Remarks on the Approximation of the Euler Equations by a Low Order Model

Angelo Iollo

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Remarks on the Approximation of the Euler Equations by a Low Order Model

Angelo Iollo

N° 3329

Decembre 1997

THÈME 4

ISSN 0249-6399

Remarks on the Approximation of the Euler Equations by a Low Order Model

Angelo Iollo

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet SINUS

Rapport de recherche n° 3329 — Decembre 1997 — 25 pages

Abstract: Fluid flows are very often governed by the dynamics of a small number of coherent structures, i.e., fluid features which keep their individuality during the evolution of the flow. The purpose of this paper is to present a way to simulate the Euler equations on the basis of the evolution of such coherent structures. One way to extract from flow simulations some basis functions which can be interpreted as coherent structures is by Proper Orthogonal Decomposition (POD). Then, by means of a Galerkin projection, it is possible to find the system of ODEs which approximates the problem in the finite dimensional space spanned by the basis functions found. Issues concerning the stability and the accuracy of such an approximation are discussed. It is found that a straight-forward Galerkin method is unstable. Some features of discontinuous Galerkin methods are therefore incorporated to achieve stability, which is proved for a linear scalar case. In addition, we propose a way to reduce the cost of the computation and to increase accuracy at the same time. Some one-dimensional computational experiments are presented, including shock tube simulations and rarefaction fans.

Key-words: POD, Euler equations, Galerkin method, low order model.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, B.P. 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : 04 93 65 77 77 - International : +33 4 93 65 77 77 — Fax : 04 93 65 77 65 - International : +33 4 93 65 77 65
à partir du 05/01/1998
Téléphone : 04 92 38 77 77 - International : +33 4 92 38 77 77 — Fax : 04 92 38 77 65 - International : +33 4 92 38 77 65

Notes sur l'Approximation des Équations d'Euler par un modèl d'ordre faible

Résumé : Les écoulements fluides sont souvent gouvernés par la dynamique d'un petit nombre de structures cohérentes, ayant la propriété de conserver leur individualité au cours de l'évolution de l'écoulement. L'objectif de cette étude est de présenter une méthode de simulation des équations d'Euler sur la base de l'évolution de ces structures cohérentes. Une des manières d'extraire, à partir de simulations d'écoulements, des fonctions de base qui peuvent être interprétées comme des structures cohérentes, est l'utilisation de la Décomposition Orthogonale aux Valeurs Propres. Dès lors, grâce à la projection de Galerkin, il est possible de déterminer le système d'EDO qui approche le problème dans l'espace à dimension finie des fonctions de base trouvées. On discutera des problèmes liés à la stabilité et à la précision d'une telle approximation. Par ailleurs, nous avons trouvé qu'une méthode directe de Galerkin n'est pas stable. Nous avons donc utilisé quelques propriétés des méthodes discontinues de Galerkin afin d'obtenir la stabilité, que nous avons prouvée dans un cas linéaire scalaire. De plus, nous introduisons un nouvel algorithme qui permet à la fois de réduire le coût de calcul et d'augmenter la précision. Nous présentons quelques simulations 1-D sur des tubes à choc et des ondes de raréfaction.

Mots-clés : POD, équations d'Euler, méthode de Galerkin, modèl d'ordre faible.

1 Introduction

Simulation schemes based on the Euler equations are used as a good compromise between model sophistication and computational attainability when dealing with complex geometries and compressible flows. Yet, the computational cost of Euler simulations is still too high in contexts such as optimization or control, where large number of computations for different configurations are needed. This is particularly true if one aims at controlling unsteady flows.

In this paper the focus is on devising an efficient way to reproduce the main features of unsteady Euler flows, but instead of using less sophisticated physical models or very coarse approximations we resort to a method, the proper orthogonal decomposition (POD), by which it is possible to extract from an existing data-base of computations a certain number of basis functions. The solution is then constructed as a time-dependent linear combination of such functions. This representation should provide a sufficient description of the flow as long as global features, such as stability or growth of small perturbations, are involved.

The POD was introduced in fluid mechanics by Lumley [1] in the study of turbulent flows (for a review see [2]). In that setting the basis functions obtained by POD are recognized to be coherent structures, i.e. spatial features which repeatedly appear in space and time. In the works cited the basic assumption is that the coherent structures capture most of the dynamics of the system, i.e. they give a good representation of the realizations of the flow in the phase space. In the background of such assumption is the discovery that turbulence might be the manifestation in the physical space of a strange attractor of limited dimension in phase space.

From this view-point the situation for unsteady Euler equations is completely different since we have no a priori result for the phase space and the questions attaining the possibility of simulating the unsteady Euler equations by means of a reduced number of basis function which retain the basic dynamic properties (low order model) will be the object of a subsequent investigation. Nonetheless, POD analysis should prove effective also for the Euler equations. Think, for example, of an oscillating airfoil in an inviscid flow: there is a continuous shedding of vortices from the trailing edge to compensate for the Kutta condition and a big unsteady vortex centered on the airfoil responsible for the lift. In this situation it seems reasonable to assume that the entire flow-field is governed by the interactions of these structures, which are limited in number.

The method of "snapshots" proposed by Sirovich [3] for incompressible Navier-Stokes equations is adapted here to the Euler equations. This method, that will be outlined below for completeness, allows for the construction of the basis functions starting from a limited number of "snapshots" of the flow at subsequent times.

The scope of this paper is to contribute to the construction of robust and efficient schemes for the integration of the system of equations resulting from the projection of the Euler equations on the set of basis functions obtained by POD. In fact, in most applications presented so far (which deal essentially with the incompressible Navier-Stokes equations) there was no discussion concerning the stability of such schemes, probably because physical viscosity tends to damp out spurious oscillations. An example of such situation is found in the application of the canonical spectral method. Consider the linear scalar advection

equation. It is well known that if trigonometric expansions are used, the resulting scheme of integration is unstable. Yet, if Burgers' equation is discretized by the same method, for high enough viscosity, the scheme is stable.

This is not the case for Euler equations: the computational solution may blow up in a few timesteps as a result of numerical instabilities. For Navier-Stokes equations as well, if the Reynolds number is too high, similar stability problems can be foreseen unless special care is taken in modeling the unresolved scales. In this respect, Aubry *et alii* [4] presented a model for the flat plate turbulent boundary layer based on a Galerkin projection of Navier-Stokes equations onto POD eigenfunctions resulting from experiments. The unresolved modes were accounted by a model for the small scales which ensures the proper amount of dissipation (and stabilization), so that the system of ordinary differential equations obtained is consistent and stable at the same time. On the other hand, for low Reynolds number applications, such that presented in [5], apparently there is no need for modeling the unresolved modes.

The stability issue has important consequences on the efficiency of the scheme. Indeed, the main advantage of a low order model over, say, a standard finite-volume scheme should be its supposed computational efficiency. Yet, it is shown in what follows that one way to make the scheme stable, is to compute upwind fluxes at the interfaces of discretization cells, so that the overall cost of computation is similar to the cost of other conventional schemes. However, there are other methods to stabilize POD Galerkin schemes. Indeed, simple arguments suggest that the Galerkin approximation of a scalar advection equation is stable if the underlying finite-differences or finite-volumes scheme is stable.

Finally, inspired by the work of Cockburn and Shu [6], we describe a method by which it is possible to cluster some discretization cells achieving higher order accuracy and computational efficiency at the same time.

2 POD and the Euler equations

In this section we present a novel view in which the POD is seen in the setting of linear integral equations. For the construction of the basis functions for the Euler equations we extend the derivation of [3] which applied to the incompressible Navier-Stokes equations. In addition, some results of the Hilbert-Schmidt theory of symmetric kernels relevant to our study are introduced.

For simplicity we consider the quasi one-dimensional Euler equations

$$W_t + F_x + S = 0 \tag{1}$$

with $W = \{\rho, \rho u, \rho e\}$, $F = \{\rho u, \rho u^2 + p, u(\rho e + p)\}$, $S = \beta\{\rho u, \rho u^2, u(\rho e + p)\}$ and where ρ is the density, u is the speed of the flow, e is the total energy per unit mass and $\beta = Q_x/Q$ with Q the nozzle cross section area. The pressure is given by $p = 1/2\,(\gamma - 1)\,(2\,\rho\,e - \rho\,u^2)$ where γ is the specific heats ratio.

Suppose that from existing data we have the solution at M different time steps t_n in terms of $W^{(n)} = W(x, t_n)$. It is required to find a function $\phi(x) = (\phi_1, \phi_2, \phi_3) \in (L^2(R))^3$

which gives the best representation of $W^{(n)}$ in the following sense

$$\max_{\psi} \frac{< (W^{(n)}, \psi)^2 >}{(\psi, \psi)} = \frac{< (W^{(n)}, \phi)^2 >}{(\phi, \phi)} \tag{2}$$

where (ϕ, ϕ) denotes the canonical inner product

$$\int_{\Omega} (\phi_1 \phi_1 + \phi_2 \phi_2 + \phi_3 \phi_3)\, dx \tag{3}$$

and the brackets $< \cdot >$ indicate the time average

$$< (W^{(n)}, \phi)^2 > = \frac{1}{M} \sum_{n=1}^{M} (W^{(n)}, \phi)^2 \tag{4}$$

In other words, one seeks a function ϕ which is most parallel in an average sense to the given solution set $W^{(n)}$.

This problem finds its natural setting in the theory of linear integral equations (see chapter III of [7]). Here some results relevant to our study are sketched in the frame of such a theory. For proofs omitted we address the reader to [7].

Consider the quadratic integral form

$$J(\psi, \psi) = \int_{\Omega} \psi(s)\, K(s,t)\, \psi(t)\, dt\, ds \tag{5}$$

where $K(s,t) = K(t,s)$ is a symmetric 3 by 3 matrix and $\psi \in (L^2(R))^3$. Suppose that the functional J takes on positive values only, i.e. $\forall \psi$: $J(\psi, \psi) > 0$; in this case $J(\psi, \psi)$ is said to be positive definite. Suppose also that $(\psi, \psi) = 1$.

If the function $\phi^{(1)}$ satisfies

$$\int_{\Omega} K(s,t)\, \phi^{(1)}(t) dt = \lambda_1\, \phi^{(1)}(s) \tag{6}$$

then one can show that $J(\phi^{(1)}, \phi^{(1)}) = \lambda_1$ is the solution of the maximization problem for the form $J(\psi, \psi)$ subject to the constraint $(\psi, \psi) = 1$.

Now, take

$$K(s,t) = \frac{1}{M} \sum_{n=1}^{M} W^{(n)}(s) \otimes W^{(n)}(t) \tag{7}$$

where $\otimes$ is the dyadic product and insert it in eq. 5. We have

$$J(\psi, \psi) = \frac{1}{M} \sum_{n=1}^{M} \left[\int_{\Omega} \psi(s)\, W^{(n)}(s)\, ds \int_{\Omega} \psi(t)\, W^{(n)}(t)\, dt \right] =< (W^{(n)}, \psi)^2 > \tag{8}$$

therefore the maximum of $J(\psi,\psi)$ with $(\psi,\psi)=1$ solves the problem stated at the beginning of this section. In this case eq. 6 becomes

$$\frac{1}{M}\sum_{n=1}^{M}\int_{\Omega} W^{(n)}(s)\otimes W^{(n)}(t)\;\phi^{(1)}(t)\,dt=\lambda_1\;\phi^{(1)}(s) \tag{9}$$

We recognize in eq. 9 an eigenvalue problem for the symmetric, degenerate, positive definite kernel of eq. 7. In this respect, every continuous symmetric kernel that does not vanish identically possesses eigenvalues and eigenfunctions. Also since the kernel of eq. 9 is real, symmetric and degenerate then the eigenvalues are M in number and all real. The eigenvalues are positive since $J(\psi,\psi)$ in eq. 8 is positive definite.

The totality of eigenvalues and eigenfunctions of eq. 9 are found sequentially. Each new eigenfunction is sought as the solution the maximization problem of the functional J subject to the constraint of being orthogonal to all previously found eigenfunctions. Any r-fold degenerate eigenvalue is associated with r linearly independent eigenfunctions.

In conclusion, once the eigenvalue problem of eq. 9 is solved, we are left with a set of M eigenvalues λ_n and M orthonormal eigenfunctions $\phi^{(n)}$ which give an optimal representation, in the sense made precise at the beginning of this section, of the solution known at different time steps. In the following sections we will project the Euler equations over the space spanned by these orthonormal functions.

2.1 Approximation properties

It is easy to show that the eigenfunctions of eq. 6 are able to represent the functions $W^{(n)}$ on average, i.e.

$$<\int_{\Omega}\left(W^{(n)}-\sum_{i=1}^{M}(W^{(n)},\phi^{(i)})\,\phi^{(i)}\right)^{2}dx>=0 \tag{10}$$

However, the main idea on which the application of the POD analysis relies, is that a given number of basis functions found by means of the POD capture more features of the flow compared to any other set of orthonormal functions.

To see this, consider the energy-like quantity

$$\Upsilon\;=\;\frac{1}{M}\sum_{n=1}^{M}\int_{\Omega} W^{(n)}(s)\,W^{(n)}(s)\,ds \tag{11}$$

On the basis of the following result known as Mercer's theorem, one has

$$\frac{1}{M}\sum_{n=1}^{M} W^{(n)}(s)\,W^{(n)}(t)=\sum_{i=1}^{M}\lambda_i\,\phi^{(i)}(s)\,\phi^{(i)}(t) \tag{12}$$

and hence $\Upsilon = \sum_{i=1}^{M} \lambda_i$. If another orthonormal basis is taken, we have $W^{(n)} = \sum_{i=1}^{M} b_i^n \psi_i$ and

$$\Upsilon = \frac{1}{M} \sum_{i=1}^{M} \sum_{n=1}^{M} (b_i^n)^2 = < \sum_{i=1}^{M} (b_i^n)^2 > \tag{13}$$

On the other hand, it was shown that $< (W^{(n)}, \phi^{(i)})^2 > = \lambda_i$. Clearly for any other ψ with $(\psi, \psi) = 1$ one has $< (W^{(n)}, \psi)^2 > \leq \lambda_i$. Therefore $< (b_i^n)^2 > \leq \lambda_i$ and hence $\sum_{i=1}^{M} < (b_i^n)^2 > \leq \sum_{i=1}^{M} \lambda_i$.

Given M, the basis functions resulting from the POD have the advantage that they give a better description of Υ compared to any other set of orthonormal functions. This is the main reason for which we try to apply POD to the construction of a low order model for Euler equations.

2.2 Solution of the eigenvalue problem

Sirovich [3] proposed a way to solve the eigenvalue problem of eq. 9 based on the fact, evident in the same equation, that the eigenfunctions can be expressed in terms of the original set of data, i.e.

$$\phi(s) = \sum_{n=1}^{M} a_n W^{(n)}(s) \tag{14}$$

where a_n is a scalar. If eq. 14 is introduced in eq. 9 we are left with the problem of linear algebra of finding the eigenvalues and the eigenvectors of

$$C\, A = \lambda\, A \tag{15}$$

where $A = (a_1, \ldots, a_M)$ is one of the M eigenvectors, λ is the correspondent eigenvalue, and $C = 1/M\, (W^{(i)}, W^{(j)})$ for $i, j = 1, \ldots, M$. When the solution W is sampled on a large number of snapshots it is very likely that a certain number of eigenvalues will be very close to zero so that the contribution of the corresponding eigenfunction to the description of the flow may be considered negligible. In this case we can consider only the first M_c eigenvalues for the projection of the Euler equations over the corresponding eigenfunctions, in order to have a limited set or ODEs to solve.

Finally it should be noted that all the analysis presented so far can be repeated all the same in two or three dimensions. Of course we will have $\phi(x) \in (L^2(R^2))^4$ and $\phi(x) \in (L^2(R^3))^5$ respectively. The definitions of other operators will change accordingly.

2.3 Computational example

One of the model problems studied throughout this work is the Laval nozzle with cross section area of equation $Q = c_1\,(x + x_0) + c_2\,/(x + x_0)$ in the interval $x \in [0, 1]$. In all of the

cases studied it was taken $x_0 = 0.1$, $c_1 = 2.5$ and $c_2 = 0.3$. The fluid in the nozzle is initially at rest and at given pressure and temperature. At time $t = 0^+$ the pressure at the boundary $x = 1$ is decreased to a fraction of the initial value. An expansion fan propagates backward inside the nozzle until it reaches the boundary $x = 0$ where total pressure and entropy are imposed. At this boundary the fan is reflected and propagates forward until it is reflected backward by the right boundary. This process ends when a steady flow is established in the nozzle.

The computer code used to simulate such a flow is based on the time-dependent, finite-volume, flux-difference splitting formulation of [8]. Higher order spatial accuracy is achieved by means of ENO reconstruction [9] and time accuracy is obtained by the Taylor series method.

The solution found with the scheme described above is saved every ΔTs time steps and the number of discretization cells is N. Once C is computed, the eigenvalues λ_i and the eigenvectors $(a_1, \ldots, a_M)$ are found numerically. Although C is positive definite, when ΔTs is large, the eigenvalues closer to 0 may be sometime negative due to round off errors. However, such eigenvalues will be disregarded since they would give a negligible contribution to the description of the solution. Only the first eigenfunctions, corresponding to the largest eigenvalues will be considered. Finally, using eq. 14, we are able to determine the eigenfunctions ϕ which are defined on the same discretization cells of the snapshots $W^{(n)}$.

For the case shown in fig. 1 and fig. 2 the computation is performed for $N = 400$ and with first order accuracy. The number of time steps is 1000 and $\Delta Ts = 20$. The CFL is 0.4.

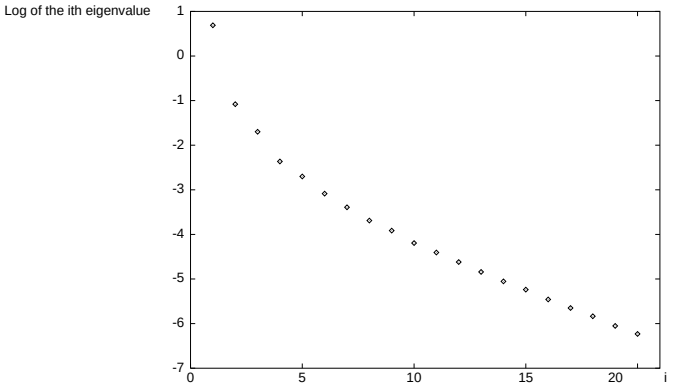

Figure 1: POD analysis of the data relative to an expansion fan. Logarithm of the first 20 eigenvalues.

Only the 20 highest eigenvalues are shown. Most of the information about the flow is in the first few eigenfunctions and relative eigenvalues, which tend fastly to 0 for increasing i. It is seen that the eigenfunctions computed have more and more extrema as the corresponding eigenvalue decreases. The eigenfunctions found are such that the space spanned is reasonably

close to that spanned by an ordinary Fourier series. The results found in this case are very similar to those found for a fan propagating into a straight duct.

3 Galerkin projection

Once an optimal basis of functions is found, it is left to actually construct an approximation to eqs. 1. We assume that the solution $W(x,t)$ belongs to the finite dimensional space spanned by the $\phi^{(j)}$, i.e.

$$W(x,t) = \sum_{j=1}^{M} w_j(t)\,\phi^{(j)}(x) \tag{16}$$

It should be noted that w_j is a scalar coefficient solely function of time. The evolution of the solution from a given initial condition is determined entirely by the ODEs satisfied by each w_j. In order to find such ODEs we project eqs. 1 onto every $\phi^{(j)}$

$$\int_\Omega (W_t + F_x + S)\;\phi^{(j)}\,dx \;= 0 \tag{17}$$

Substituting eq. 16 in the previous equation and considering that $(\phi^{(i)}, \phi^{(j)}) = \delta_{i\,j}$, we have

$$\dot{w}_j + \int_\Omega F_x\,\phi^{(j)}\,dx + \int_\Omega S\,\phi^{(j)}\,dx \;=\; 0 \tag{18}$$

Let us discretize this equation. Denoting by I_i the generic computational cell and integrating by parts we have

$$\dot{w}_j + \sum_{i=1}^{N}\left[(\phi^{(j)}_{i+\frac{1}{2}})^-\,F_{i+\frac{1}{2}} - (\phi^{(j)}_{i-\frac{1}{2}})^+\,F_{i-\frac{1}{2}}\right] - \sum_{i=1}^{N}\int_{I_i} F\,\phi^{(j)}_x\,dx + \sum_{i=1}^{N}\int_{I_i} S\,\phi^{(j)}\,dx = 0 \tag{19}$$

The locations $i+\frac{1}{2}$ and $i-\frac{1}{2}$ indicate the edges of I_i. The eigenfunctions $\phi^{(j)}$ are in general discontinuous at such boundaries, therefore it is necessary to discriminate between $(\phi^{(j)}_{i+\frac{1}{2}})^-$ and $(\phi^{(j)}_{i+\frac{1}{2}})^+$. Equation 19 is an ODE and governs the evolution of the coefficient w_j.

The boundary conditions are imposed explicitly taking the appropriate fluxes at the boundary edges of the discretization cells. If not explicitly imposed, the boundary conditions would be respected in an average sense because of the way the basis functions are derived. However, if one aims at using such simulations in conjunction with boundary control, it is necessary to impose the boundary conditions explicitly.

Taking

$$\Theta_j = \sum_{i=1}^{N}\left((\phi^{(j)}_{i+\frac{1}{2}})^-\,F_{i+\frac{1}{2}} - (\phi^{(j)}_{i-\frac{1}{2}})^+\,F_{i-\frac{1}{2}}\right) - \sum_{i=1}^{N}\int_{I_i} F\,\phi^{(j)}_x\,dx + \sum_{i=1}^{N}\int_{I_i} S\,\phi^{(j)}\,dx \tag{20}$$

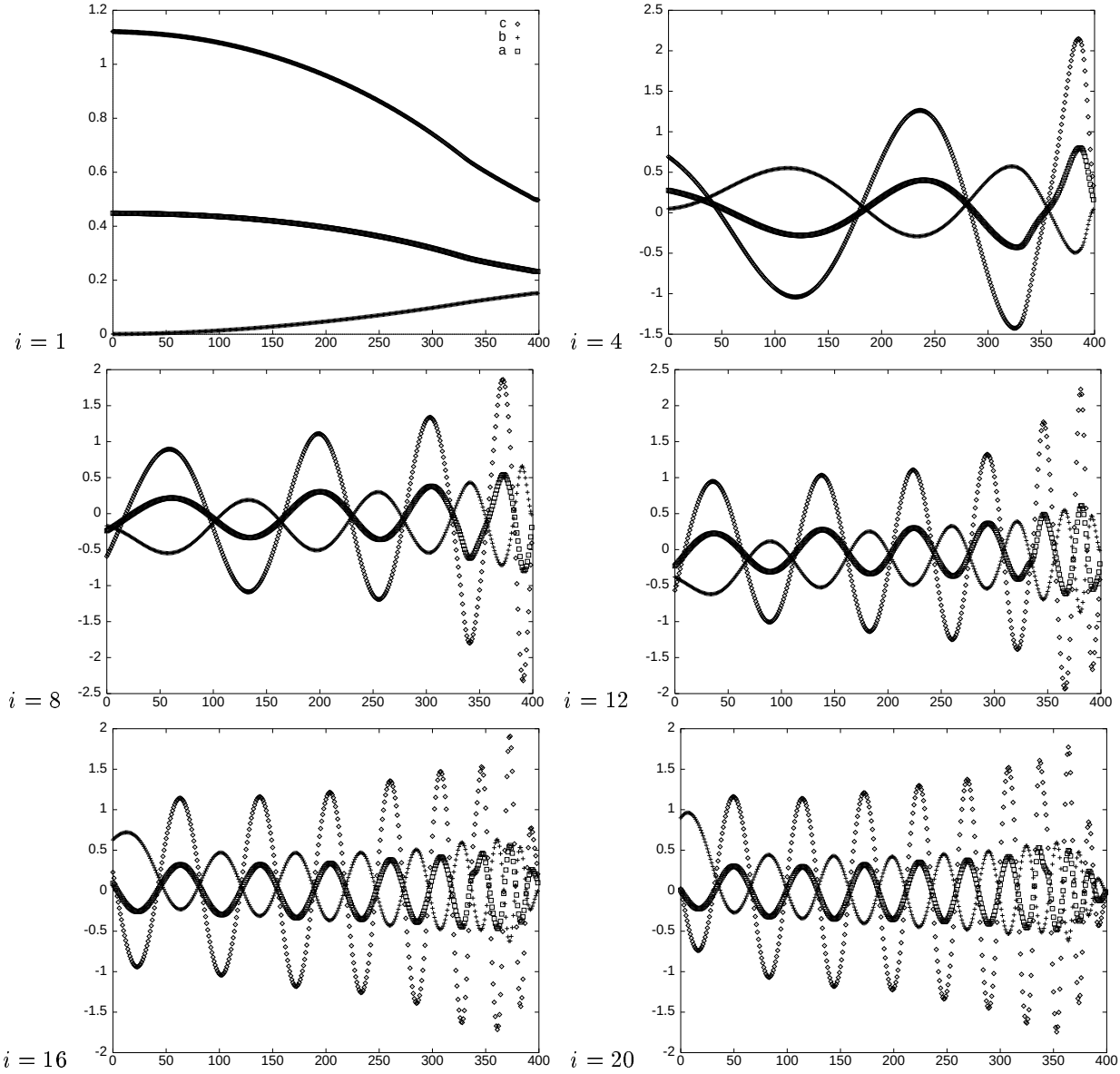

Figure 2: POD analysis of the data relative to an expansion fan. Eigenfunctions components for the ith eigenvalue: a) ϕ_1, b) ϕ_2, c) ϕ_3.

it is possible to write eq. 19 as

$$\dot{w}_j + \Theta_j = 0 \tag{21}$$

Of course we would like to have a stable integration scheme for this equation. Stability depends on how we choose the fluxes at the cell interfaces $F_{i+\frac{1}{2}}$ and $F_{i-\frac{1}{2}}$. Indeed, this stability issue is similar to that arising in discontinuous Galerkin methods.

3.1 Stability

For the sake of a simple analysis, let us consider a linear scalar equation of the same form of eq. 21 in which the flux and the conservative variable are linked by

$$F = c\,W \tag{22}$$

where c is a scalar constant. For simplicity let $S = 0$. Multiply eq. 21 by w_j and sum over j to get

$$\sum_{j=1}^{M} \left[\frac{d}{dt} \left(\frac{w_j^2}{2} \right) + w_j\,\Theta_j \right] = 0 \tag{23}$$

In order to have a stable scheme, the norm of the vector whose elements are the coefficients w_j should always be bounded from above. Therefore we request the following inequality to be true

$$\sum_{j=1}^{M} \frac{d}{dt} \left(\frac{w_j^2}{2} \right) \leq 0 \tag{24}$$

Such a condition implies that

$$\sum_{j=1}^{M} w_j\,\Theta_j \geq 0 \tag{25}$$

Let us see when this is the case. We have

$$\sum_{j=1}^{M} w_j\,\Theta_j = \sum_{i=1}^{N} \left((W_{i+\frac{1}{2}})^-\,F_{i+\frac{1}{2}} - (W_{i-\frac{1}{2}})^+\,F_{i-\frac{1}{2}} \right) - \sum_{i=1}^{N} \int_{I_i} F\,W_x\,dx \tag{26}$$

since $(W_{i+\frac{1}{2}})$ may as well be discontinuous across the cell edge.

Using the linear relation of eq. 22 we can integrate the last integral of the previous equation so that

$$\sum_{j=1}^{M} w_j\,\Theta_j = \sum_{i=1}^{N} \left\{ (W_{i+\frac{1}{2}})^-\,F_{i+\frac{1}{2}} - (W_{i-\frac{1}{2}})^+\,F_{i-\frac{1}{2}} - \frac{1}{2}\,c\,\left[(W^2_{i+\frac{1}{2}})^- - (W^2_{i-\frac{1}{2}})^+ \right] \right\} \tag{27}$$

Suppose that periodic boundary conditions are imposed. Rearranging the summation around the edge at $i+\frac{1}{2}$ one obtains

$$\sum_{j=1}^{M} w_j \, \Theta_j = \sum_{i=1}^{N-1} [W]_{i+\frac{1}{2}} \left(c \, \overline{W}_{i+\frac{1}{2}} - F_{i+\frac{1}{2}} \right) \tag{28}$$

where $\overline{W}_{i+\frac{1}{2}} = \frac{1}{2} \left(\left(W_{i+\frac{1}{2}} \right)^+ + \left(W_{i+\frac{1}{2}} \right)^- \right)$ and $[W]_{i+\frac{1}{2}} = \left(W_{i+\frac{1}{2}} \right)^+ - \left(W_{i+\frac{1}{2}} \right)^-$. Now, if the flux at cell interfaces is taken to be

$$F_{i+\frac{1}{2}} = c \, \overline{W}_{i+\frac{1}{2}} - \frac{|c|}{2} [W]_{i+\frac{1}{2}} = \frac{c-|c|}{2} \left(W_{i+\frac{1}{2}} \right)^+ + \frac{c+|c|}{2} \left(W_{i+\frac{1}{2}} \right)^- \tag{29}$$

that is the upwind flux according to the sign of c, from eq. 28 we have finally

$$\sum_{j=1}^{M} w_j \, \Theta_j = \sum_{i=1}^{N-1} \frac{|c|}{2} \, [W]^2_{i+\frac{1}{2}} \geq 0 \tag{30}$$

Therefore, in the case of a linear scalar equation, if the numerical flux at cell interfaces is upwinded, the norm of the vector whose components are the coefficients w_i is bounded from above. From eq. 23, one obtains

$$\sum_{j=1}^{M} w_j^2(T) \leq \sum_{j=1}^{M} w_j^2(0) \tag{31}$$

It should be noted, however, that for a centered scheme there is at least marginal stability.

The stability result above suggests the conjecture that *any POD-Galerkin scheme based on an underlying stable finite-differences or finite-volumes scheme is also stable.* In fact, consider a stable first-order accurate integration scheme for a linear scalar conservation law. The norm of solution at time step $n+1$ is bounded by the norm of the solution at time step n, i.e., $\|W^{n+1}\| = \|AW^n\| \leq \|W^n\|$, where A is the discretization matrix of the scheme considered. We have $\|A\| \leq 1$. The POD-Galerkin scheme can be viewed as a projection of this scheme onto a low dimensional space. Taking into account that $\|A\| \leq 1$ and $\|\phi^{(j)}\| = 1$, it is $\|\phi^{(j)} A W^n\| \leq \|W^n\|$. If the span of the POD eigenfunction covers the solution space, i.e., if the Galerkin projection amounts to a change of bases, we have $\|W^n\| = \|\phi^{(j)} \, W^n\|$ and finally $\|\phi^{(j)} A W^n\| \leq \|\phi^{(j)} \, W^n\|$ which is a stability bound for the POD-Galerkin scheme.

We used these results for the Euler equations as explained in the next paragraph.

3.2 First order scheme

Let us consider eq. 19 where we take $S = 0$. Because of the way we have constructed the functions ϕ, we have $\phi_x = 0$ inside each I_i, see fig. 3. Therefore $(\phi^{(j)}_{i+\frac{1}{2}})^- = (\phi^{(j)}_{i-\frac{1}{2}})^+ =$

$\phi^{(j)}(x_i) = \phi_i^{(j)}$ and hence eq. 19

$$\dot{w}_j + \sum_{i=1}^{N} \phi_i^{(j)} \left(F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}} \right) = 0 \tag{32}$$

for every $j = 1, \ldots, M$.

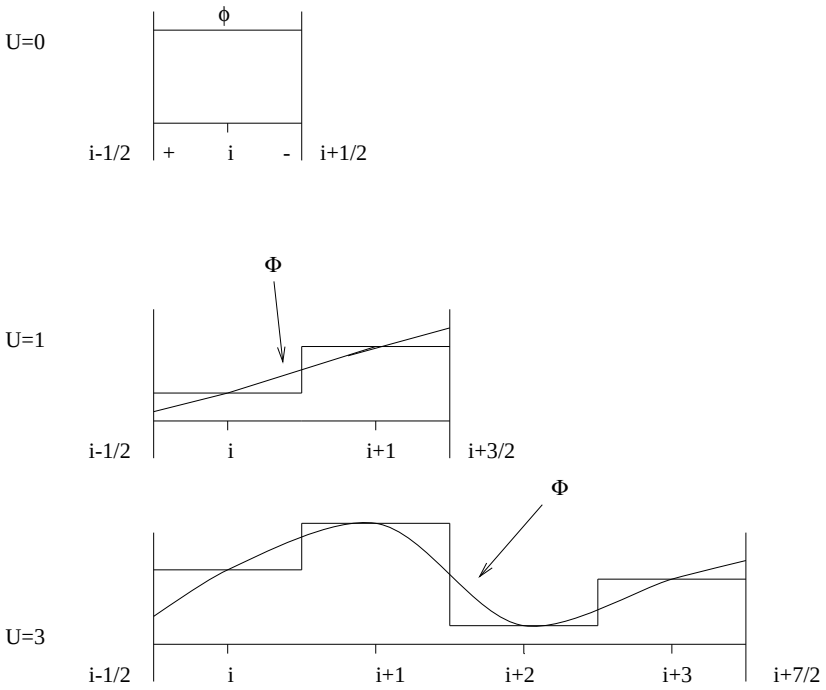

Figure 3: Discretization cells for first order and higher order approximation (Section 3.3).

From the discussion on the linear scalar equation we conclude that in the present case, the non linear system of Euler equations, it is necessary that the numerical fluxes at the cell edges are computed by an upwind type method. We tested some of the well known flux-vector or flux-difference splitting formulations and we found that they preserve the stability of the scheme. In these numerical experiments the time integration was done by a forward Euler step.

In order to have an idea of the timestep allowed with such scheme we went back to the linear scalar case. It can be seen that the POD eigenfunctions for such equation are nothing else than the trigonometric functions, i.e., $\phi_i^{(j)} = 1/\sqrt{2\pi}\ \exp(\iota j i 2\pi/N)$ where $\iota = \sqrt{-1}$ and the computational field extends between 0 and 2π. Assuming periodic boundary conditions,

from eq. 32 we have

$$\dot{w}_j + \frac{c\,N}{2\pi} \sum_{k=-M}^{M} w_k \sum_{i=1}^{N} \frac{2\pi}{N} \bar{\phi}_i^{(j)} \left(\phi_i^{(k)} - \phi_{i-1}^{(k)} \right) = 0 \tag{33}$$

The eigenvalues of

$$A_u = \{a_{j\,k}\} = -\left\{ \sum_{i=1}^{N} \frac{2\pi}{N} \bar{\phi}_i^{(j)} \left(\phi_i^{(k)} - \phi_{i-1}^{(k)} \right) \right\} \tag{34}$$

are plotted on the complex plane ($N = 201$, $M = 100$) in fig. 4. Let the CFL number be

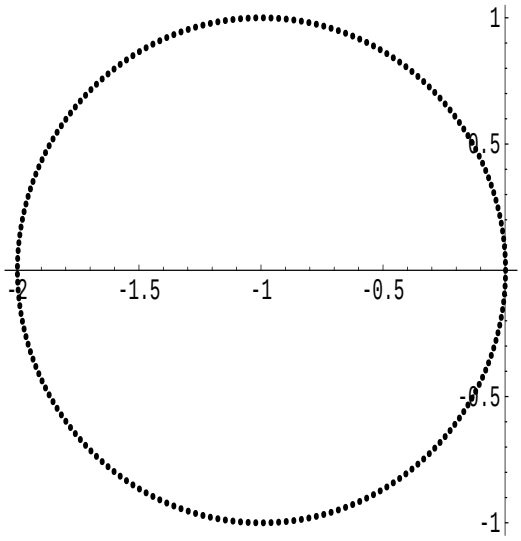

Figure 4: Eigenvalues of the matrix A_u on the complex plane.

$r = c\Delta t N/2\pi$. On the complex plane, the stability region of an explicit first order time integration is a circle of radius 1, with center in $(-1, 0)$. It practically overlaps the region bounded by the eigenvalues of the matrix A_u, therefore when $r \leq 1$ the scheme is stable. If the fluxes in eq. 32 are computed by a centered approximation, i.e., $F_{i+1/2} = 1/2\,(F_i + F_{i+1})$ and $F_{i-1/2} = 1/2\,(F_i + F_{i-1})$, we get

$$A_c = \{a_{j\,k}\} = -\left\{ \sum_{i=1}^{N} \frac{\pi}{N} \bar{\phi}_i^{(j)} \left(\phi_{i+1}^{(k)} - \phi_{i-1}^{(k)} \right) \right\} \tag{35}$$

The eigenvalues of A_c are scattered around the imaginary axis, see fig. 5, indicating marginal stability for such spatial discretization. The same result is obtained from eq. 28. Indeed, in

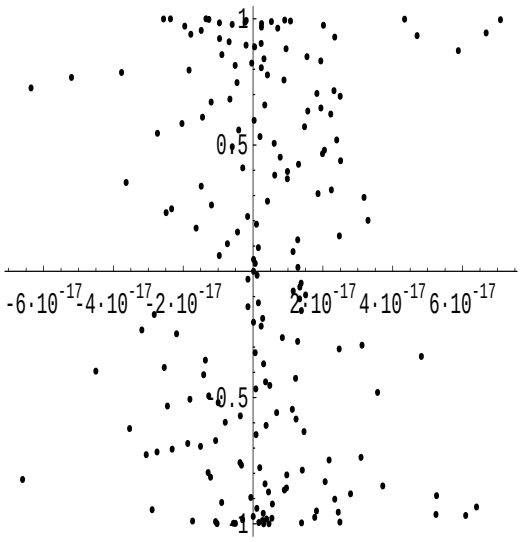

Figure 5: Eigenvalues of the matrix A_c on the complex plane.

the case of a centered scheme we get

$$\sum_{j=1}^{M} w_j \, \Theta_j = 0 \tag{36}$$

Hence it seems that if we do not upwind there is at least marginal stability. Yet, in the experiments on the Euler equations we were not able to obtain a stable solution with Runge-Kutta time integrations up to fourth order.

Although this is a specific example, it is important to notice that in the stable case, since the low frequencies are the one closest to the imaginary axis, using a limited number of POD eigenfunctions we are be able to march with a higher timestep. This is, in fact, the result of several numerical experiments with the Euler equations. Marching explicitly in time and taking 5 eigenfunctions we were able to reach a CFL of 10. Indeed, in relation with fig. 4, it is seen that the stability circle can be shrunk many times and still include the first few eigenvalues.

Let us compare the scheme of eq. 32 to a finite-volumes (FV) first order scheme, for the reason that such schemes are known to be monotone and to converge to the entropy solution. We go back to the linear scalar equation and check under which conditions we are able to recover the first order scheme

$$W_{i,t} + \left(F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}} \right) = 0 \tag{37}$$

where $W_{i,t} = \partial W(x_i)/\partial t$. Take eq. 32 multiply by $\phi_k^{(j)}$ and sum over j from 1 to M to get

$$W_{k,t} + \sum_{i=1}^{N} A_{i\,k} \left(F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}} \right) = 0 \tag{38}$$

where $A_{i\,k} = \sum_{j=1}^{M} \phi_k^{(j)} \phi_i^{(j)}$ is a symmetric positive semidefinite matrix. If $A_{i\,k} = \delta_{i\,k}$ then eq. 37 and eq. 38 would be identical. This is true only if $\phi_i^{(j)}$ is an orthonormal matrix, which is not the case in general. Yet, if $F = c\,W$ and the vector space of the solution $W(x_i)$ is spanned by $\phi_i^{(j)}$ then we have

$$\sum_{i=1}^{N} A_{i\,k} \left(F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}} \right) = F_{k+\frac{1}{2}} - F_{k-\frac{1}{2}} \tag{39}$$

Applying a POD analysis it is expected that if the snapshots are properly chosen, i.e., if the solution space is sampled on a representative set, a limited number of basis functions is sufficient to span the solution vector space.

In general it is possible to prove that if eq. 39 is not satisfied then the total variation of W, i.e. $|W|_{TV} = \sum_i |W_i - W_{i-1}|$, may be increasing with time. We sketch here the proof following [6].

Suppose that the fluxes at cell interfaces can be written as $F_{i+1/2} = F^+(x_i) + F^-(x_{i+1})$ with F^+ and F^- non-decreasing functions of their arguments. Multiply eq. 38 by

$$V_k = - \left(|W_{k+1} - W_k|' - |W_k - W_{k-1}|' \right) \tag{40}$$

where $'$ denotes differentiation with respect to the argument, and sum over k from 1 to N. We have

$$\frac{d}{dt} |W|_{TV} + \sum_{k=1}^{N} \sum_{i=1}^{N} \left\{ V_k\, A_{i\,k} \left[F^+(x_i) - F^+(x_{i-1}) \right] + V_k\, A_{i\,k} \left[F^-(x_{i+1}) - F^-(x_i) \right] \right\} = 0 \tag{41}$$

Multiply also eq. 37 by V_i and sum over i from 1 to N, we have

$$\frac{d}{dt} |W|_{TV} + \sum_{i=1}^{N} \left\{ V_i \left[F^+(x_i) - F^+(x_{i-1}) \right] + V_i \left[F^-(x_{i+1}) - F^-(x_i) \right] \right\} = 0 \tag{42}$$

In eq. 42 it is well known that the sign of the sum is positive, yielding a bound for the total variation. This is not the case for eq. 41: the double summation can be either positive or negative.

Let us see this through a computational example. The computational set up is that explained previously. In this case we have a duct of constant unit area. The pressure at the left boundary is dropped to 0.4 the reference value. We take 100 snapshots of the flow

between $t = 0$ and $t = 1$. In the simulation only 50 eigenfunctions are employed. As mentioned, for this advection type problem the POD basis can yield a good representation of the flow when $0 < t < 1$. In this time interval we expect to have a monotone solution, whose approximation is comparable to that obtained by a first-order FV scheme.

Accordingly compare in fig. 6 the density resulting from the two schemes for $t = 0.7$ and $t = 1.5$. It is seen that the solution of the POD based scheme and the FV scheme overlap for $t = 0.7$. For $t = 1.5$ the eigenfunctions are not able anymore to give a good representation of the solution space and as shown above the scheme is not monotone. From $t = 0$ to $t = 0.7$

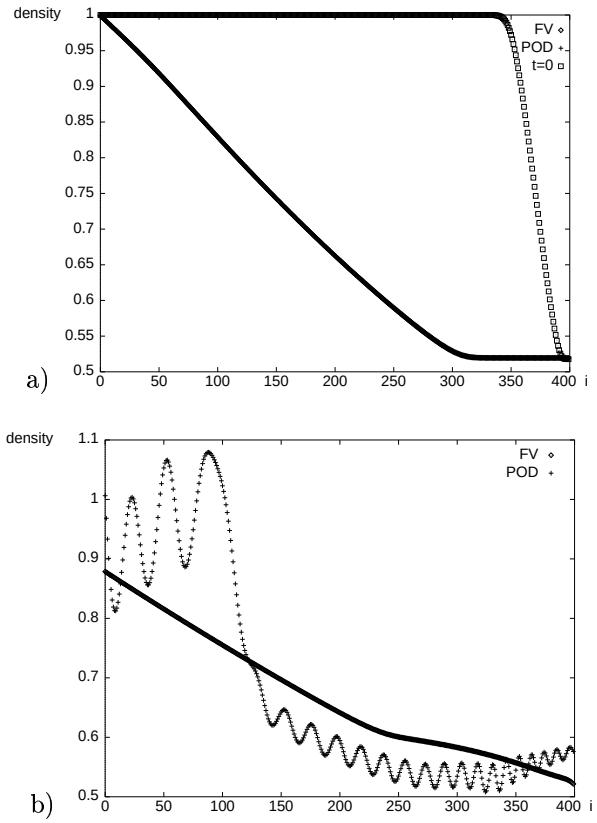

Figure 6: Comparison between solutions computed using the first order POD and the FV scheme: a) $t = 0.7$, b) $t = 1.5$.

the total variation of the density goes from 0.4823 to 0.4794 for the POD scheme and to 0.4796 for the FV scheme. Inversely, when $t = 1.5$ the total variation of the density is 2.836 for the POD scheme and 0.3571 for the FV scheme. It should be noted that the residuals of the POD scheme converge to 0 nevertheless.

As another example, a Sod test case was computed with 400 cells and the results for the density are shown in fig. 7. We took 50 snapshots of the solution and used all of the 50 eigenfunctions. The results shown are obtained after 1200 iterations with CFL= 0.4. From the shock resolution it appears that the POD based scheme is slightly more dissipative compared to the FV scheme. If the simulation is carried beyond the last snapshot taken to build the eigenfunctions, the approximation becomes very poor, like in the previous example. As mentioned, this situation is typical of advection dominated problems. Periodic or quasi-periodic flows are not subject to this limitation.

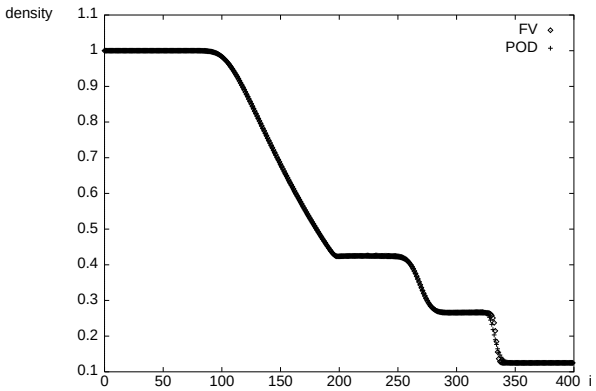

Figure 7: Comparison between solutions computed using the first order POD and the finite-volume scheme: Sod test case.

The first order scheme of this paragraph can easily be extended to two and three dimensions. For example in two dimensions we have

$$\dot{w}_j + \sum_{i=1}^{N} \left\{ \int_{\partial\Omega_i} \phi^{(j)} \left(F\, n_x + G\, n_y\right) d\sigma - \int_{\Omega_i} \left[F\, \phi_x^{(j)} + G\, \phi_y^{(j)} \right] d\Omega \right\} = 0 \tag{43}$$

where $F = [\rho, \rho u^2 + p, \rho uv, u(\rho e + p)]$, $[\rho, \rho uv, \rho v^2 + p, v(\rho e + p)]$, (n_x, n_y) is the unit vector normal to the boundary $\partial\Omega_i$ of the ith discretization cell of area Ω_i. If $\phi^{(j)}$ is taken constant inside the cells we are left with an ODE similar to eq. 32 where the fluxes at cell interfaces must be upwinded.

3.3 Higher-order scheme

The need for upwinding the fluxes at the cell interfaces has a major drawback as long as efficiency is concerned. In fact, compared to a FV scheme, the POD based scheme has the advantage that only a small number of coefficients w_j is updated at each time step, whereas for FV schemes it is necessary to update each cell average. This advantage becomes significant in two and three dimensions. Yet, the computation of the upwind fluxes at

cell interfaces represents the substantial portion of the cost of the simulation. One way to combine efficiency and accuracy is to resort to a higher resolution scheme.

Consider eq. 19. The integrals over the domain are computed as sums of integrals over the intervals $[x_{i-1/2}, x_{i+1/2}]$. These intervals represent the discretization cells of the snapshots from which we are building by POD the set of basis functions we use for the Galerkin projection. The idea is to cluster several cells together so that the basis functions inside the elements are not simply constants, but have some "structure" originating from the POD analysis. In the context of finite elements Cockburn and Shu [10] studied a similar scheme, the Runge-Kutta Discontinuous Galerkin method (RKDG). We refer the reader to their work for what concerns the analysis of the scheme. We adapt here some of those ideas to our scopes.

Suppose we want to create an element which contains two cells. Then the Galerkin projection is made over a set of basis functions which are piecewise linear, whereas in the first order algorithm the basis functions where considered piecewise constant. Similarly, over four cells we fit a Lagrange polynomial of third degree to the cell values of each basis function. In this case the basis functions are considered as piecewise polynomials, see fig. 3. By doing such interpolations we expect second and fourth order accuracy respectively.

Let us denote $L^j_{U,k}$ the Lagrange polynomial which fits the points $\phi^{(j)}(x_k), \phi^{(j)}(x_{k+1}), \ldots, \phi^{(j)}(x_{k+U})$. In this paper we have considered $U = 0, 1, 3$. The approximate solution is sought in the space of discontinuous functions

$$\Phi^{(j)} = \{ \Phi^{(j)}\Big|_k = L^j_{U,k}, \forall k \in I_U \} \tag{44}$$

where I_U is the discretization of the domain in elements containing $U+1$ cells. Replacing in eq. 19 ϕ_j by Φ_j and computing the last two integrals using quadrature rules, we have

$$\begin{aligned} &\dot{w}_j + \sum_{i=1}^{N} \left[\Phi^{(j)}(x_{i+U+\frac{1}{2}})\, F_{i+U+\frac{1}{2}} - \Phi^{(j)}(x_{i-\frac{1}{2}})\, F_{i-\frac{1}{2}} \right] + \\ &- \sum_{i=1}^{N} \sum_{l=0}^{U} \omega_l\, F(x_{i+l})\, \Phi^{(j)}(x_{i+l}) + \sum_{i=1}^{N} \sum_{l=0}^{U} \omega_l\, S(x_{i+l})\, \Phi^{(j)}(x_{i+l}) = 0 \end{aligned} \tag{45}$$

where ω_l are the coefficients of the quadrature rule exact for polynomials of order U.

Concerning time integration, we employed a simple forward Euler step. One might object that this is not consistent with a high resolution in space. However, our aim is not that of increasing time and space resolution. Instead, we aim at reducing the cost of the computation by decreasing the number of fluxes at cell interfaces to be computed. Therefore, given a first order simulation made with, say, 400 cells, we want to recover a similar quality of approximation computing the fluxes over 100 cells. Since the time step is the same in the two cases, it is useless to resort to a more accurate time integration. However, since the eigenvalues of the discrete higher-order operator fall on a curve much flatter on the imaginary axis, with a multi-step time-integration scheme a larger CFL would be allowed.

The test case of a pressure wave propagating inside a straight duct was studied comparing different schemes. We compared the results of the following methods: the second order ENO FV scheme (FV II), the first order ($U = 0$), the second order ($U = 1$), the fourth order ($U = 3$) POD Galerkin scheme. Using 100 snapshots, 70 basis functions were constructed and employed for the Galerkin projection. The pressure at the right end of the duct is set to 0.4 the reference value at $t = 0$. The results after 1200 iterations are shown in fig. 8.

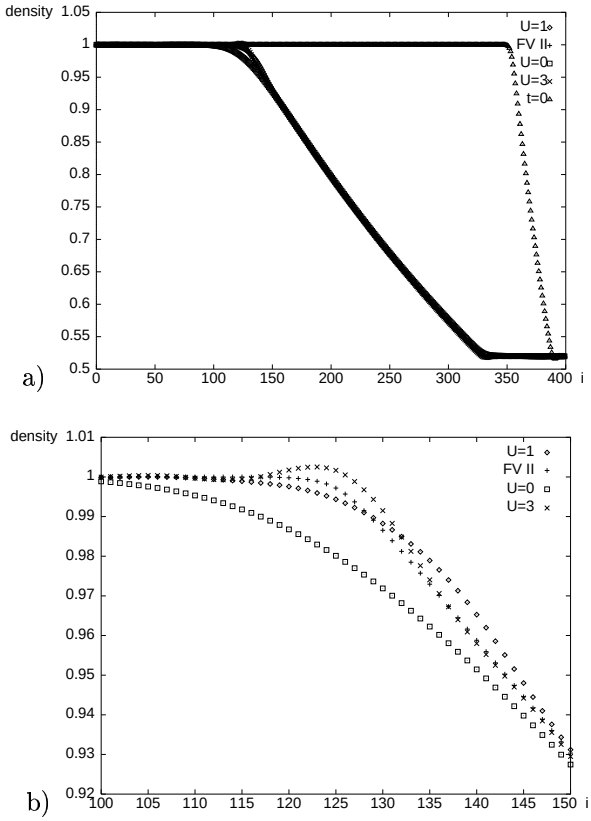

Figure 8: Comparision between solutions computed using different schemes: a) propagation of pressure wave inside a straight duct, b) close up of the results in the vicinity of the front of propagation.

It can be seen that the solution computed with $U = 0$ is the most diffusive. There is some tendency to develop some oscillations as the degree of the approximation increases. Yet, we did not implement any slope limiter as it is suggested in [10]. For $U = 1$ the number of fluxes computed is half that computed for $U = 0$ and this results in about the same ratio between the global time of computation. For $U = 3$, the ratio is $1/4$.

Special care is necessary when dealing with shocks. The higher order scheme proposed above, shows explosive instabilities in the region of the shock. This is true also for contact discontinuities. In order to stabilize the computation in presence of discontinuities, when the entropy gradient s_x reaches a threshold value in a given cell, the contributions of the last two summations in eq. 45 are neglected. The boundary terms corresponding to that cell are computed as if $\Phi^{(j)}$ were constant inside I_U: for $s_{i,x} > (s_{i,x})_{tr}$ we have

$$\Phi^{(j)}(x_{i+U+\frac{1}{2}})\, F_{i+U+\frac{1}{2}} - \Phi^{(j)}(x_{i-\frac{1}{2}})\, F_{i-\frac{1}{2}} = \frac{1}{U+1}\left(F_{i+U+\frac{1}{2}} - F_{i-\frac{1}{2}}\right)\sum_{l=0}^{U}\Phi^{(j)}(x_{i+l}) \quad (46)$$

However, as it seen in the following computational example, this local stabilization degrades the overall accuracy of this scheme.

The Sod test case was studied taking 50 snapshots and 50 eigenfunctions. After 300 iterations the results are those shown in fig. 9. The contact surface is resolved in a similar way by all the schemes, while for the shock wave, the ENO FV scheme is by far more accurate. The stabilization acts only in the region of the discontinuities, nevertheless the resolution of the expansion wave is affected. The nature of this approximation is in a sense similar to the flux-corrected pseudospectral method [11]. It is seen, however, that the resolution of the higher order scheme is better than that of the first-order scheme.

The way to extend these ideas to two and three dimensional flows is clear, yet there are some issues related to the interpolation of the basis functions over the cluster of computational cells and the corresponding degree of spatial approximation which must be investigated.

4 Discussion

One important issue is the possibility of simulating the dynamics of the most energetic modes using the system of ODEs resulting from the projection of the Euler equations on these modes only. Consider the usual expansion test case presented in the previous paragraph and suppose one is interested in the evolution of the first 5 modes. The Euler equations are projected on this five modes only, resulting in a set of 5 ODEs. We compare the values of the coefficients w_j for $j = 1, \ldots, 5$ at several times, to that obtained projecting the second order ENO FV solution on the ϕ_j for $j = 1, \ldots, 5$.

The results are shown in fig. 10. It is seen that there is no substantial difference between the values of w_j and those corresponding to the projection of the ENO FV solution for the first 2 or 3 eigenvalues. For the same test-case, in fig. 11 the solution obtained using 5 modes and the ENO FV solution are compared, after 1600 time steps. In the physical space the accuracy is poor. This is due to the fact that the high frequency components of the solution are missing. However, we conclude that these components have small effect on the dynamic of the first modes which are by far the most energetic.

Another interesting test case is the advection of an entropy layer. We took 100 snapshots for $0 < t < 1.0$ and used only the first 20 eigenfunctions. The results at $t = 0.8$ are shown

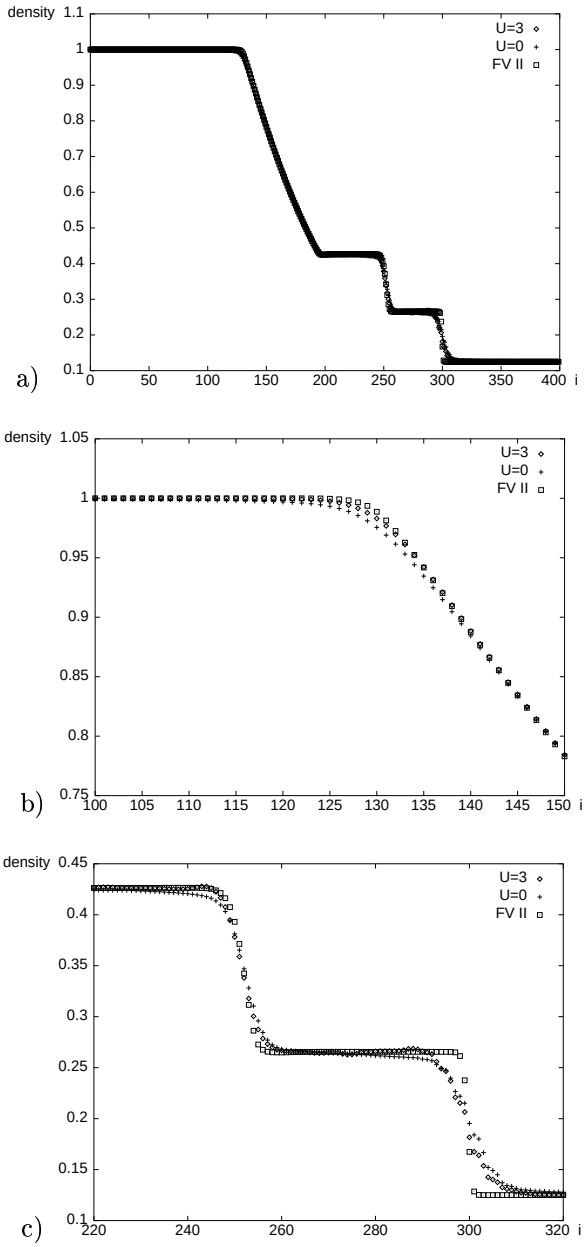

Figure 9: Comparison between solutions computed using different schemes: a) Sod test case, b) close up of the results in the vicinity of the expansion wave, c) close up of the solution in the vicinity of the discontinuities.

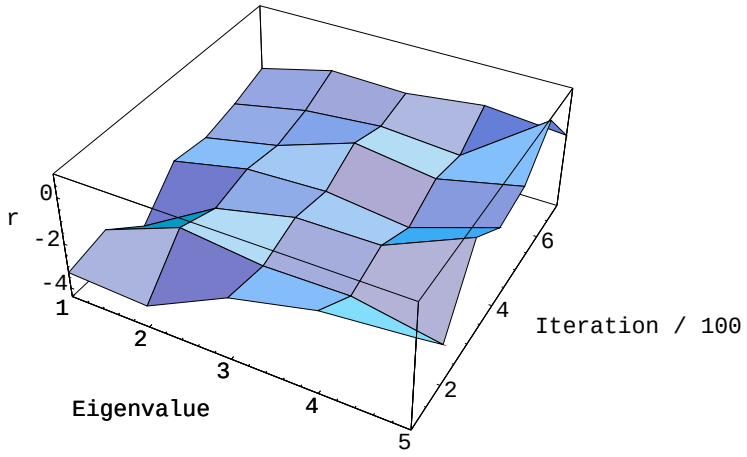

Figure 10: Expansion fan: relative difference between the coefficients w_j computed by time integration of eq. 45 and p_j the projection over ϕ_j of the ENO FV solution at the same time step over ϕ_j. Eigenvalues considered: $j = 1, \ldots, 5$; $r = \log |w_j - p_j|/p_j$.

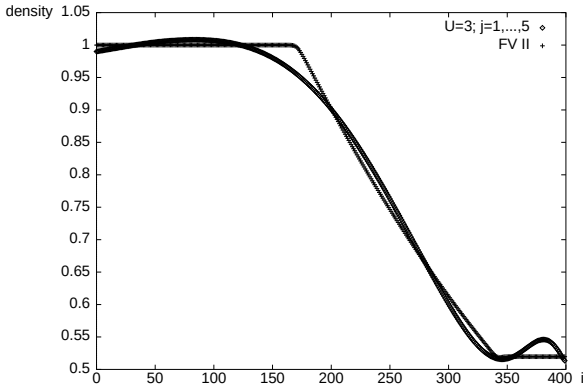

Figure 11: Comparison between ENO FV solution and solution obtained considered only 5 eigenvalues, at given time.

in fig. 12. It is seen that the scheme based on the POD Galerkin formulation ($U = 3$) shows tiny oscillations at the foot of the density layer, apart from that it overlaps the solution obtained by the ENO FV second order scheme.

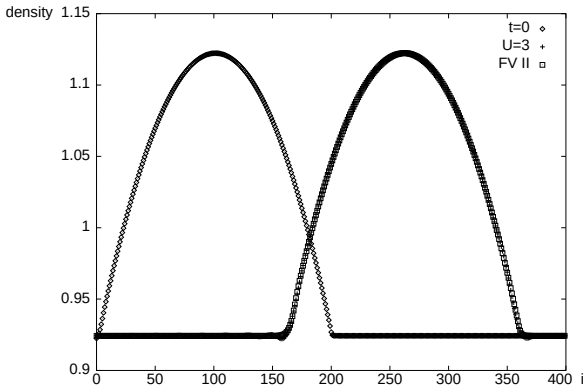

Figure 12: Comparison between ENO FV solution and solution obtained considered 20 eigenvalues after 1000 timesteps ($t = 0.8$).

The main conclusion of this work is that POD based Galerkin schemes for Euler equations need additional stabilization with respect to that provided by a straight-forward dicretization of the projection of the equations in finite dimensional space. For a linear scalar equation we show that upwinding is a satisfactory stabilization. Numerical experiments show that by this mean it is possible to find a stable discretization for Euler equations as well. Besides that, we have shown a way to reduce the computational cost reducing the number of upwind fluxes to be computed at cell interfaces. This was done resorting to a higher-order accuracy approximation. Unfortunately the resolution of shocks by this method is poor.

Aknowledgments

At various stages of the work I had the pleasure of consulting with Dr. Alain Dervieux, INRIA, France, to whom I also owe the interest in this subject. Prof. Bernardo Cockburn, University of Minnesota, USA, introduced me to the numerical analysis techniques for discontinuous Galerkin methods, and I am pleased to thank him. I would like to thank also Dr. Michel Mallet, Dr. Bruno Stoufflet and Mr. Guillaume Vigo of Dassault Aviation, France, for fruitfull discussions. The author was funded by the European Commission under the Marie Curie fellowship program (TMR).

References

[1] J.L. Lumley, 1967. *The structure of inhomogeneous turbulent flow.* In Atmospheric Turbulence and Wave Propagation, eds. A.M. Yaglom and V.I. Tatarski, pp.166-178. Nauka, Moscow.

[2] G. Berkooz, P. Holmes, J.L. Lumley, 1993. *The proper orthogonal decomposition in the analysis of turbulent flows.* Annual Review of Fluid Mechanics, Vol. 25, pp. 539-575.

[3] L. Sirovich, 1987. *Turbulence and the dynamics of coherent structures. Part I: coherent structures.* Quarterly of Applied mathematics, Vol. XLV, n. 3, pp. 561-571.

[4] N. Aubry, P. Holmes, J.L. Lumley, E. Stone, 1988. *The dynamics of coherent structures in the wall region of a turbulent boundary layer.* Journal of Fluid Mechanics, Vol. 192, pp. 115-173.

[5] K.Y. Tang, W.R. Graham, J. Peraire, 1996. *Active flow control using a reduced order model and optimum control.* AIAA Paper 96-1946.

[6] B. Cockburn, 1997. *An introduction to the discontinuous Galerkin method for convection-dominated problems.* In proceedings of the CIME summer school on "Advanced Numerical Approximation of Nonlinear Hyperbolic Equations", ed. A. Quarteroni.

[7] R. Courant, D. Hilbert, 1953. *Methods of mathematical physics.* Vol. I. John Wiley & Sons, New York.

[8] M. Pandolfi, 1984. *A contribution to the numerical prediction of unsteady flows.* AIAA Journal, Vol. 22, n. 5, pp. 602-610.

[9] A. Harten, B. Engquist, S. Chakravarthy, 1987. *Uniformly high order accurate essentially non-oscillatory schemes,iii.* Journal of Computational Physics, Vol. 71, pp. 231-303.

[10] B. Cockburn, C.W. Shu, 1997. *The Runge-Kutta discontinuous Galerkin method for conservation laws V: multidimensional systems.* ICASE Report 97-43. NASA CR 201737.

[11] B.E. McDonald, 1989. *Flux-corrected pseudospectral method for scalar hyperbolic conservation laws.* Journal of Computational Physics, Vol. 82, pp. 413-428.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - B.P. 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - B.P. 101 - 54602 Villers lès Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot St Martin (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, B.P. 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399