D. VICTORY and E. J. ALLEN. The convergence theory of particle-in-cell methods for multidimensional vlasov-poisson systems. *SIAM J. Numer. Anal.*, **28**(5):1207–1241, 1991.

B. VIDAKOVIC. *Statistical Modeling by Wavelets*. Wiley, 1999.

C. VILLANI. A review of mathematical topics in collisional kinetic theory. In S. FRIEDLANDER and D. SERRE, editors, *Handbook of Mathematical Fluid Dynamics*, volume 1, chapter 2, pages 71–305. Elsevier, 2002.

T. VON KÁRMÁN. über laminare und turbulent Reibung. *ZS ang. Math. Mech.*, **1**(4): 233–252, 1921. English translation in NACA TM 1092 (1946).

T. VON KÁRMÁN. Mechanische ähnlichkeit und turbulenz. *Nach. Ges. Wiss. Goettingen Math.-Phys. Klasse*, :58–76, 1930. (eng. trans. in NACA tech. mem. 611, 1931).

J. VON NEUMANN. *Recent theories of turbulence*, pages 437–471. Volume 5 of Taub (1963).

R. VON SACHS and K. SCHNEIDER. Wavelet smoothing of evolutionary spectra by non–linear thresholding. *Appl. Comput. Harmon. Anal.*, **3**:268–282, 1996.

W. X. WANG, W. M. LIN, Z. TANG, W. W. LEE, S. ETHIER, J. L. V. LEWANDOWSKI, G. REWOLDT, T. S. HAHM, and M. J. Gyro-kinetic simulation of global turbulent transport properties in tokamak experiments. *Phys. Plasmas*, **13**: 092505, 2006.

Y. WANG. Function estimation via wavelet shrinkage for long memory data. *Ann. Statist.*, **24**(1):466–484, 1996.

T. WELLER, K. SCHNEIDER, M. OBERLACK, and M. FARGE. DNS and wavelet analysis of a turbulent channel flow rotating about the streamwise direction. In K. HANJALIC, Y. NAGANO, and S. JARKIRLIC, editors, *Turbulence, Heat and Mass Transfer*, volume 1 of *5*, pages 163–166. 2006.

W. WOLIBNER. Un theorème sur l'existence du mouvement plan d'un fluide parfait, homogène, incompressible, pendant un temps infiniment long. *Mathematische Zeitschrift*, **37**(1):698–726, 1933.

X. WU and P. MOIN. DNS statistics data of zero-pressure-gradient flat-plate boundary layer. *J. Fluid Mech.*, **630**:5–41, 2009.

Z. P. XIN. Zero viscosity limit for incompressible Navier-Stokes equations with the Navier boundary conditions. Communication, IMA, Minneapolis, 2010.

L. H. YANG and M. MISRA. Coarse-grained parallel algorithms for multi-dimensional wavelet transforms. *J. Supercomputing*, **12**:99–118, 1998.

V. I. YUDOVICH. Non-stationary flows of an ideal incompressible fluid. *USSR Comput. Math. and Math. Phys.*, **3**:1407–1456, 1963.

V. I. Yudovich. On the loss of smoothness of the solutions of the Euler equations and the inherent instability of flows of an ideal fluid. *Chaos*, **10**(3):705–719, 2000.

I. Zaliapin and M. Ghil. Another look at climate sensitivity. *Nonlinear Processes Geophys.*, **17**:113–122, 2010.