von Karman Institute for Fluid Dynamics

Lecture Series 2003-04

POST-PROCESSING OF EXPERIMENTAL AND NUMERICAL DATA

February 17-21, 2003

PROPER ORTHOGONAL DECOMPOSITION: AN OVERVIEW

L. Cordier & M. Bergmann
LEMTA, France

Proper Orthogonal Decomposition: an overview

Cordier Laurent & Bergmann Michel
Laboratoire d'Energétique et de Mécanique Théorique et Appliquée
UMR 7563 (CNRS - INPL - UHP)
ENSEM - 2, avenue de la Forêt de Haye
BP 160 - 54504 Vandoeuvre Cedex, France

Contents

1 Introduction

Collecting very large amounts of data by numerical simulations or experimental approaches is a common situation in almost any scientific field. There is therefore a great need to have specific post-processing techniques able to extract from these large quantities of high dimensional data, synthetic information essential to understand and eventually to model the processes under study. The Proper Orthogonal Decomposition (POD) is one of the most powerful method of data-analysis for multivariate and non linear phenomena. Essentially, POD is a linear procedure that takes a given collection of input data and creates an orthogonal basis constituted by functions estimated as the solutions of an integral eigenvalue problem known as a Fredholm equation (see equation 18). These eigenfunctions are by definition (equation 16) characteristic of the most probable realizations of the input data. Moreover, it can be shown that they are optimal in terms of the representation of energy present within the data (see § 4.3).

1.1 Historical background of POD

Historically, the Proper Orthogonal Decomposition was introduced in the context of turbulence by Lumley (see Lumley, 1967) as an objective definition of what was previously called "big eddies" by Townsend (1976) and what is now widely known as Coherent Structures (hereafter denoted CS for simplicity, see Eddy Structure Identification Course, 1996 for a detailed discussion of CS and an overview of their detection methods). According to Yaglom (see Lumley, 1970), the POD is a natural idea to replace the usual Fourier decomposition in nonhomogeneous directions. The POD method was then introduced for different purposes independently by several scientists, in particular, by Kosambi (1943), Loève (1945, 1955), Karhunen (1946), Pougachev (1953) and Obukhov (1941, 1954). This technique is then known under a variety of names: Karhunen-Loève decomposition or expansion, Principal Component Analysis (Joliffe, 1986) or Hotelling Analysis (Hotelling, 1933), Singular Value Decomposition (Golub and Van Loan, 1990). Naturally, the Proper Orthogonal Decomposition has been widely used in studies of turbulence but other popular applications involve random variables (Papoulis, 1965), image processing as for example characterization of human faces (Kirby and Sirovich, 1990), signal analysis (Algazi and Sakrison, 1969), data compression (Andrews *et al.*, 1967) and recently optimal control (Ravindran, 2000a,b).

From a mathematical point of view, the Proper Orthogonal Decomposition is just a transformation which diagonalizes a given matrix A and brings it to a canonical form $A = U\Sigma V^\dagger$ where Σ is a diagonal matrix (see §3 for a complete description). The mathematical content of POD is therefore classical and is based in the spectral theory of compact, self adjoint operators (see Courant and Hilbert, 1953). Two geometric interpretations of this mathematical procedure are discussed in §3.2.

1.2 POD and turbulent flows

A complete literature review on applications of POD to turbulence is far beyond the scope of these lecture notes: really good reviews can be found in Holmes *et al.* (1996), Delville *et al.* (1999) and in the appendix of Gordeyev (1999). In the following, we remind briefly what insight can be gained from the use of POD for eduction and modelling of the Coherent Structures observed in most turbulent flows.

For our purposes, it is sufficient to have in mind (see Bonnet and Delville, 2002) that CS identification has to be done at least for two reasons: first,

from an energetic point of view because the relative energy content of the CS as compared with the total turbulent energy can represent from 10% (for boundary layers, far jets), up to 20% (far wakes, plane mixing layers) or 25% (near wakes or jets) (see Fiedler, 1998). Second, because the dynamical properties of CS play an essential role in mixing processes, drag, noise emission, etc... For these reasons, the idea of controlling turbulent flows by means of influencing their coherent structures seems promising (see Aubry *et al.*, 1988; Ukeiley *et al.*, 2001).

Several characteristics of the Proper Orthogonal Decomposition technique, as introduced by Lumley (1967), are quite attractive in terms of CS identification. Firstly, compared to many other classical methods used for large scale identification (flow visualization, conditional methods, VITA, Pattern Recognition Analysis), no a priori is needed for the eduction scheme. CS are defined in an objective and unique manner as the flow realization that possesses the largest projection onto the flow field (see equation 16). Secondly, the POD yields to an optimal set of basis functions in the sense that no other decomposition of the same order captures an equivalent amount of kinetic energy (§4.3). Up to now, POD is only presented as a data analysis method which takes as input an ensemble of data, obtained from physical experiments or from detailed numerical simulations and extracts basis functions optimal for the representativeness of the data. For illustrative purposes of the ability of the Proper Orthogonal Decomposition to educe CS, POD is applied in Cordier and Bergmann (2002) to a data-base obtained by Large Eddy Simulation of a three-dimensional plane turbulent mixing layer. However, the Proper Orthogonal Decomposition can be used as well as an efficient procedure to compute low-dimensional dynamical models of the CS. The reduced-order modelling by POD is based on projecting the governing equation of motion onto subspaces spanned by the POD basis functions (Galerkin projection) yielding to a simple set of Ordinary Differential Equations (ODE). Finally, due to the optimality of convergence in terms of kinetic energy of the POD functions, only a small number of POD modes are necessary to correctly represent the dynamical evolution of the flow. Reduced order modelling based on POD has recently received an increasing amount of attention for applications to optimal control problems for partial differential equations (see Hinze, 2000; Volkwein, 2001; Fahl, 2000). In Cordier and Bergmann (2002), a low-order model based on POD is developed for the incompressible unsteady wake flow behind a circular cylinder at a Reynold's number of 200. In particular, it will be demonstrated how the control action can be incorporated in the low-dimensional model.

This course is devoted to a rigorous presentation of the Proper Orthogonal

Decomposition. First, the Proper Orthogonal Decomposition is introduced in the general context of the approximation theory (§2). Since POD can be approached as an application of the Singular Value Decomposition (SVD), this decomposition is then presented in §3. The relationship between POD and SVD is discussed in § 3.3. In §4, POD is now described in a statistical setting using an averaging operation for use with turbulent flows. The different POD approaches is then extensively discussed in §5 where the main differences between the classical POD (§5.3) and the snapshot POD (§5.4) are particularly highlighted. Before to conclude in §7 by an evaluative summary of the POD approach, §6 presents the Proper Orthogonal Decomposition as a generalization of the classical Fourier analysis to inhomogeneous directions.

2 Proper Orthogonal Decomposition as an approximation method

In these lecture notes, we decide to follow the view of Chatterjee (2000) and to introduce the Singular Value Decomposition (§3) and his generalization, the Proper Orthogonal Decomposition (§4) in the general context of approximation theory (see Rivlin, 1981).

Suppose we want to approximate a possibly vector-valued function $u(\vec{\boldsymbol{x}}, t)$ over some domain of interest $\mathcal{D} = \Omega \times [0; T]$ as a finite sum in the variables-separated form:

$$u(\vec{\boldsymbol{x}}, t) \simeq \sum_{k=1}^{K} a^{(k)}(t)\phi^{(k)}(\vec{\boldsymbol{x}}) \tag{1}$$

For simplicity and because it will be the case in Fluid Mechanics applications, $\vec{\boldsymbol{x}}$ can be viewed as a spatial coordinate and t as a temporal coordinate.

Our expectation is that this approximation becomes exact as $K \longrightarrow +\infty$. The representation (1) is clearly not unique. A classic way to solve this approximation problem is to use for the basis functions $\phi^{(k)}(\vec{\boldsymbol{x}})$, functions given a priori as for example Fourier series, Legendre polynomials or Chebyshev polynomials... An alternative way could be to determine functions $\phi^{(k)}(\vec{\boldsymbol{x}})$ intrinsic by nature of the function $u(\vec{\boldsymbol{x}}, t)$ to approximate. As it will be explained in the following, this particular approach corresponds to the Proper Orthogonal Decomposition (POD).

An additional difficulty is that for each choice of basis functions $\phi^{(k)}(\vec{\boldsymbol{x}})$ corresponds a different sequence of time-functions $a^{(k)}(t)$. Then, given $\phi^{(k)}(\vec{\boldsymbol{x}})$,

how can we determine the coefficients $a^{(k)}(t)$? Suppose we have chosen orthonormal basis functions, i.e.:

$$\int_\Omega \phi^{(k_1)}(\vec{\boldsymbol{x}})\, \phi^{(k_2)}(\vec{\boldsymbol{x}})\, d\vec{\boldsymbol{x}} = \delta_{k_1 k_2} \tag{2}$$

where

$$\delta_{k_1 k_2} = \begin{cases} 0 & \text{for } k_1 \neq k_2, \\ 1 & \text{for } k_1 = k_2, \end{cases}$$

is the Kronecker delta symbol, then:

$$a^{(k)}(t) = \int_\Omega u(\vec{\boldsymbol{x}}, t)\phi^{(k)}(\vec{\boldsymbol{x}})\, d\vec{\boldsymbol{x}}.$$

Therefore for orthonormal basis functions, $a^{(k)}(t)$ depends only on $\phi^{(k)}(\vec{\boldsymbol{x}})$ and not on the other ϕ's. So for selecting the function $\phi^{(k)}(\vec{\boldsymbol{x}})$, it would be useful to use orthonormality condition.

Moreover, while an approximation to any desired accuracy can always be obtained if K can be chosen large enough, we may like to find, once and for all, a sequence of orthonormal functions $\phi^{(k)}(\vec{\boldsymbol{x}})$ in such a way that the approximation for each K is as good as possible in a least square sense. Now consider that we can measure (experimentally or numerically) at N_t different instants of time, M realizations of $u(\vec{\boldsymbol{x}}, t)$ in M different locations $x_1, x_2, \cdots, x_M$. The approximation problem (1) is then equivalent to find orthonormal functions $\{\phi^{(m)}(\vec{\boldsymbol{x}})\}_{k=1}^{K}$ with $K \leq N_t$ solving:

$$\min \sum_{i=1}^{N_t} \|u(\vec{\boldsymbol{x}}, t_i) - \sum_{k=1}^{K} \left(u(\vec{\boldsymbol{x}}, t_i), \phi^{(k)}(\vec{\boldsymbol{x}})\right) \phi^{(k)}(\vec{\boldsymbol{x}})\|_2^2 \tag{3}$$

where $\|.\|_2$ define the norm associated to the usual L^2 inner product $(.,.)$. Remind that for any vector $\vec{\boldsymbol{y}} \in \mathbb{R}^M$, we have:

$$\vec{\boldsymbol{y}} = \begin{bmatrix} y_1 \\ \vdots \\ y_M \end{bmatrix} \Longrightarrow \|\vec{\boldsymbol{y}}\|_2 = (\vec{\boldsymbol{y}}, \vec{\boldsymbol{y}})^{1/2} = \sqrt{\vec{\boldsymbol{y}}^T \vec{\boldsymbol{y}}} = \sqrt{y_1^2 + \cdots + y_M^2} \tag{4}$$

The practical method to solve the minimization problem (3) is to arrange the set of data $\mathcal{U} = \{u(\vec{\boldsymbol{x}}, t_1), \cdots, u(\vec{\boldsymbol{x}}, t_{N_t})\}$ in an $M \times N_t$ matrix A called

the Snapshot Data Matrix.

$$A = \begin{pmatrix} u(x_1,t_1) & u(x_1,t_2) & \cdots & u(x_1,t_{N_t-1}) & u(x_1,t_{N_t}) \\ u(x_2,t_1) & u(x_2,t_2) & \cdots & u(x_2,t_{N_t-1}) & u(x_2,t_{N_t}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ u(x_M,t_1) & u(x_M,t_2) & \cdots & u(x_M,t_{N_t-1}) & u(x_M,t_{N_t}) \end{pmatrix} \in \mathbb{R}^{M\times N_t} \tag{5}$$

Each column $A_{:,i} \in \mathbb{R}^M$ of the snapshot data matrix represents a single snapshot $u(\vec{\boldsymbol{x}}, t_i)$ of the input ensemble $\mathcal{U}$. We can note that if the snapshot data are assumed to be linearly independent (this will be the case in particular for the snapshot POD method for reasons explained in §5.4.1) then the snapshot data matrix has full column rank.

The solutions of the problem of minimization (3) are given by the truncated Singular Value Decomposition of length K of the matrix A. For this reason, the Singular Value Decomposition of a matrix is reviewed in §3. The relationship between the Proper Orthogonal Decomposition and the Singular Value Decomposition is addressed in §3.5.

3 The Singular Value Decomposition (SVD)

3.1 Definition of SVD

Let A be a general complex M by N_t matrix. The Singular Value Decomposition (SVD) of A is the factorization (see Golub and Van Loan, 1990):

$$A = U\Sigma V^\dagger \tag{6}$$

where[1] U and V are (non unique) unitary[2] $M \times M$ respectively $N_t \times N_t$ matrices, i.e. $UU^\dagger = I_M$ and $VV^\dagger = I_{N_t}$, and $\Sigma = diag(\sigma_1, \cdots, \sigma_r,)$ with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq 0$ where $r = \min(M, N_t)$. The rank of A equals the number of nonzero singular values it has.

The σ_i are called the singular values of A (and also of $A^\dagger$), the first r columns of $V = (v_1, v_2, \cdots, v_{N_t})$ the right singular vectors and the first r columns of $U = (u_1, u_2, \cdots, u_M)$ the left singular vectors. Since the singular values are arranged in a specific order, the index i of the i^{th} singular value will be called the singular value number.

[1]Here, $V^\dagger$ denotes the adjoint matrix of V defined as the conjugate transpose of V.

[2]Remind that for a unitary matrix $A^{-1} = A^\dagger$. If $A \in \mathbb{R}^{M\times N_t}$ then $V^\dagger = V^T$ and V is said orthogonal.

3.2 Geometric interpretations of SVD

3.2.1 Geometric structure of a matrix

By definition of a matrix, an $M \times N_t$ matrix A is a linear operator that maps vectors from an N_t dimensional space, say $\mathcal{E}_{N_t}$, to an M dimensional space, say $\mathcal{E}_M$. Imagine the unit sphere in $\mathcal{E}_{N_t}$ (the set of vectors of unit magnitude). Multiplication of these vectors by the matrix A results in a set of vectors which defines a r dimensional ellipsoid in $\mathcal{E}_M$, where r is the number of non-zero singular values. The singular values $\sigma_1, \sigma_2, \cdots, \sigma_r$ are the lengths of the principal radii of that ellipsoid (see figure 1). Intuitively, the singular values of a matrix describe the extent to which multiplication by the matrix distorts the original vector. Moreover, since the matrix V is unitary, equation (6) writes $AV = U\Sigma$. The consequences are that the directions of these principal radii are given by the columns of U and the pre-images of these principal radii are the columns of V. A second geometric interpretation is given in the next section.

Due to the interpretation of the matrix A in terms of linear algebra, it is now obvious that the 2-induced norm of A is σ_1:

$$\|A\|_2 = \max_{\|x\|=1} \|Ax\|_2 = \sigma_1 \tag{7}$$

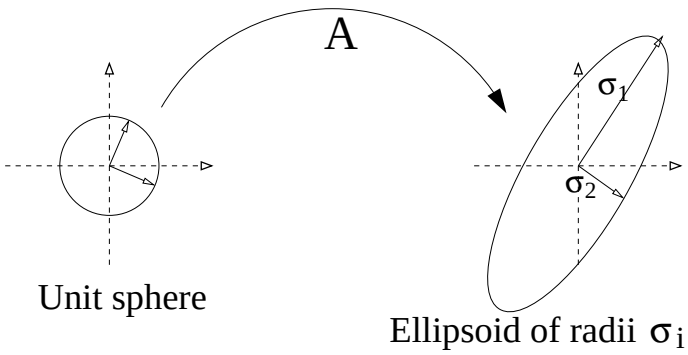


Figure 1: Geometric interpretation of the SVD of the matrix A: image by A of a unit sphere.

3.2.2 SVD as a phase space rotation

A second geometric interpretation may be attributed to SVD applications. We now view the $M \times N_t$ matrix A as a list of coordinates of M points denoted $P_1, P_2, \cdots, P_M$ in an N_t dimensional space. Each point P_i is represented in

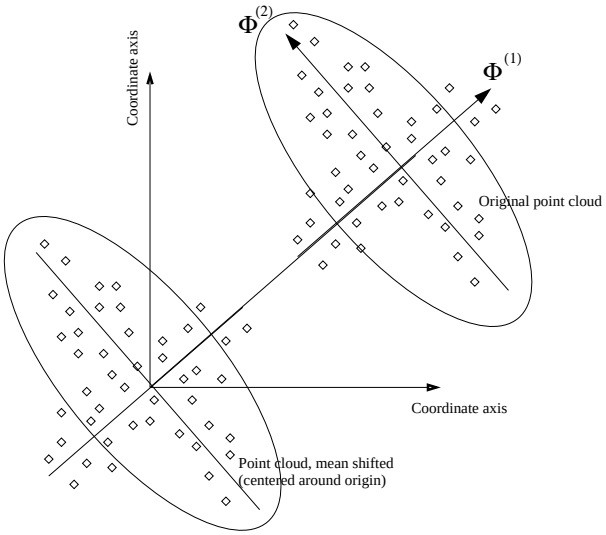


Figure 2: Geometric interpretation of the SVD of the matrix A: phase space rotation.

figure 2 by a diamond. For any $k \leq N_t$, we seek a k-dimensional subspace for which the mean square distance of the points, from the subspace, is minimized i.e. we search a vector $\vec{\mathbf{\Phi}}^{(1)}$ (see figure 2) such that $\sum_{i=1}^{M} |\vec{\boldsymbol{H_i P_i}}|^2$ is minimized where H_i are the orthogonal projection of P_i onto the line of direction vector $\vec{\mathbf{\Phi}}^{(1)}$. This mathematical procedure can be geometrically interpreted (see figure 2) as a rotation of the phase space from the original basis into a new coordinate system whose orthogonal axes coincide with the axes of inertia of the data. This formulation of the SVD problem corresponds exactly to the way the Principal Component Analysis is commonly introduced in the literature (see Joliffe, 1986).

When the Singular Value Decomposition is used for data-analysis, the SVD algorithm is generally applied to a matrix deduced from the snapshot matrix A by subtracting from each column of A the mean of that column. This mean shift ensures that the M point cloud is now centered around the origin of the coordinate (see figure 2).

3.3 Relationships between SVD and eigenvalue problems

In this section, we present how the singular values and the right and left singular vectors of a rectangular matrix A can also be computed by solving symmetric eigenproblems with e.g., matrices $A^\dagger A$ or $AA^\dagger$, instead of comput-

ing the SVD of A. In this case, $A^\dagger A$ and $AA^\dagger$ represent a finite-dimensional version of the two-point space-time correlation R introduced in §4.1. The results of this section will be used in §3.5.

Let $A = U\Sigma V^\dagger$ be a singular value decomposition of $A \in \mathbb{R}^{M\times N_t}$. Then $A^\dagger A = V\Sigma U^\dagger U\Sigma V^\dagger = V\Sigma^2 V^\dagger$ where Σ^2 is a diagonal matrix. Since $A^\dagger A$ is an hermitian matrix, its eigenvalue-decomposition writes: $A^\dagger A = W\Lambda W^{-1} = W\Lambda W^\dagger$ where W is an $N_t \times N_t$ unitary matrix. By comparing the two expression of A, we conclude that $\Sigma^2 = \Lambda$, and $W = V$. In other words: $\sigma_i = \sqrt{\lambda_i}$, and (V, Λ) is the eigenvector-eigenvalue decomposition of $A^\dagger A \in \mathbb{R}^{N_t\times N_t}$.

The same development applied to the matrix $AA^\dagger$ leads to: $AA^\dagger = U\Sigma V^\dagger V\Sigma U^\dagger = U\Sigma^2 U^\dagger = W\Lambda W^\dagger$, so (U, Λ) is the eigenvector-eigenvalue decomposition of $AA^\dagger \in \mathbb{R}^{M\times M}$.

At this point, we remark that the eigenvalue problem associated to $A^\dagger A$ is more practical to solve than the eigenvalue problem associated to $AA^\dagger$ in cases where the input collection N_t is significantly smaller than the number of coefficients needed to represent each item of the collection M. This remark explains that two different POD approaches exist: the classical POD (§5.3) and the snapshot POD (§5.4).

3.4 Lower-rank approximation to A

Given $A \in \mathbb{R}^{M\times N_t}$, the computation of a matrix $X \in \mathbb{R}^{M\times N_t}$ with $\mathrm{rank}\,(X) = k < \mathrm{rank}\,(A)$ such that an appropriate norm of the error $E = A - X$ is minimized, is a classical problem. This problem can be solved explicitly if we take as norm the Frobenius norm, defined as the square root of the sums of squares of all the elements and denoted as $\|.\|_F$ or any unitarily invariant norm[3]. The solution is given by the Eckart-Young theorem (see Higham, 1989) which states that:

$$\min_{\mathrm{rank}\,(X)\le k} \|A - X\|_F = \|A - A_k\|_F = \sqrt{\sum_{i=k+1}^{r} \sigma_i^2(A)} \qquad (8)$$

[3]For example, the 2-norm defined by equation (7) can be used. In this case, the Eckart-Young theorem (8) writes (see Hubert *et al.*, 2000):

$$\min_{\mathrm{rank}\,(X)\le k} \|A - X\|_2 = \|A - A_k\|_2 = \sigma_{k+1}(A) \ .$$

where:

$$A_k = U \begin{pmatrix} \Sigma_k & 0 \\ 0 & 0 \end{pmatrix} V^\dagger = \sigma_1 u_1 v_1^\dagger + \cdots + \sigma_k u_k v_k^\dagger$$

with Σ_k the matrix obtained by setting $\sigma_{k+1} = \sigma_{k+2} = \cdots = \sigma_r = 0$ in Σ.

Remark: This theorem establishes a relationship between the rank k of the approximant, and the $(k+1)^{th}$ largest singular value of A. Therefore, if the singular values decrease is fast, we can hope to find an approximant with small rank (see §3.6 for applications of SVD to data-image processing).

3.5 Relationship between POD and SVD

Here, we discuss the close relationship between POD and SVD. Our presentation follow the view of Fahl (2000) but similar treatments can be found in Atwell and King (1999). The reader is referred to Volkwein (1999) for the mathematical demonstrations.

Suppose that each member of the input collection $\mathcal{U}$ defined in §2 can be written in terms of an n^{th} order finite element basis functions $\{\varphi^{(j)}(\vec{\boldsymbol{x}})\}_{j=1}^n$, i.e.:

$$u(\vec{\boldsymbol{x}}, t_i) = u^n(\vec{\boldsymbol{x}}, t_i) = \sum_{j=1}^{n} u^{(j)}(t_i)\varphi^{(j)}(\vec{\boldsymbol{x}})$$

where the superscript n denotes a high-order finite element discretization.

The inner product can then be defined by:

$$(u, v)_{\mathcal{M}} = \vec{\boldsymbol{u}}^T \mathcal{M}\, \vec{\boldsymbol{v}} \tag{9}$$

where $\mathcal{M} \in \mathbb{R}^{n\times n}$ is the finite element mass matrix and $\vec{\boldsymbol{u}}$, $\vec{\boldsymbol{v}} \in \mathbb{R}^n$ are the finite element coefficient vectors for a given t_i. Employing a Cholesky factorization $\mathcal{M} = \mathcal{M}^{1/2}\left(\mathcal{M}^{1/2}\right)^T$, the $\mathcal{M}$ inner product (9) can be transformed to the standard Euclidean inner product (4) such that the condition

$$\|u\|_{\mathcal{M}} = (u, u)_{\mathcal{M}}^{1/2} = \| \left(\mathcal{M}^{1/2}\right)^T \vec{\boldsymbol{u}}\|_2$$

holds. The minimization problem (3) can then be reformulated for the $\mathcal{M}$ inner product as:

$$\min \sum_{i=1}^{N_t} \|u^n(\vec{\boldsymbol{x}}, t_i) - \sum_{k=1}^{K} \left(u^n(\vec{\boldsymbol{x}}, t_i), \phi^{(k)}(\vec{\boldsymbol{x}})\right)_{\mathcal{M}} \phi^{(k)}(\vec{\boldsymbol{x}})\|_{\mathcal{M}}^2 \tag{10}$$

where the POD basis functions $\{\phi^{(k)}\}_{k=1}^{K}$ are assumed to be in the linear space spanned by the finite element basis functions $\{\varphi^{(j)}(\vec{\boldsymbol{x}})\}_{j=1}^{n}$, i.e.:

$$\phi^{(k)}(\vec{\boldsymbol{x}}) = \sum_{j=1}^{n} \phi_j^{(k)} \varphi^{(j)}(\vec{\boldsymbol{x}})$$

In order to reformulate the minimization problem (10) in a matrix approximation context, let $\Phi \in \mathbb{R}^{n \times K}$ denote a matrix collecting the finite element coefficients of the unknown POD functions. Since for any matrix $\hat{A} \in \mathbb{R}^{n \times N_t}$, $\sum_{i=1}^{N_t} \|\hat{A}_{:,i}\|_2^2 = \|\hat{A}\|_F$, where $\|.\|_F$ denotes the Frobenius norm defined in §3.4, the problem (10) is equivalent to solving:

$$\min_{Z \in \mathbb{R}^{n \times K}} \|\hat{A} - ZZ^T\hat{A}\|_F^2 \quad \text{s.t.} \quad Z^T Z = I_K \tag{11}$$

with $\hat{A} = \left(\mathcal{M}^{1/2}\right)^T A$ and $Z = \left(\mathcal{M}^{1/2}\right)^T \Phi \in \mathbb{R}^{n \times K}$.

Equation (11) indicates that we are looking for a K dimensional subspace with orthogonal matrix Z such that $X = ZZ^T\hat{A}$ is the best approximation to $\hat{A}$ compared to all subspaces of dimension K. According to the Eckart-Young theorem (8), the solution to problem (11) is given by a truncated Singular Value Decomposition of $\hat{A}$ of length K:

$$\hat{A}_K = U_K \Sigma_K V_K^T \tag{12}$$

where U_K, V_K correspond to the first K columns of U, V respectively. Finally, comparing $\hat{A}_K$ and the form of X, we found that the matrix Φ solves:

$$\left(\mathcal{M}^{1/2}\right)^T \Phi = U_K \in \mathbb{R}^{n \times K}. \tag{13}$$

The finite element coefficients of the POD basis functions can then be computed by solving the linear system (13) where the left singular vectors U of $\hat{A} = U\Sigma V^T$ can be obtained directly as the eigenvalues of the matrix $\hat{A}\hat{A}^T$ (see §3.3). However, as it was previously remarked at the end of §3.3, when N_t is significantly smaller than n then it is more practical to solve the eigenvalue problem $\hat{A}^T\hat{A}$. It follows that, in this case, the right singular vectors V of $\hat{A}$ is obtained and U must be deduced from V by the equation $U = \Sigma^{-1}\hat{A}V$.

Remarks: The eigenvalue problems can be solved with the library LAPACK[4] and efficient algorithms for POD computations based on Lanczos iterations can be found in Fahl (2000).

[4] http://www.netlib.org/lapack/

3.6 Examples of image processing by SVD

As an illustration of the SVD process for computing low-rank approximations to data matrices (§3.4), consider a time-independent problem where the input collection consists of greyscale images. In figure 5(a) and 6(a) the "clown" picture and the "trees" picture from MatLab are considered. These images can be represented by means of a 200×330, and a 128×128 matrix respectively, each entry (pixel) having a value between 0 (white) and 1 (black) in 64 levels of gray. Both matrices have full rank, i.e. 200 and 128 respectively. Their numerical rank however are much lower. The singular values of these two images are shown in figure 3 on a semi-log scale; both sets of singular values fall-off rapidly, and hence low-rank approximations with small error are possible.

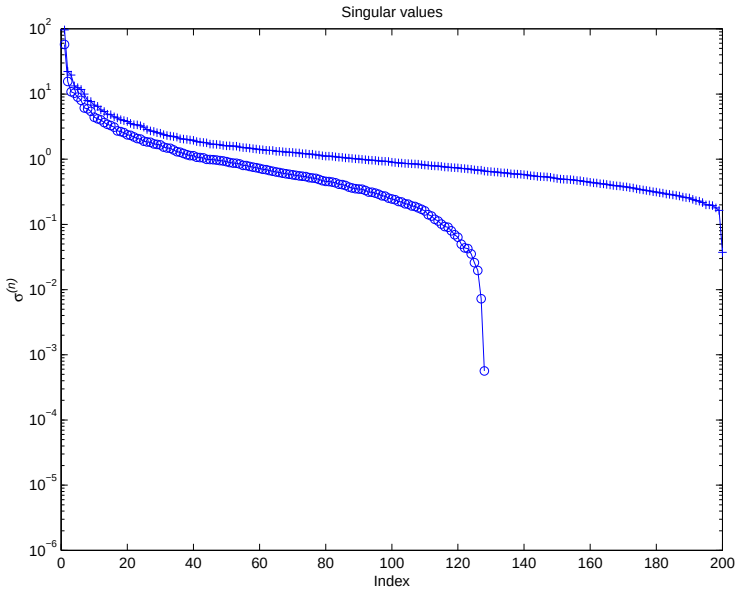


Figure 3: Singular values for the "clown" image (+) and the "trees" image (∘).

By comparing the spectrum of the two singular value plots, we can determine that the relative error for approximants of the same rank is greater for the "clown" image than for the "trees" image. Thus the "trees" image is easier to approximate.

Eckart-Young theorem (see §3.4) stated that for any matrix A of rank N,

an approximation of rank $k \leq N$ of the matrix A can be obtained by:

$$A = \sigma_1 u_1 v_1^\dagger + \sigma_2 u_2 v_2^\dagger + \cdots + \sigma_k u_k v_k^\dagger$$

Thus using the Singular Value Decomposition, one can obtain a high fidelity model perhaps with large k. In order to obtain a lower rank representation of these images, singular modes corresponding to small singular values are neglected. So if the spectrum of the singular values decays fast, one can choose a cutoff value $M \ll N$ and carry out an approximation of A with a reduced number of singular modes. To make this idea more precise, one can define the *relative information content* of the Singular Value Decomposition of A by:

$$\mathrm{RIC}(M) = \frac{\sum_{i=1}^{M} \sigma_i}{\sum_{i=1}^{N} \sigma_i} \tag{14}$$

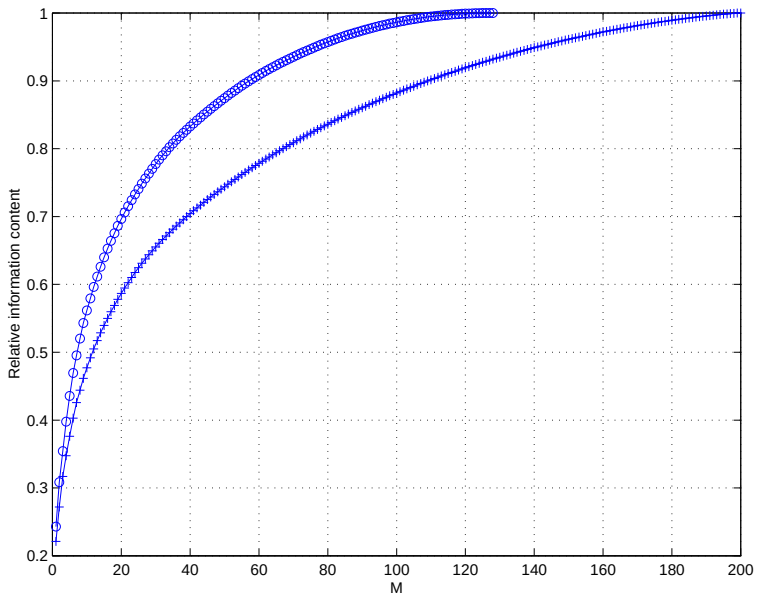


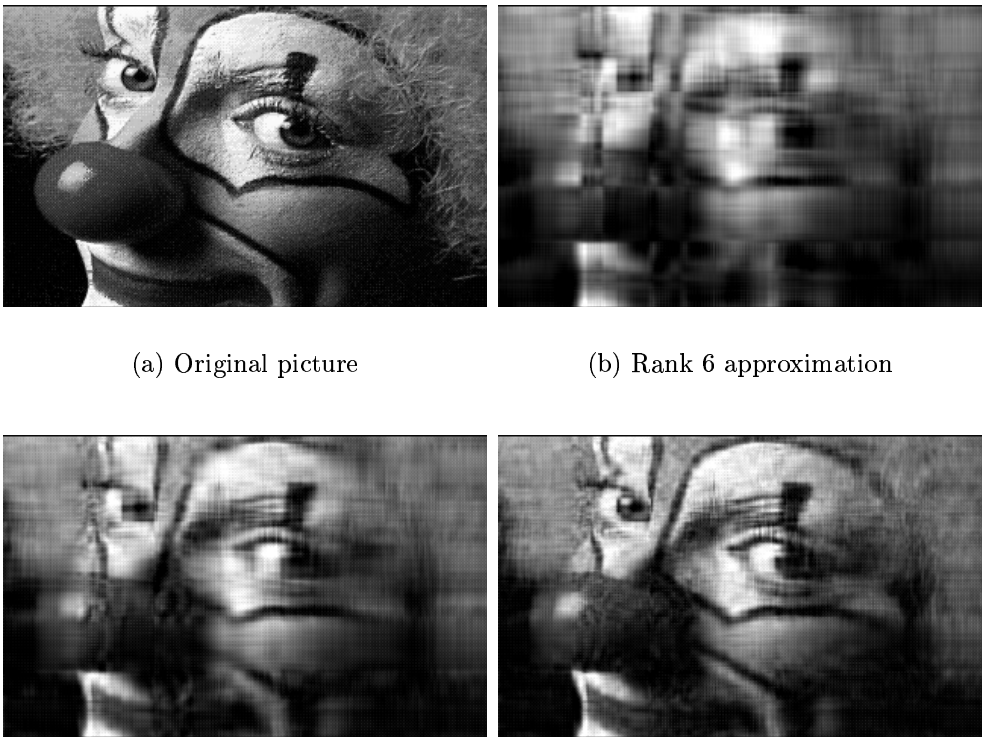
Figure 4: Relative information content for the "clown" image (+) and the "trees" image (∘).

If the low rank approximation is required to contain $\delta\%$ of the total information contained in the original image, the dimension M of the subspace

D_M^{SVD} spanned by the M first singular modes is determined by:

$$M = argmin\{\mathrm{RIC}(M); \mathrm{RIC}(M) \geq \delta\}. \tag{15}$$

In figure 4, the relative information content for the "clown" image and the "trees" image are shown. The same result as previously mentionned for the two images when the singular values spectrum was discussed is evidenced. For a given number of singular modes, say $M = 20$ for example, respectively 60 % and 70 % of the information content of the original "clown" image and "trees" image are contained in the approximation. This clearly demonstrates that the "trees" image is easier to approximate by a lower rank image than the "clown" image.



(a) Original picture

(b) Rank 6 approximation

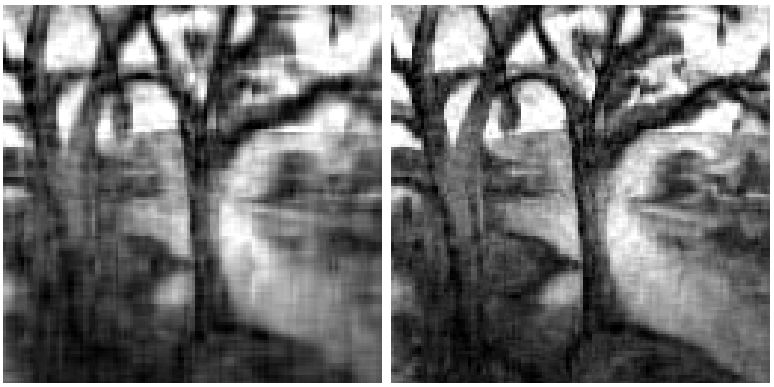
(c) Rank 12 approximation

(d) Rank 20 approximation

Figure 5: Approximations of the "clown" image from MatLab by images of lower rank.

Lastly, we present in figures 5 and 6, clockwise from top, the original picture, and approximants of rank 6, rank 20, and rank 12, for the "clown" image and "trees" image, respectively.

(a) Original picture (b) Rank 6 approximation



(c) Rank 12 approximation (d) Rank 20 approximation

Figure 6: Approximations of the "trees" image from MatLab by images of lower rank.

4 The Proper Orthogonal Decomposition (POD)

This section introduces the Proper Orthogonal Decomposition in the spirit of Holmes *et al.* (1996), as a technique which can contribute to a better understanding of turbulent flows. Here, POD is not reduced to an advanced processing method that allow extracting coherent structures from experimental or numerical data. Rather, POD is used to provide a set of basis functions with which can be identified a low-dimensional subspace on which to construct a dynamical model of the coherent structures by projection on the governing equations. This idea was first applied in Aubry *et al.* (1988) to model the near-wall region of a turbulent boundary layer and more recently by Ukeiley *et al.* (2001) to study the dynamics of the coherent structures in a plane turbulent mixing layer.

4.1 The Fredholm equation

Let $\{\vec{\boldsymbol{u}}(\vec{\boldsymbol{X}}),\ \vec{\boldsymbol{X}} = (\vec{\boldsymbol{x}}, t_n) \in \mathcal{D} = \mathbb{R}^3 \times \mathbb{R}^+\}$ denote the set of observations (also called *snapshots*) obtained at N_t different time steps t_n over a spatial domain of interest Ω ($\vec{\boldsymbol{x}} = (x, y, z) \in \Omega$). These snapshots could be experimental measurements or numerical solutions of velocity fields, vorticity fields, temperatures, etc. taken at different time steps and/or different physical parameters, for example Reynolds number (see Christensen *et al.*, 1998). The underlying problem is to extract from this ensemble of random vector field a coherent structure. Following Lumley (1967), a coherent structure is defined as *the deterministic function which is best correlated on average with the realizations* $\vec{\boldsymbol{u}}(\vec{\boldsymbol{X}})$. In other words, we look for a function $\vec{\boldsymbol{\Phi}}$ which has the largest mean square projection on the observations $|(\vec{\boldsymbol{u}}, \vec{\boldsymbol{\Psi}})|^2$. Since it is only the parallelism between $\vec{\boldsymbol{\Phi}}$ and the observations that is of interest, the dependence on the amplitude of $\vec{\boldsymbol{\Phi}}$ must be removed. One way is to normalize the amplitude of $\vec{\boldsymbol{\Phi}}$. It is then natural to look at a space of functions $\vec{\boldsymbol{\Phi}}$ for which the inner-product exists, i.e. to impose $\vec{\boldsymbol{\Phi}}$ to be an element of $L^2(\mathcal{D})$, the collection of square-integrable functions defined on flow region $\mathcal{D}$. Finally, in order to include the statistics, we must maximize the expression:

$$\frac{\langle |(\vec{\boldsymbol{u}}, \vec{\boldsymbol{\Phi}})|^2 \rangle}{\|\vec{\boldsymbol{\Phi}}\|^2}$$

in some average sense (temporal, spatial, ensemble or phase-average) denoted here by $\langle . \rangle$ and to be specified for each application. The choice of the average operator is at the heart of the different POD approaches and a detailed discussion of this point is postponed to §5.

Hence, mathematically, the function $\vec{\mathbf{\Phi}}$ corresponds to the solution of the constrained optimization problem:

$$\max_{\vec{\mathbf{\Psi}} \in L^2(\mathcal{D})} \frac{\langle |(\vec{\boldsymbol{u}}, \vec{\mathbf{\Psi}})|^2 \rangle}{\|\vec{\mathbf{\Psi}}\|^2} = \frac{\langle |(\vec{\boldsymbol{u}}, \vec{\mathbf{\Phi}})|^2 \rangle}{\|\vec{\mathbf{\Phi}}\|^2} \tag{16}$$

with respect to:

$$(\vec{\mathbf{\Phi}}, \vec{\mathbf{\Phi}}) = \|\vec{\mathbf{\Phi}}\|^2 = 1.$$

Here[5] $(.,.)$ and $\|.\|$ denote the usual L^2 inner product and L^2 norm over $\mathcal{D}$:

$$(\vec{\boldsymbol{u}}, \vec{\mathbf{\Phi}}) = \int_{\mathcal{D}} \vec{\boldsymbol{u}}(\vec{\boldsymbol{X}}) \cdot \vec{\mathbf{\Phi}}^*(\vec{\boldsymbol{X}}) d\vec{\boldsymbol{X}} = \sum_{i=1}^{n_c} \int_{\mathcal{D}} u_i(\vec{\boldsymbol{X}}) \Phi_i^*(\vec{\boldsymbol{X}}) d\vec{\boldsymbol{X}} \quad ; \quad \|\vec{\boldsymbol{u}}\|^2 = (\vec{\boldsymbol{u}}, \vec{\boldsymbol{u}})$$

where the $*$ superscript indicates conjugate complex and n_c is the number of vectorial component of $\vec{\boldsymbol{u}}(\vec{\boldsymbol{X}})$.

The maximization problem (16) can be cast in an equivalent eigenvalue problem. To see this, let us define the operator $\mathcal{R} : L^2(\mathcal{D}) \longrightarrow L^2(\mathcal{D})$ by:

$$\mathcal{R}\vec{\mathbf{\Phi}}(\vec{\boldsymbol{X}}) = \int_{\mathcal{D}} R(\vec{\boldsymbol{X}}, \vec{\boldsymbol{X}'}) \vec{\mathbf{\Phi}}(\vec{\boldsymbol{X}'}) d\vec{\boldsymbol{X}'}$$

where $R(\vec{\boldsymbol{X}}, \vec{\boldsymbol{X}'}) = \langle \vec{\boldsymbol{u}}(\vec{\boldsymbol{X}}) \otimes \vec{\boldsymbol{u}}^*(\vec{\boldsymbol{X}'}) \rangle$ is the two-point space-time correlation tensor ($\otimes$ is the dyadic product).

Then, straightforward calculations[6] reveal that:

$$\begin{aligned}
(\mathcal{R}\vec{\mathbf{\Phi}}, \vec{\mathbf{\Phi}}) &= \Big(\int_{\mathcal{D}} \langle \vec{\boldsymbol{u}}(\vec{\boldsymbol{X}}) \otimes \vec{\boldsymbol{u}}^*(\vec{\boldsymbol{X}'}) \rangle \vec{\mathbf{\Phi}}(\vec{\boldsymbol{X}'}) d\vec{\boldsymbol{X}'}, \vec{\mathbf{\Phi}}(\vec{\boldsymbol{X}}) \Big) \\
&= \int_{\mathcal{D}} \int_{\mathcal{D}} \langle \vec{\boldsymbol{u}}(\vec{\boldsymbol{X}}) \otimes \vec{\boldsymbol{u}}^*(\vec{\boldsymbol{X}'}) \rangle \vec{\mathbf{\Phi}}(\vec{\boldsymbol{X}'}) d\vec{\boldsymbol{X}'} \cdot \vec{\mathbf{\Phi}}^*(\vec{\boldsymbol{X}}) d\vec{\boldsymbol{X}} \\
&= \langle \int_{\mathcal{D}} \vec{\boldsymbol{u}}(\vec{\boldsymbol{X}}) \cdot \vec{\mathbf{\Phi}}^*(\vec{\boldsymbol{X}}) d\vec{\boldsymbol{X}} \int_{\mathcal{D}} \vec{\boldsymbol{u}}^*(\vec{\boldsymbol{X}'}) \cdot \vec{\mathbf{\Phi}}(\vec{\boldsymbol{X}'}) d\vec{\boldsymbol{X}'} \rangle \\
&= \langle |(\vec{\boldsymbol{u}}, \vec{\mathbf{\Phi}})|^2 \rangle \geq 0
\end{aligned}$$

[5] L^2 seems to be a natural space in which to do Fluid Mechanics since it corresponds to flow having finite kinetic energy, but the choice of other norms for the POD basis computation is possible, see § 5.2 for a discussion.

[6] We suppose that the probabilistic structure of the ensemble of observations is such that the average and integrating operations can be interchanged (see Berkooz, 1991).

Furthermore, it follows that:

$$(\mathcal{R}\vec{\mathbf{\Phi}}, \vec{\mathbf{\Psi}}) = (\vec{\mathbf{\Phi}}, \mathcal{R}\vec{\mathbf{\Psi}}) \qquad \text{for any} \quad (\vec{\mathbf{\Phi}}, \vec{\mathbf{\Psi}}) \in [L^2(\mathcal{D})]^2.$$

Then $\mathcal{R}$ is a linear, self-adjoint[7], non-negative[8] operator on $L^2(\mathcal{D})$. Consequently, spectral theory applies (see Riesz and Nagy, 1955; Courant and Hilbert, 1953) and guarantees that the maximization problem (16) admits a solution, equal to the largest eigenvalue of the problem:

$$\mathcal{R}\vec{\mathbf{\Phi}} = \lambda\vec{\mathbf{\Phi}} \tag{17}$$

which can be written as a Fredholm integral eigenvalue problem:

$$\boxed{\sum_{j=1}^{n_c} \int_{\mathcal{D}} R_{ij}(\vec{\boldsymbol{X}}, \vec{\boldsymbol{X}'})\Phi_j(\vec{\boldsymbol{X}'})d\vec{\boldsymbol{X}'} = \lambda\Phi_i(\vec{\boldsymbol{X}}).} \tag{18}$$

The properties of the empirical eigenfunctions $\Phi_i(\vec{\boldsymbol{X}})$ obtained by solving the Fredholm equation (18) are fully discussed in §4.2. Here, it is sufficient to make some comments shedding light on the constraints linked to the POD method.

In equation (18), the integral $\int_{\mathcal{D}} .\, d\vec{\boldsymbol{X}'}$ is over the entire domain $\mathcal{D}$ of interest. The consequence is that the two-point correlation tensor R_{ij} has to be known over all $\mathcal{D}$. Therefore, the data volume to handle can be very important (several gigabytes are not rare) and sometimes a data compression is necessary to reduce the data storage requirements (see Cordier and Bergmann, 2002, for an example). Due to the important size of the data sets necessary to apply POD, renewed interest for POD appears only in the 1990's explained by the great advances in numerical simulation capability and in measurement techniques.

Remark: An alternative approach for finding the solution to maximization of (16) is by directly solving a classical problem in the calculus of variations. Since $(\mathcal{R}\vec{\mathbf{\Phi}}, \vec{\mathbf{\Phi}}) = \langle|(\vec{\boldsymbol{u}}, \vec{\mathbf{\Phi}})|^2\rangle$, the problem (16) is equivalent to determine $\vec{\mathbf{\Phi}}$ that maximizes λ where:

$$\lambda = \frac{\langle|(\vec{\boldsymbol{u}}, \vec{\mathbf{\Phi}})|^2\rangle}{(\vec{\mathbf{\Phi}}, \vec{\mathbf{\Phi}})} = \frac{(\mathcal{R}\vec{\mathbf{\Phi}}, \vec{\mathbf{\Phi}})}{(\vec{\mathbf{\Phi}}, \vec{\mathbf{\Phi}})} \tag{19}$$

[7] i.e. $\mathcal{R}^\dagger = \mathcal{R}$ where the adjoint of $\mathcal{R}$, $\mathcal{R}^\dagger$ is defined by:

$$(\mathcal{R}\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}) = \left(\vec{\boldsymbol{u}}, \mathcal{R}^\dagger\vec{\boldsymbol{v}}\right) \qquad \text{for all } \vec{\boldsymbol{u}} \in L^2(\mathcal{D}) \text{ and } \vec{\boldsymbol{v}} \in L^2(\mathcal{D}).$$

[8] i.e. $(\mathcal{R}\vec{\boldsymbol{u}}, \vec{\boldsymbol{u}}) \geq 0 \qquad \text{for all } \vec{\boldsymbol{u}} \in L^2(\mathcal{D})$.

Using the calculus of variations, $\vec{\mathbf{\Phi}}$ is determining by imposing the condition $\frac{dF(\epsilon)}{d\epsilon}\Big|_{\epsilon=0} = 0$ with:

$$F(\epsilon) = \frac{\left(\mathcal{R}(\vec{\mathbf{\Phi}} + \epsilon\vec{\mathbf{\Upsilon}}), (\vec{\mathbf{\Phi}} + \epsilon\vec{\mathbf{\Upsilon}})\right)}{\left((\vec{\mathbf{\Phi}} + \epsilon\vec{\mathbf{\Upsilon}}), (\vec{\mathbf{\Phi}} + \epsilon\vec{\mathbf{\Upsilon}})\right)} = \frac{(\mathcal{R}\vec{\mathbf{\Phi}}, \vec{\mathbf{\Phi}}) + \epsilon(\mathcal{R}\vec{\mathbf{\Phi}}, \vec{\mathbf{\Upsilon}}) + \epsilon(\mathcal{R}\vec{\mathbf{\Upsilon}}, \vec{\mathbf{\Phi}}) + \epsilon^2(\mathcal{R}\vec{\mathbf{\Upsilon}}, \vec{\mathbf{\Upsilon}})}{(\vec{\mathbf{\Phi}}, \vec{\mathbf{\Phi}}) + \epsilon(\vec{\mathbf{\Phi}}, \vec{\mathbf{\Upsilon}}) + \epsilon(\vec{\mathbf{\Upsilon}}, \vec{\mathbf{\Phi}}) + \epsilon^2(\vec{\mathbf{\Upsilon}}, \vec{\mathbf{\Upsilon}})}$$

This leads one to verify for any $\vec{\mathbf{\Upsilon}}$ the condition:

$$\left(\mathcal{R}\vec{\mathbf{\Phi}}, \vec{\mathbf{\Upsilon}}\right) = \lambda \left(\vec{\mathbf{\Phi}}, \vec{\mathbf{\Upsilon}}\right)$$

which is equivalent as finding the eigenvalue of the eigenvalue problem (17).

4.2 Properties of the POD basis functions

1. For bounded integration domain $\mathcal{D}$, Hilbert-Schmidt theory applies (Riesz and Nagy, 1955) and assures us that there is not one, but a denumerable infinity of solutions of (18). Then, the Fredholm equation (18) has a discrete set of solutions satisfying:

$$\sum_{j=1}^{n_c} \int_{\mathcal{D}} R_{ij}(\vec{\boldsymbol{X}}, \vec{\boldsymbol{X}}')\, \Phi_j^{(n)}(\vec{\boldsymbol{X}}')\, d\vec{\boldsymbol{X}}' = \lambda^{(n)}\, \Phi_i^{(n)}(\vec{\boldsymbol{X}}) \tag{20}$$

   where $\lambda^{(n)}$ and $\Phi_i^{(n)}$ denote respectively the POD eigenvalues and POD eigenvectors or eigenfunctions of order $n = 1,\ 2,\ 3, \cdots, +\infty$. Each new eigenfunction is sought as the solution problem of the maximization problem (16) subject to the constraint of being orthogonal to all previously found eigenfunctions. Hence, by construction, the eigenfunctions are mutually orthogonal but they can be choosen orthonormal (see property 4). Any d-fold degenerate eigenvalue is associated with d linearly independent eigenfunctions.

2. $\mathcal{R}$ is a self-adjoint and non negative operator then all eigenvalues are real and positive:

$$\lambda^{(1)} \geq \lambda^{(2)} \geq \lambda^{(3)} \geq, \cdots \lambda^{(+\infty)} \geq 0 \tag{21}$$

   and the corresponding series converges:

$$\sum_{n=1}^{+\infty} \lambda^{(n)} < +\infty$$

3. The eigenfunctions $\vec{\mathbf{\Phi}}^{(n)}$ form a complete orthogonal set, which means that almost every member (except possibly on a set of measure zero, see Berkooz *et al.*, 1993) of the snapshots can be reconstructed in the following way:

$$u_i(\vec{\boldsymbol{X}}) = \sum_{n=1}^{+\infty} a^{(n)} \Phi_i^{(n)}(\vec{\boldsymbol{X}}) \tag{22}$$

4. Eigenfunctions $\vec{\mathbf{\Phi}}^{(n)}$ can be chosen mutually orthonormal[9]:

$$\sum_{i=1}^{n_c} \int_{\mathcal{D}} \Phi_i^{(m)}(\vec{\boldsymbol{X}})\, \Phi_i^{*(n)}(\vec{\boldsymbol{X}})\, d\vec{\boldsymbol{X}} = \delta_{mn} = \begin{cases} 0 & \text{for } m \neq n; \\ 1 & \text{for m=n.} \end{cases} \tag{24}$$

5. The random coefficients $a^{(n)}$, projections of $\vec{\boldsymbol{u}}$ onto $\vec{\mathbf{\Phi}}$, are then calculated by using the orthonormality of the eigenfunctions $\vec{\mathbf{\Phi}}$.

$$a^{(n)} = (\vec{\boldsymbol{u}}, \vec{\mathbf{\Phi}}) = \sum_{i=1}^{n_c} \int_{\mathcal{D}} u_i(\vec{\boldsymbol{X}})\, \Phi_i^{*(n)}(\vec{\boldsymbol{X}})\, d\vec{\boldsymbol{X}}. \tag{25}$$

6. The two-point space-time correlation tensor R_{ij} can be decomposed as a uniformly convergent series (see Courant and Hilbert, 1953):

$$R_{ij}(\vec{\boldsymbol{X}}, \vec{\boldsymbol{X}}') = \sum_{n=1}^{+\infty} \lambda^{(n)} \Phi_i^{(n)}(\vec{\boldsymbol{X}}) \Phi_j^{*(n)}(\vec{\boldsymbol{X}}') \tag{26}$$

   This result is known as the Mercer's theorem.

7. The diagonal representation of the tensor R_{ij} combined with the decomposition of $\vec{\boldsymbol{u}}$ on the eigenfunctions $\vec{\mathbf{\Phi}}$ and their orthogonality assure

[9] Since $\mathcal{R}$ is a self-adjoint operator the orthogonality is verified necessarily. On the other hand, the choice of orthonormality for the eigenfunctions are rather arbitrary because they are determined relative to a real multiplicative constant. Hence, it is numerically equivalent to impose:

$$\sum_{i=1}^{n_c} \int_{\mathcal{D}} \Phi_i^{(m)}(\vec{\boldsymbol{X}})\, \Phi_i^{*(n)}(\vec{\boldsymbol{X}})\, d\vec{\boldsymbol{X}} = \lambda^{(m)} \delta_{mn} \tag{23}$$

for the eigenfunctions $\Phi_i^{(m)}(\vec{\boldsymbol{X}})$ and the condition $\langle a^{(n)} a^{*(m)} \rangle = \delta_{mn}$ for the projection coefficient $a^{(n)}$ or to impose for the eigenfunctions the orthonormality condition (24) and the orthogonality condition (27) for the coefficients. For numerical reasons, it is easier to use condition (24) for the "classical POD", and condition (24) for the "snapshot POD" (e.g. Rempfer and Fasel, 1994).

that the coefficients $a^{(n)}$ are mutually uncorrelated and that their mean square values are the eigenvalues themselves.

$$\langle a^{(n)} a^{*(m)} \rangle = \delta_{mn} \lambda^{(n)}. \tag{27}$$

Proof: This assertion derives directly from the representation of $R_{ij}(\vec{\boldsymbol{X}}, \vec{\boldsymbol{X}}')$, given in equation (26):

$$\begin{aligned} R_{ij}(\vec{\boldsymbol{X}}, \vec{\boldsymbol{X}}') = \langle u_i(\vec{\boldsymbol{X}})\, u_j^*(\vec{\boldsymbol{X}}') \rangle &= \left\langle \sum_{n=1}^{+\infty} a^{(n)} \Phi_i^{(n)}(\vec{\boldsymbol{X}}) \sum_{m=1}^{+\infty} a^{*(m)} \Phi_j^{*(m)}(\vec{\boldsymbol{X}}') \right\rangle \\ &= \sum_{n=1}^{+\infty} \sum_{m=1}^{+\infty} \left\langle a^{(n)} a^{*(m)} \right\rangle \Phi_i^{(n)}(\vec{\boldsymbol{X}})\, \Phi_j^{*(m)}(\vec{\boldsymbol{X}}') \end{aligned}$$

But we know from the Mercer's theorem that:

$$R_{ij}(\vec{\boldsymbol{X}}, \vec{\boldsymbol{X}}') = \sum_{n=1}^{+\infty} \lambda^{(n)} \Phi_i^{(n)}(\vec{\boldsymbol{X}}).\Phi_j^{*(n)}(\vec{\boldsymbol{X}}'),$$

and so, since the $\vec{\boldsymbol{\Phi}}^{(n)}(\vec{\boldsymbol{X}})$ are an orthonormal family in $L^2(\mathcal{D})$, we see that $\langle a^{(n)} a^{*(m)} \rangle = \delta_{mn} \lambda^{(n)}$.

□

8. Finally, the Mercer's theorem and the orthonormality of $\vec{\boldsymbol{\Phi}}^{(n)}$ lead to:

$$\sum_{i=1}^{n_c} \int_{\mathcal{D}} R_{ii}(\vec{\boldsymbol{X}}, \vec{\boldsymbol{X}})\, d\vec{\boldsymbol{X}} = \sum_{n=1}^{+\infty} \lambda^{(n)} = E \tag{28}$$

If $\vec{\boldsymbol{u}}(\vec{\boldsymbol{X}})$ is a velocity field[10], then E corresponds to the Turbulent Kinetic Energy (TKE) integrated over the domain $\mathcal{D}$. The interpretation of this equation is that every structure of order (n) makes an independent contribution to the TKE. Then, the amplitude of the eigenvalues $\lambda^{(n)}$ measure the relative importance of the different structures present within the flow.

[10] In the same way, if $\vec{\boldsymbol{u}}(\vec{\boldsymbol{X}})$ is a vorticity field as in Sanghi (1991), this relation leads to the system enstrophy. So, whatever variable is considered for the POD, the eigenvalues $\lambda^{(n)}$ obtained by solving the Fredholm equation (18) are always homogeneous to energy but are not strictly speaking energy. Thinking of the POD eigenvalues as energy in a general mechanical context is incorrect in principle and may lead to misleading results.

4.3 Optimality of the POD basis

Suppose that we have a signal $\vec{\boldsymbol{u}}(\vec{\boldsymbol{X}})$ with $\vec{\boldsymbol{u}} \in L^2(\mathcal{D})$ and an approximation $\vec{\boldsymbol{u}}^a$ of $\vec{\boldsymbol{u}}$ with respect to an arbitrary orthonormal basis $\vec{\boldsymbol{\Psi}}^{(n)}(\vec{\boldsymbol{X}})$, $n = 1, 2, \cdots, +\infty$, one can write:

$$u_i^a(\vec{\boldsymbol{X}}) = \sum_{n=1}^{+\infty} b^{(n)} \Psi_i^{(n)}(\vec{\boldsymbol{X}}).$$

The equations (28) and (27) clearly stated that if the $\Psi_i^{(n)}(\vec{\boldsymbol{X}})$ have been nondimensionalized then the expression $\langle b^{(n)} b^{*(n)} \rangle$ represents the average kinetic energy in the n^{th} mode. The following lemma establishes the notion of optimality of the POD approach.

Lemma Let $\{\vec{\boldsymbol{\Phi}}^{(1)}(\vec{\boldsymbol{X}}), \vec{\boldsymbol{\Phi}}^{(2)}(\vec{\boldsymbol{X}}), \cdots, \vec{\boldsymbol{\Phi}}^{(\infty)}(\vec{\boldsymbol{X}})\}$ denote an orthonormal set of POD basis elements, and $\{\lambda_1, \lambda_2, \cdots, \lambda_\infty\}$ denote the corresponding set of eigenvalues. If:

$$u_i^P(\vec{\boldsymbol{X}}) = \sum_{n=1}^{+\infty} a^{(n)} \Phi_i^{(n)}(\vec{\boldsymbol{X}})$$

denotes the approximation to $\vec{\boldsymbol{u}}$ with respect to this basis, then for any value of N (see Holmes *et al.*, 1996):

$$\sum_{n=1}^{N} \langle a^{(n)} a^{*(n)} \rangle = \sum_{n=1}^{N} \lambda^{(n)} \geq \sum_{n=1}^{N} \langle b^{(n)} b^{*(n)} \rangle$$

Proof: It is straightforward (see the proof of equation (27) in §4.2) that the kernel R_{ij} can be expressed in terms of $\vec{\boldsymbol{\Psi}}^{(n)}$, $n = 1, \cdots, +\infty$ as:

$$R_{ij}(\vec{\boldsymbol{X}}, \vec{\boldsymbol{X}}') = \sum_{n=1}^{+\infty} \sum_{m=1}^{+\infty} \left\langle b^{(n)} b^{*(m)} \right\rangle \Psi_i^{(n)}(\vec{\boldsymbol{X}}) \, \Psi_j^{*(m)}(\vec{\boldsymbol{X}}')$$

Therefore, the projection of the kernel R_{ij} in an N-dimensional space spanned by $\{\vec{\boldsymbol{\Psi}}^{(n)}\}_{n=1}^{N}$ can be written in a matrix form as:

$$R = \begin{bmatrix} \langle b^{(1)} b^{*(1)} \rangle & \langle b^{(1)} b^{*(2)} \rangle & \cdots & \langle b^{(1)} b^{*(N)} \rangle & 0 & \cdots & 0 \\ \langle b^{(2)} b^{*(1)} \rangle & \langle b^{(2)} b^{*(2)} \rangle & \cdots & \langle b^{(2)} b^{*(N)} \rangle & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \langle b^{(N)} b^{*(1)} \rangle & \langle b^{(N)} b^{*(2)} \rangle & \cdots & \langle b^{(N)} b^{*(N)} \rangle & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

The proof finally relies on a result on linear operators (see Temam, 1988, p. 260) which states that the sum of the first N eigenvalues of a self-adjoint operator is greater than or equal to the sum of the diagonal terms in any N-dimensional projection of it:

$$\sum_{n=1}^{N} \lambda^{(n)} \geq \mathrm{Tr}(R) = \sum_{n=1}^{N} \langle b^{(n)} b^{*(m)} \rangle$$

□

This lemma establishes that among all linear decompositions[11], the POD is the most efficient, in the sense that, for a given number of modes, N, the projection on the subspace spanned by the N leading eigenfunctions will contain the greatest possible kinetic energy on average.

4.4 Model reduction aspects

The energetic optimality of the POD basis functions suggests that only a very small number of POD modes, say M, may be necessary to describe efficiently any signal $\vec{u}(\vec{X})$ of the input data. The choice of M is then an important and critical task and adequate criteria for choosing M must be introduced.

Let N_{POD} denote the number of POD modes obtained by solving the Fredholm equation (18). The truncation error $\epsilon(M)$, of using M instead of N_{POD} POD basis functions in representing the input data is given by[12]:

$$\begin{aligned} \epsilon(M) =& \|\vec{u}(\vec{X}) - \sum_{n=1}^{M} \left(\vec{u}(\vec{X}), \vec{\Phi}^{(n)}(\vec{X})\right) \vec{\Phi}^{(n)}(\vec{X})\|^2 \\ =& \| \sum_{n=M+1}^{N_{POD}} \left(\vec{u}(\vec{X}), \vec{\Phi}^{(n)}(\vec{X})\right) \vec{\Phi}^{(n)}(\vec{X})\|^2 \end{aligned} \tag{29}$$

The quantity $\epsilon(M)$ measures the accumulated squarred error in representing the input snapshots, due to neglecting POD basis elements that correspond to small POD eigenvalues.

However, in practice, this criterion is never used and the choice of M is rather based on heuristic considerations. As we indicated in point 8 of §4.2, $\sum_{i=1}^{M} \lambda^{(i)}$ represents in some sense (see footnote 10) the average energy[13]

[11] The reader must remember that optimality of the POD functions is obtained only with respect to other linear representations.

[12] It's immediate to deduce from equation (29), equivalent forms for the two particular approaches of POD described in §5. The reader is referred to Fahl (2000) where ϵ is defined for the snapshot POD.

[13] For turbulent flows, it corresponds exactly to the average Turbulent Kinetic Energy.

contained in the first M POD modes. Therefore, to capture most of the energy contained in the N_{POD} POD modes, it suffices to choose M so that $\sum_{i=1}^{M} \lambda^{(i)} \simeq \sum_{i=1}^{N_{POD}} \lambda^{(i)}$. By definition, the ratio $\sum_{i=1}^{M} \lambda^{(i)} / \sum_{i=1}^{N_{POD}} \lambda^{(i)}$ yields the percentage of the total kinetic energy in the N_{POD} POD modes that is contained in the first M POD basis functions. For a predefined percentage of energy δ , the dimension M of the subspace spanned by the M first POD functions is chosen such that the condition

$$\frac{\sum_{i=1}^{M} \lambda^{(i)}}{\sum_{i=1}^{N_{POD}} \lambda^{(i)}} \geq \delta \tag{30}$$

holds (see Cordier and Bergmann, 2002; Ravindran, 2000b; Fahl, 2000). The criterion (30) is equivalent to the one based on the relative information content used in §3.6 for the Singular Value Decomposition (see equation 15). The POD reduced basis subspace is defined as $D_M^{POD} = \text{span}\{\vec{\mathbf{\Phi}}^{(1)}, \vec{\mathbf{\Phi}}^{(2)}, \cdots, \vec{\mathbf{\Phi}}^{(M)}\}$.

To this point we have only discussed of the model reduction associated with using POD basis functions in approximation of the input collection. Dynamical models based on POD were not discussed. Nevertheless, the optimal energetic convergence of the POD basis functions suggests that only a very small number of modes may be necessary to describe the dynamics of the system. Therefore, starting from data issued from high-dimensional models (experimental data or detailed simulations), it seems conceivable that POD modes can be efficiently used in Galerkin projections that yield low-dimensional dynamical models. Even though there are no theoretical guarantees of optimality in dynamical modelling, this method was already used in many cases, for turbulent flows[14] or optimal control of fluids [15] and reasonable to excellent models were obtained. The presentation of this approach is

[14]For turbulent flows, POD is used to weave low-dimensional models that address the role of coherent structures in turbulence generation (see Aubry *et al.*, 1988; Ukeiley *et al.*, 2001).

[15]For optimal control of fluids, POD is used for obtaining reduced order models of dynamics that reduces computational complexity associated with high complexity models like the Navier-Stokes equations (see Ravindran, 2000a,b; Hinze, 2000; Fahl, 2000). In the control literature (see Atwell and King, 1999), several philosophies exist for using a reduced basis obtained by applying POD in low order control design. A "reduce-then-design" approach involves reduction of the system model before control design, and the "design-then-reduce" approach, in which full order model design is followed by full order control design, and then by control order reduction.

postponed to Cordier and Bergmann (2002).

As a partial conclusion, note that the reduced order models based on POD belong to a wider class of approximation methods called Singular Value based methods by Antoulas and Sorensen (2001). These authors recently review the state of affairs in the area of model reduction of dynamical systems and distinguish three broad categories of approximation methods (see figure 7): Singular Value based methods, Krylov based methods and iterative methods combining aspects of both the SVD and Krylov methods. Since, the strengths and weaknesses of these methods are different, new insights can certainly be gained by applying these approximation methods for fluid flow control. For example, the reader is referred to Allan (2000) for an application of the Krylov subspace method to derive an optimal feedback control design for the driven cavity flow.

SVD based methods		**Krylov based methods**
Nonlinear systems	Linear systems	
POD methods Empirical grammians	Balanced truncation Hankel approximation	Lanczos Arnoldi Interpolation
SVD-Krylov based methods		

Figure 7: Overview of approximation methods for dynamical systems after Antoulas and Sorensen (2001).

5 The different POD approaches

Except for the inner product, defined as the standard L^2 inner product for simplicity of presentation, the POD was derived in §4 in a general setting. The fundamental questions of the choice of:

- ▷ the input collection,
- ▷ the inner product,
- ▷ the averaging operation $\langle . \rangle$ (spatial or temporal),
- ▷ the variable $\vec{\boldsymbol{X}}$ (spatial $\vec{\boldsymbol{x}} = (x, y, z)$ or temporal t),

were not discussed. This section demonstrates that different orthogonal decomposition can be obtained depending, for example, on the way the averaging operator $\langle . \rangle$ is defined for calculating the kernel of the Fredholm equation

(18). In what follows, only two methods: the classical POD (§5.3) and the snapshot POD (§5.4) will be fully described. The reader is referred to Aubry *et al.* (1991) for a presentation of the generalization of these two methods called: *the Biorthogonal Decomposition.*

5.1 Choice of input collection

Choosing an input collection is a vital part of the Proper Orthogonal Decomposition process since the POD basis only reflects information provided by the input collection. The POD eigenfunctions are intrinsically linked to the input data used to extract them. This is the source of the method's strengths as well as its limitations: extrapolation of the POD functions to different geometry or control parameters (Reynolds number, ...), can be difficult to undertake (see Delville *et al.*, 1998, §4.6, p. 254 for a discussion).

When the POD basis is used for model reduction (§4.4), an input collection of time snapshots is frequently chosen (see §5.4). Typically, this data sets comes from experimental measurement or numerical computations. Hence the data have some error associated with them. Therefore it is important to study the effect of these errors, assimilated to infinitesimal perturbations, on the outcome of the POD model reduction procedure. This fundamental question have been only recently investigated theoretically by Rathinam and Petzold (2001). These authors introduced the POD sensitivity factor as a non dimensional measure of the sensitivity of the resulting projection with respect to perturbations in data. They found that the POD sensitivity is relevant in some applications of POD while it is not in some other applications. These theoretical results still need to be illustrated by realistic examples issued from Fluid Mechanics for example. Now, consider the ideal case with no perturbations in the input collection. Choosing a time snapshot input collection relevant for dynamical system description remains a difficult task because there is no definitive way to decide: how many snapshots are necessary to capture the information content of the system, how long numerical simulations or experiments should be run to generate sufficiently resolved snapshots, which initial conditions should be used[16].

For control problems based on reduced order models, an open question is how to incorporate control information. A simple solution, generally used,

[16]The reader should remember that the input collection corresponds to solutions belonging to the attractor of a dynamical system such as the Navier-Stokes equation in Fluid Mechanics. If this attractor is ergodic, the initial conditions are forgotten as time proceeds (see Holmes *et al.*, 1996).

is to generate snapshots using a variety of control inputs to excite system dynamics that arise when a control is applied (see Graham *et al.*, 1997).

5.2 Choice of inner product and norm

So far, POD was described in the context of the standard L^2 inner product for reason of simplicity and more importantly because it corresponds to the general case for fluid flow applications for reasons explained in §5.2.1. However, in few cases, it could be useful to use different inner product, to obtain different notions of optimality (see § 5.2.2 and § 5.2.3).

5.2.1 L^2 inner product

Let $L^2(\Omega)$ be the Hilbert space of square integrable[17] complex-valued functions defined on Ω. For vector-valued functions $\vec{\boldsymbol{u}}$, such as the velocity field in a fluid flow, the inner product on $L^2(\Omega)$ is defined by:

$$(\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}) = \int_\Omega (u_1 v_1^* + u_2 v_2^* + u_3 v_3^*)\ d\vec{\boldsymbol{x}} \quad ; \quad \|\vec{\boldsymbol{u}}\|^2 = (\vec{\boldsymbol{u}}, \vec{\boldsymbol{u}})\,, \tag{31}$$

where Ω denotes the spatial domain occupied by the fluid. Moreover, his kinetic energy is proportionnal to $\|\vec{\boldsymbol{u}}\|^2$. Therefore, L^2 is a natural space in which to do Fluid Mechanics since it corresponds to flow having finite kinetic energy. This is the reason which explains that the L^2 inner product is the most used for defining the Proper Orthogonal Decomposition.

5.2.2 H^1 inner product

Let $H^1(\Omega)$ be the Sobolev space of functions that, along with their first derivatives belong to $L^2(\Omega)$.

In Iollo (1997), it is found that the low order model developed for the Euler equations by a straightforward Galerkin projection (see Cordier and Bergmann, 2002, for a description of the method) was unstable. Therefore, Iollo *et al.* (1998) proposed a way to improve the numerical stability of the low-order model developped by Galerkin POD by redefining the norms involved in the POD definition as:

$$(u, v)_\epsilon = \int_\Omega u\, v\, d\vec{\boldsymbol{x}} + \epsilon \int_\Omega \left(\vec{\boldsymbol{\nabla}} u \cdot \vec{\boldsymbol{\nabla}} v\right)\ d\vec{\boldsymbol{x}} \tag{32}$$

[17] Square integrable means that the functions $f(x)$ belonging to $L^2(\Omega)$ satisfy:

$$\|f\| = (f, f)^{1/2} = \left[\int_\Omega |f|^2\, dx\right]^{1/2} < +\infty.$$

where ϵ is a parameter to take into account different metrics. Numerical experiments conclude to definite benefit in employing the H^1 formulation of the POD. Even though the use of the H^1 inner product seems beneficial for the robustness of the reduced order model, we believe that it has not been given sufficient attention in the literature.

5.2.3 Inner product for compressible flow

For compressible flow configurations, the velocity variables $\vec{\boldsymbol{u}} = (u, v, w)$ and thermodynamic variables (e.g. density ρ, pressure p, enthalpy h) are dynamically coupled. This introduces questions of whether to treat the thermodynamic variables separately from the velocity, or together as a single vector-valued variable (e.g. $\vec{\boldsymbol{q}}(\vec{\boldsymbol{x}}) = (\rho, u, v, w, p)(\vec{\boldsymbol{x}})$). For a scalar-valued POD, where separate POD modes is computed for each flow variable, the standard L^2 inner product defined by (31) can be used (see Rowley, 2002). For vector-valued POD, where all the flow variables are written as a single vector $\vec{\boldsymbol{q}}$, the standard-inner product:

$$(\vec{\boldsymbol{q}}_{\mathbf{1}}, \vec{\boldsymbol{q}}_{\mathbf{2}}) = \int_\Omega (\rho_1\rho_2 + u_1u_2 + v_1v_2 + w_1w_2 + p_1p_2)\ d\,\vec{\boldsymbol{x}}$$

may not be a sensible choice for dimensional reason. Of course, one could simply nondimensionalize the variables, but then the sense in which projections are optimal is rather arbitrary and depends on the nondimensionalization. Rowley (2002) sought an inner product for compressible flow which makes intuitive sense, in that the "energy" defined by the induced norm is meaningful physical quantity. For a two-dimensional configuration, Rowley introduced a vector-valued variable $\vec{\boldsymbol{q}} = (u, v, a)$ where u and v are the velocities and a is the local sound speed, and defined a family of inner products as:

$$(\vec{\boldsymbol{q}}_{\mathbf{1}}, \vec{\boldsymbol{q}}_{\mathbf{2}})_\epsilon = \int_\Omega \left(u_1u_2 + v_1v_2 + \frac{2\epsilon}{\gamma(\gamma-1)} a_1a_2 \right)\ d\,\vec{\boldsymbol{x}} \tag{33}$$

where γ is the ratio of specific heats and ϵ is a parameter. If $\epsilon = \gamma$ then the induced norm gives $\|\vec{\boldsymbol{q}}\|^2 = 2h_0$ i.e. twice the total enthalpy of the flow, and if $\epsilon = 1$ then the induced norm gives twice the total energy of the flow.

5.3 Classical POD or direct method

This approach is the one originated by Lumley (1967). In this case, the average $\langle . \rangle$ is temporal:

$$\langle . \rangle = \frac{1}{T} \int_T . \, dt$$

and is evaluated as an ensemble average, based on the assumptions of stationarity and ergodicity. In the other hand, the variable $\vec{\boldsymbol{X}}$ is assimilated to the space variable $\vec{\boldsymbol{x}} = (x, y, z)$ defined over the domain Ω.

Figure 8 describes schematically the principle of the classical POD.

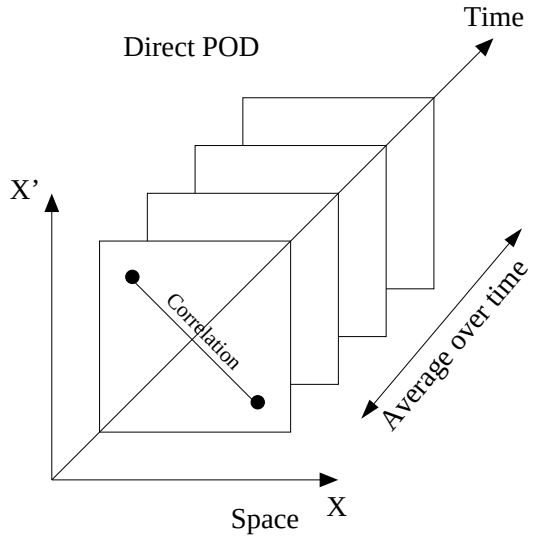


Figure 8: Schematic view of the classical POD.

The corresponding eigenvalue problem is deduced easily from equation (20) by replacing the domain of integration $\mathcal{D}$ by Ω and the variable $\vec{\boldsymbol{X}}$ by $\vec{\boldsymbol{x}}$. The integral Fredholm equation to be solved is then given by:

$$\boxed{\sum_{j=1}^{n_c} \int_{\Omega} R_{ij}(\vec{\boldsymbol{x}}, \vec{\boldsymbol{x}}')\, \Phi_j^{(n)}(\vec{\boldsymbol{x}}')\, d\vec{\boldsymbol{x}}' = \lambda^{(n)}\, \Phi_i^{(n)}(\vec{\boldsymbol{x}})} \tag{34}$$

where $R_{ij}(\vec{\boldsymbol{x}}, \vec{\boldsymbol{x}'})$ is the **two-point spatial correlation tensor** defined as:

$$R_{ij}(\vec{\boldsymbol{x}}, \vec{\boldsymbol{x}'}) = \frac{1}{T} \int_T u_i(\vec{\boldsymbol{x}}, t) u_j(\vec{\boldsymbol{x}'}, t)\; d\,t = \sum_{n=1}^{N_{POD}} \lambda^{(n)} \Phi_i^{(n)}(\vec{\boldsymbol{x}}) \Phi_j^{(n)*}(\vec{\boldsymbol{x}'})$$

with T a sufficiently long period of time for which the space-time signal $\vec{\boldsymbol{u}}(\vec{\boldsymbol{x}}, t)$ is known and with N_{POD} the number of POD modes i.e. the size of the eigenvalue problem (34). Note that the eigenfunctions arising from this decomposition are purely spatial.

Discussion of the size of the eigenvalue problem. Given M the number of spatial points[18] of the snapshots data and n_c the number of components of the variable $\vec{\boldsymbol{u}}$ used for the decomposition, $N_{POD} = M \times n_c$. Now, suppose one performs a detailed numerical simulation or one employs a modern measurement technique as Particle Image Velocimetry in Fluid Mechanics. In each case, a large number of gridpoints M can be obtained and the size of the POD problem can then quickly become too large to be solved with a good numerical precision even with numerical library dedicated to this kind of problem, like the ARPACK library[19].

Nevertheless, as it will be demonstrated in §6, the POD method can be viewed as the generalization of the harmonic decomposition to the inhomogeneous directions. So one way to take into account this size constraint with the classical POD approach, is to decompose the flow directions in homogeneous directions and inhomogeneous directions as it was done in general in experimental approaches (see Delville *et al.*, 1999; Ukeiley *et al.*, 2001).

Now, suppose the number of ensemble members deemed adequate for a description of the process is N_t with $N_t \ll M$ (the question of determining N_t is not addressed), then even if the eigenvalue problem can be accurately solved, time can be saved in solving an eigenvalue problem of size N_t. This remark is at the heart of the method of snapshots presented in §5.4.

5.4 Snapshot POD

The snapshot POD method, suggested by Sirovich (1987a,b,c), is the exact symmetry of the classical POD. The average operator $\langle . \rangle$ is evaluated as a space average over the domain Ω of interest:

$$\langle . \rangle = \int_{\Omega} . \, d\vec{\boldsymbol{x}}$$

and the variable $\vec{\boldsymbol{X}}$ is assimilated to time t.

The principle of the snapshot POD method is schematically described in figure 9.

[18] $M = N_x \times N_y \times N_z$ where N_x, N_y and N_z are the number of nodes of the experimental or numerical grid respectively in directions X, Y and Z.

[19] http://www.caam.rice.edu/software/ARPACK

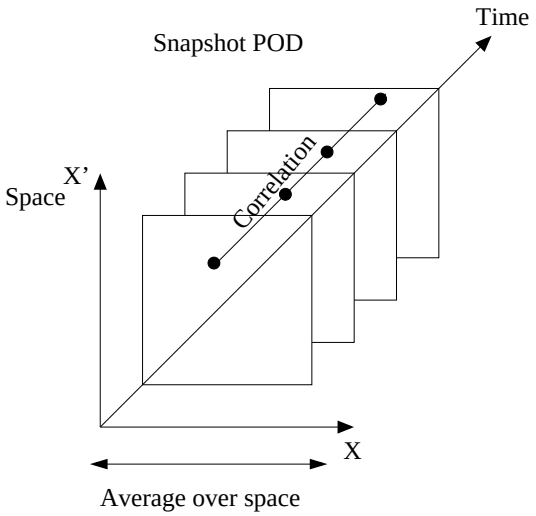


Figure 9: Schematic view of the snapshot POD.

5.4.1 Discrete eigenvalue problem

To derive the discrete eigenvalue problem corresponding to the snapshot POD, we assume[20] that $\vec{\mathbf{\Phi}}$ has a special form in terms of the original data:

$$\vec{\mathbf{\Phi}}(\vec{\boldsymbol{x}}) = \sum_{k=1}^{N_t} a(t_k)\vec{\boldsymbol{u}}(\vec{\boldsymbol{x}}, t_k) \tag{35}$$

where the coefficients $a(t_k)$, $k = 1, \cdots, N_t$ are to be determined so that $\vec{\mathbf{\Phi}}$ given by the expression (35) provides a maximum for (16) i.e. is the solution of equation (34) written here for convenience as:

$$\int_\Omega R(\vec{\boldsymbol{x}}, \vec{\boldsymbol{x}}')\, \vec{\mathbf{\Phi}}(\vec{\boldsymbol{x}}')\, d\vec{\boldsymbol{x}}' = \lambda\, \vec{\mathbf{\Phi}}(\vec{\boldsymbol{x}}). \tag{36}$$

The two-point spatial correlation tensor $R(\vec{\boldsymbol{x}}, \vec{\boldsymbol{x}'})$ is estimated under stationarity and ergodicity assumptions as:

$$R(\vec{\boldsymbol{x}}, \vec{\boldsymbol{x}'}) = \frac{1}{T}\int_T \vec{\boldsymbol{u}}(\vec{\boldsymbol{x}}, t) \otimes \vec{\boldsymbol{u}}^*(\vec{\boldsymbol{x}'}, t)\; d\,t = \frac{1}{N_t}\sum_{i=1}^{N_t} \vec{\boldsymbol{u}}(\vec{\boldsymbol{x}}, t_i) \otimes \vec{\boldsymbol{u}}^*(\vec{\boldsymbol{x}'}, t_i)$$

Substituting this expression of R and the decomposition (35) of $\vec{\mathbf{\Phi}}$ into

[20] More exactly, the properties of the span of the POD eigenfunctions guarantee that such a development exists (see Holmes *et al.*, 1996).

equation (36), we obtain:

$$\sum_{i=1}^{N_t}\left(\sum_{k=1}^{N_t}\frac{1}{N_t}\left(\int_\Omega \vec{\boldsymbol{u}}(\vec{\boldsymbol{x}}',t_k)\cdot\vec{\boldsymbol{u}}^*(\vec{\boldsymbol{x}'},t_i)\,d\vec{\boldsymbol{x}}'\right)\,a(t_k)\right)\vec{\boldsymbol{u}}(\vec{\boldsymbol{x}},t_i)=\lambda\sum_{k=1}^{N_t}a(t_k)\vec{\boldsymbol{u}}(\vec{\boldsymbol{x}},t_k)$$

and we conclude that a sufficient condition for the coefficients $a(t_k)$ to be a solution of equation (36) is to verify:

$$\sum_{k=1}^{N_t}\frac{1}{N_t}\left(\vec{\boldsymbol{u}}(\vec{\boldsymbol{x}}',t_k),\vec{\boldsymbol{u}}^*(\vec{\boldsymbol{x}'},t_i)\right)\,a(t_k)=\lambda\,a(t_i)\quad i=1,\cdots,N_t \tag{37}$$

This can be rewritten as the eigenvalue problem:

$$\boxed{C\vec{\boldsymbol{V}}=\lambda\vec{\boldsymbol{V}}} \tag{38}$$

where

$$C_{ki}=\frac{1}{N_t}\int_\Omega \vec{\boldsymbol{u}}(\vec{\boldsymbol{x}},t_k)\cdot\vec{\boldsymbol{u}}^*(\vec{\boldsymbol{x}},t_i)\,d\vec{\boldsymbol{x}}\quad\text{and}\quad\vec{\boldsymbol{V}}=\begin{bmatrix}a(t_1)\\a(t_2)\\\vdots\\a(t_{N_t})\end{bmatrix}$$

Note that in order for (37) to be a necessary condition, one needs to assume that the observations $\vec{\boldsymbol{u}}(\vec{\boldsymbol{x}},t_i)$, $i=1,\cdots,N_t$ are linearly independent.

Since C is a nonnegative Hermitian matrix, it has a complete set of orthogonal eigenvectors:

$$\vec{\boldsymbol{V}}^{(1)}=\begin{bmatrix}a^{(1)}(t_1)\\a^{(1)}(t_2)\\\vdots\\a^{(1)}(t_{N_t})\end{bmatrix},\ \vec{\boldsymbol{V}}^{(2)}=\begin{bmatrix}a^{(2)}(t_1)\\a^{(2)}(t_2)\\\vdots\\a^{(2)}(t_{N_t})\end{bmatrix},\cdots,\ \vec{\boldsymbol{V}}^{(N_t)}=\begin{bmatrix}a^{(N_t)}(t_1)\\a^{(N_t)}(t_2)\\\vdots\\a^{(N_t)}(t_{N_t})\end{bmatrix}$$

along with a set of eigenvalues $\lambda^{(1)}\geq\lambda^{(2)}\geq\cdots\geq\lambda^{(N_t)}\geq 0$. Now, for reason of simplicity, we can impose that the projection coefficients $a(t_k)$, $k=1,\cdots,N_t$ verify the same orthonogonality conditions as for the classical POD. Then we can normalize the temporal eigenfunctions $\vec{\boldsymbol{V}}^{(i)}$ by requiring that:

$$\frac{1}{N_t}\left(\vec{\boldsymbol{V}}^{(n)},\vec{\boldsymbol{V}}^{(m)}\right)=\frac{1}{N_t}\sum_{k=1}^{N_t}a^{(n)}(t_k)a^{(m)*}(t_k)=\lambda^{(n)}\delta_{nm}$$

It is now easy to check that if the POD eigenfunctions $\vec{\mathbf{\Phi}}^{(n)}(\vec{\boldsymbol{x}})$ are not estimated via equation (35) but as:

$$\vec{\mathbf{\Phi}}^{(n)}(\vec{\boldsymbol{x}}) = \frac{1}{N_t\,\lambda^{(n)}} \sum_{k=1}^{N_t} a^{(n)}(t_k)\vec{\boldsymbol{u}}(\vec{\boldsymbol{x}}, t_k) \tag{39}$$

then the spatial modes are orthonormal:

$$\int_\Omega \vec{\mathbf{\Phi}}^{(n)}(\vec{\boldsymbol{x}}) \cdot \vec{\mathbf{\Phi}}^{(n)}(\vec{\boldsymbol{x}})\, d\,\vec{\boldsymbol{x}} = \delta_{nm}.$$

5.4.2 Continuous eigenvalue problem

So far, the snapshot POD method was presented as Sirovich did in his original work (see Sirovich, 1987a,b,c). Therefore the eigenvalue problem (38) is discrete and not continuous as was defined the eigenvalue problem (34) derived for the classical POD. However, deducing from equation (38) an integral Fredholm equation is immediate, we obtain:

$$\boxed{\int_T C(t,t')a^{(n)}(t')\; dt' = \lambda^{(n)}a^{(n)}(t)} \tag{40}$$

where $C(t,t')$ is the **two-point temporal correlation tensor** defined[21] as:

$$C(t,t') = \frac{1}{T}\int_\Omega u_i(\vec{\boldsymbol{x}},t)u_i(\vec{\boldsymbol{x}},t')\; d\vec{\boldsymbol{x}} = \frac{1}{T}\sum_{n=1}^{N_{POD}} a^{(n)}(t)a^{(n)*}(t').$$

The main properties of the snapshot POD are the following:

1. The eigenfunctions are purely time dependent.
2. No cross-correlations appear in the kernel.
3. Homogeneity hypothesis is not required to lower the size of the eigenvalue problem.
4. Linear independence of the snapshots is assumed.
5. The size of the eigenvalue problem (40) is $N_{POD} = N_t$. Then as was already mentionned in §5.3, the snapshot POD reduces drastically computationnal effort when M the number of spatial points of the snapshots data is much greater than N_t. For this reason, every time this condition is verified, the snapshot POD will be preferred.

[21] In this definition, the summation over i is implicit.

5.5 Common properties of the two POD approaches

5.5.1 General properties

Whatever the particular method used to determine the spatial and temporal POD eigenfunctions, they verified the following properties:

1. Each space-time realization $u_i(\vec{\boldsymbol{x}}, t)$ can be expanded into orthogonal basis functions $\Phi_i^{(n)}(\vec{\boldsymbol{x}})$ with uncorrelated coefficients $a^{(n)}(t)$:

$$u_i(\vec{\boldsymbol{x}}, t) = \sum_{n=1}^{N_{POD}} a^{(n)}(t)\Phi_i^{(n)}(\vec{\boldsymbol{x}})$$

2. The spatial modes $\vec{\boldsymbol{\Phi}}^{(n)}(\vec{\boldsymbol{x}})$ are specified to be orthonormal:

$$\int_\Omega \vec{\boldsymbol{\Phi}}^{(n)}(\vec{\boldsymbol{x}}) \cdot \vec{\boldsymbol{\Phi}}^{(m)}(\vec{\boldsymbol{x}})\ d\vec{\boldsymbol{x}} = \delta_{nm}$$

3. The temporal modes $a^{(n)}(t)$ are orthogonal:

$$\frac{1}{T}\int_T a^{(n)}(t)a^{(m)*}(t)\ dt = \lambda^{(n)}\delta_{nm}$$

5.5.2 Incompressibility and boundary conditions

The spatial basis functions $\Phi_i^{(n)}(\vec{\boldsymbol{x}})$ can be calculated from the velocities $u_i(\vec{\boldsymbol{x}}, t)$ and the coefficients $a^{(n)}(t)$, by integrating over a sufficiently long period of time T and normalizing by the eigenvalues $\lambda^{(n)}$:

$$\Phi_i^{(n)}(\boldsymbol{x}) = \frac{1}{T\,\lambda^{(n)}}\int_T u_i(\boldsymbol{x}, t)a^{(n)*}(t)\ dt \tag{41}$$

The POD basis functions are then represented as linear combinations of instantaneous velocity fields. Therefore, all the properties of the snapshots that can be written as linear and homogeneous equations pass directly to the POD basis functions. For example, if the snapshots are divergence free, then we obtain divergence-free POD basis functions:

$$\vec{\boldsymbol{\nabla}} \cdot \vec{\boldsymbol{u}} = 0 \Longrightarrow \vec{\boldsymbol{\nabla}} \cdot \vec{\boldsymbol{\Phi}}^{(n)} = 0 \quad \forall n = 1, \cdots, N_{POD}$$

If the snapshots satisfy homogeneous Dirichlet boundary conditions then we also obtain POD basis functions satisfying homogeneous boundary conditions.

5.6 Snapshot POD or "classical" POD ?

As it was presented respectively in §5.3 and §5.4, two different POD approaches exist, the classical POD and the snapshot POD, then how can we choose for each practical configurations the pertinent method? The answer is mainly determined by the particular data set available for the evaluation of the kernels.

In one hand, data obtained by numerical simulations like Direct Numerical Simulation or Large Eddy Simulation can be highly resolved in space and time but due to cost considerations only a very short time sample is simulated. In the same vein, a good spatial resolution can be obtained by Particle Image Velocimetry, but associated to a poor temporal resolution. These two configurations, characterized by a moderate time history and a high spatial resolution, correspond to situations for which the two-point temporal correlation tensor $C(t,t')$ is statistically well converged.

On the other hand, experimental approaches like Hot Wire Anemometry or Laser Doppler Anemometry provide a well defined time description but with limited spatial resolution. These measurements techniques enabled long time history and moderate spatial resolution. Therefore, the two-point spatial correlation tensor $R_{ij}(\vec{\boldsymbol{x}},\vec{\boldsymbol{x}'})$ is statiscally well converged.

In conclusion, the data issued from an experimental approach will be generally[22] treated by using the classical method and data issued from numerical simulations by the snapshots method.

6 POD and harmonic analysis

As long as the domain $\mathcal{D}$ defined in (18) is bounded, the Hilbert-Schmidt theory applies (see Riesz and Nagy, 1955) and all the properties stated in §4.2 hold. It is thus necessary to pay special attention to flow directions assumed to be homogeneous, stationary or periodic.

6.1 A first approach: homogeneity in one direction

As a first approach, we can assume, for example, that the spatial direction OX_3 is homogeneous (a generalization including other directions is straightforward). If OX_3 is homogeneous then the two-point correlation $R(\vec{\boldsymbol{x}},\vec{\boldsymbol{x}'})$

[22] An exception is the case of data sets obtained from Particle Image Velocimetry.

depends only on the difference $r_3 = x_3' - x_3$ of the two coordinates in the OX_3 direction:

$$R_{ij}\left(x_1, x_1', x_2, x_2', x_3, x_3', t, t'\right) = R_{ij}\left(x_1, x_1', x_2, x_2', t, t'; r_3\right).$$

Splitting the space-time variable $\vec{\boldsymbol{X}} = (x_1, x_2, x_3, t)$ into an homogeneous variable x_3 and an inhomogeneous variable $\vec{\boldsymbol{\chi}} = (x_1, x_2, t)$, the integral Fredholm equation (18) writes:

$$\sum_{j=1}^{n_c} \int_{\mathcal{D}'} \int_{-\infty}^{+\infty} R_{ij}(\vec{\boldsymbol{\chi}}, \vec{\boldsymbol{\chi}}'; x_3 - x_3') \Phi_j(\vec{\boldsymbol{\chi}}', x_3') d\vec{\boldsymbol{\chi}}' dx_3' = \lambda \Phi_i(\vec{\boldsymbol{\chi}}; x_3) \tag{42}$$

Under homogeneity hypothesis, we may develop the spatial eigenfunction Φ_l in a Fourier series decomposition as:

$$\Phi_l(\vec{\boldsymbol{\chi}}; r_3) = \sum_{k_3=-\infty}^{+\infty} \hat{\Phi}_l(\vec{\boldsymbol{\chi}}; k_3) \cdot \exp(2\pi j k_3 r_3) \tag{43}$$

and introduce Π_{ij} the Fourier transform of R_{ij} in the direction OX_3:

$$\begin{aligned} \Pi_{ij}(\vec{\boldsymbol{\chi}}, \vec{\boldsymbol{\chi}}'; k_3) &= \int_{-\infty}^{+\infty} R_{ij}(\vec{\boldsymbol{\chi}}, \vec{\boldsymbol{\chi}}'; r_3) \exp(-2\pi j k_3 r_3)\, dr_3 \\ &= \int_{-\infty}^{+\infty} R_{ij}(\vec{\boldsymbol{\chi}}, \vec{\boldsymbol{\chi}}'; -r_3) \exp(2\pi j k_3 r_3)\, dr_3 \end{aligned} \tag{44}$$

where k_3 is the spatial wavenumber associated to r_3.

Substituting expression (43) in equation (42), we first obtain:

$$\begin{aligned} \sum_{j=1}^{n_c} \sum_{k_3=-\infty}^{+\infty} \int_{\mathcal{D}'} \int_{-\infty}^{+\infty} R_{ij}(\vec{\boldsymbol{\chi}}, \vec{\boldsymbol{\chi}}'; -r_3) \hat{\Phi}_j(\vec{\boldsymbol{\chi}}', k_3)\,.\, \exp[2\pi j k_3 (x_3 + r_3)] d\vec{\boldsymbol{\chi}}' dr_3 = \\ \sum_{k_3=-\infty}^{+\infty} \lambda(k_3) \hat{\Phi}_i(\vec{\boldsymbol{\chi}}; k_3)\,.\, \exp(2\pi j k_3 x_3) \end{aligned} \tag{45}$$

Then replacing the two point correlation R_{ij} by his Fourier transform Π_{ij} defined by equation (44) the Fredholm equation becomes:

$$\begin{aligned} \sum_{k_3=-\infty}^{+\infty} \Big[\sum_{j=1}^{n_c} \int_{\mathcal{D}'} \Pi_{ij}(\vec{\boldsymbol{\chi}}, \vec{\boldsymbol{\chi}}'; k_3) \hat{\Phi}_j(\vec{\boldsymbol{\chi}}', k_3) d\,\vec{\boldsymbol{\chi}}' \Big] \exp(2\pi j k_3 x_3) = \\ \sum_{k_3=-\infty}^{+\infty} \lambda(k_3) \hat{\Phi}_i(\vec{\boldsymbol{\chi}}; k_3)\,.\, \exp(2\pi j k_3 x_3) \end{aligned} \tag{46}$$

Finally, unicity of the Fourier series coefficients implies that the Fredholm equation (18) is equivalent to:

$$\sum_{j=1}^{n_c} \int_{\mathcal{D}'} \Pi_{ij}(\vec{\boldsymbol{\chi}}, \vec{\boldsymbol{\chi}}'; k_3) \hat{\Phi}_j(\vec{\boldsymbol{\chi}}', k_3) d\ \vec{\boldsymbol{\chi}}' = \lambda(k_3) \hat{\Phi}_i(\vec{\boldsymbol{\chi}}; k_3) \tag{47}$$

The conclusion is that homogeneity hypothesis in OX_3 direction decouples the initial POD problem into a set of lower dimensional problems. For each Fourier wavenumber k_3 the eigenvalue problem to solve writes:

$$\boxed{\sum_{j=1}^{n_c} \int_{\mathcal{D}'} \Pi_{ij}(\vec{\boldsymbol{\chi}}, \vec{\boldsymbol{\chi}}') \hat{\Phi}_j(\vec{\boldsymbol{\chi}}') d\ \vec{\boldsymbol{\chi}}' = \lambda \hat{\Phi}_i(\vec{\boldsymbol{\chi}}) \quad \forall k_3} \tag{48}$$

Another key result is that in each homogeneous (or stationary) direction, harmonic functions are solutions of the integral Fredholm equations. Then, as a first approximation, the Proper Orthogonal Decomposition can be viewed as the generalization of the harmonic decomposition to the inhomogeneous directions.

This result is especially useful in systems where the domain $\mathcal{D}$ is of higher dimension. For example, in the study of the three-dimensional turbulent plane mixing layer realized via the classical POD in Delville *et al.* (1999) and Ukeiley *et al.* (2001), we appeal to homogeneity in the spanwise (x_3) and streamwise[23] (x_1) directions. Selecting the finite domain $[0, L_1] \times [0, L_3]$ in these variables, we used a mixed Fourier-empirical decomposition of the form:

$$\vec{\boldsymbol{u}}(\vec{\boldsymbol{x}}, t) = (L_1 L_3)^{1/2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \sum_{n=1}^{N_{POD}} a^{(n)}_{k_1,k_3}(t) \vec{\boldsymbol{\Phi}}^{(n)}(x_2; k_1, k_3)\, e^{2\pi i (k_1 x_1 + k_3 x_3)}\, dk_1\, dk_3$$

The vector-valued eigenfunctions $\vec{\boldsymbol{\Phi}}^{(n)}(x_2; k_1, k_3)$ are obtained by solving a Fredholm equation analogous to (18) in which the kernel R_{ij} is replaced by the cross-spectral tensor $\Psi_{ij}(x_2, x_2'; k_1, k_3)$ defined as the streamwise and spanwise transform of the cross-correlation tensor. More details are given in Cordier (1996); see also Delville *et al.* (1999) and Ukeiley *et al.* (2001).

[23] As detailed in Delville *et al.* (1999), time is mapped to the streamwise direction through Taylor's hypothesis.

6.2 Discussion of the "phase indetermination"

The "phase indetermination" is one of the most important limitations of the POD. This indetermination is due to the use of two-point correlations and, as it will be demonstrated in the following, appears only for directions where an harmonic decomposition was used.

Suppose that the eigenfunction $\hat{\Phi}_j(\vec{\chi}'; k_3)$ is a solution of equation (47). Then it can be easily proven that every function $\hat{\Phi}_j(\vec{\chi}'; k_3)\,\theta(k_3)$, where $\theta(k_3)$ is a random phase function, will also be a solution. The phase information between the different modes is lost, the eigenfunctions $\hat{\Phi}_i(\vec{\chi}; k_3)$ are known up to an arbitrary function $\theta(k_3)$ which needs to be determined. In particular, for the classical POD, it is impossible to obtain directly a description of the preferred modes in the physical space. However, description of the dominant modes can be obtained by using a complementary method called *the shot-noise theory* fully described in Herzog (1986) and in Moin and Moser (1989). The reader is referred to Delville *et al.* (1999) for an application of the shot-noise theory to recover from the POD eigenfunctions determined via the classical POD the dominant modes of a three-dimensional turbulent mixing layer.

An alternative way is to build, from the dominant POD eigenfunctions, a low-order model by use of a Galerkin projection of the governing equations onto the POD modes, leading to a low-order dynamical system described by a set of Ordinary Differential Equations. In this case, these equations themselves drive the missing spectral phase information. This approach has been succesfully addressed for the wall region of a turbulent boundary layer in Aubry *et al.* (1988) and for a plane turbulent mixing layer in Ukeiley *et al.* (2001). Naturally, this kind of low-order models is particularly suited for active flow control studies (see §4.4 for a discussion).

7 Evaluative summary of the POD approach

The Proper Orthogonal Decomposition is a powerful and elegant method of data analysis aimed at obtaining low-dimensional approximations of high-dimensional processes. For turbulent flows, the POD approach by itself is neither a theory nor a closure method. However, a better understanding of the role of Coherent Structures in turbulence generation can be gained with low-order dynamical system developed by Galerkin Projection of the governing equations onto the POD basis functions (see Aubry *et al.*, 1988, for example). On the other hand, the recent invention of Micro Electro Mechan-

ical Systems has generated substantial interest for control methods for fluid dynamics. The design of reduced-order controllers for fluid system is essential for real-time implementation and POD method is particularly suited for deriving reduced-order models (see Ravindran, 2000a,b; Fahl, 2000; Atwell and King, 1999).

Among the **advantages** related to the Proper Orthogonal Decomposition, the following points can be underlined:

- ▷ The method is objective, methodic and rigorous: a mathematical framework is provided by the Hilbert-Schmidt theory.
- ▷ The POD is a linear method but no linear hypothesis is imposed on the process. The fact that this approach always looks for linear or affine subspaces instead of curved submanifolds makes it computationnally tractable. However the POD does not neglect the nonlinearities of the original vector field. If the original dynamical system is nonlinear, then the resulting POD reduced order model will also typically be nonlinear.
- ▷ The POD basis functions are optimal in terms of energy.
- ▷ The efficiency of POD increases with the level of inhomogeneity of the process. Then, this method is particularly suited for the analysis of turbulent shear flows. Moreover, as the generalization of the Fourier methods to inhomogeneous directions, POD is complementary to harmonic methods.
- ▷ Combined with the Galerkin Projection procedure, POD provides a powerful method for generating lower dimensional models of dynamical systems that have a very large or even infinite dimensional space.

Among the **disadvantages** related to POD are the following:

- ▷ This technique requires the knowledge of a two-point correlation tensor over a great number of points. Its use can then be limited by the size of the data sets that can quickly becomes huge (see Cordier and Bergmann, 2002, for an example).
- ▷ Due to the use of two-point correlations, the phase indetermination appears for directions where an harmonic decomposition has to be used. For the classical POD, in particular, complementary techniques are necessary to obtain a description of the preferred modes in the physical space.

- ▷ The POD basis functions are intrinsic[24] by nature to the flow configurations from which they have been derived. Therefore, it can happen that a POD base, determined with a set of realizations of the flow model for a specified control input, can perfectly reproduce the dynamics of the flow for a fixed system and may not be sufficient when the system is under the action of a control. In these cases, the POD base need to be improved through an adaptative procedure (see Fahl, 2000; Ravindran, 2000a).

References

Algazi V.R. and Sakrison D.J. (1969): On the optimality of the Karhunen-Loève expansion. *IEEE Trans. Inform. Theory*, **15**, pp. 319-321.

Allan B.G. (2000): A reduced order model of the linearized incompressible Navier-Stokes equations for the sensor/actuator placement problem. *ICASE report N° 2000-19*. Downloadable in http://www.icase.edu/docs/library/itrs.html.

Andrews C.A., Davies J.M. and Schwartz G.R. (1967): Adaptative data compression. *Proc. IEEE*, **55**, pp. 267-277.

Atwell J.A. and King B.B. (1999): Reduced order controllers for spatially distributed systems via Proper Orthogonal Decomposition. *Technical report ICAM 99-07-01*, Virginia Tech. Downloadable in http://www.math.vt.edu/people/bbking/bbkpubs.html.

Antoulas A.C. and Sorensen D.C. (2001): Approximation of large-scale dynamical systems: an overview. *Technical report*, Dpt. of Electrical Computer Engineering, Rice University, Houston.

Aubry N., Guyonnet R. and Lima R. (1991): Spatio-temporal analysis of complex signals: theory and applications. *J. Statis. Phys.*, **64** (3/4), pp. 683-739.

Aubry N., Holmes P., Lumley J.L. and Stone E. (1988): The dynamics of coherent structures in the wall region of a turbulent boundary layer. *J. Fluid Mech.*, **192**, pp. 115-173.

Berkooz G. (1991): Turbulence, coherent structures, and low dimensional models. *PhD dissertation, Cornell university.*

Berkooz G., Holmes P. and Lumley J.L. (1993): The Proper Orthogonal Decomposition in the analysis of turbulent flows. *Annual Review of Fluid Mech.*, **25**, pp. 539-575.

[24] The same argument can be used to explain the energetic optimality of the POD basis functions.

Bonnet J.-P. and Delville J. (2002): Coherent structures in turbulent flows and numerical simulations approaches. *Lecture series 2002-04* on post-processing of experimental and numerical data, Von Karman Institute for Fluid Dynamics.

Chatterjee A. (2000): An introduction to the Proper Orthogonal Decomposition. *Current Science*, **78**, n° 7, pp. 808-817.

Christensen E.A., Brøns M. and Sørensen J.N. (1998): Evaluation of POD-based decomposition techniques applied to parameter-dependent non turbulent flows. *DCAMM, report N°. 573*, Technical University of Denmark.

Cordier L. (1996): Etude de systèmes dynamiques basés sur la décomposition orthogonale aux valeurs propres (POD). Application à la couche de mélange turbulente et à l'écoulement entre deux disques contra-rotatifs. *PhD dissertation, Poitiers university*, in french.

Cordier L. and Bergmann M. (2002): Two typical applications of POD: coherent structures eduction and reduced order modelling. *Lecture series 2002-04* on post-processing of experimental and numerical data, Von Karman Institute for Fluid Dynamics.

Courant R. and Hilbert D. (1953): Methods of mathematical physics. Vol. 1. John Wiley & Sons, New-York.

Delville J., Cordier L. and Bonnet J.-P. (1998): Large-scale structure identification and control in turbulent shear flows. In *Flow Control: Fundamental and Practices*, M. Gad-el-Hak, A. Pollard and J.-P. Bonnet eds, Lecture Notes in Physics, m 53, Springer, pp. 199-273.

Delville J., Ukeiley L., Cordier L., Bonnet J.-P. and Glauser M. (1999): Examination of large-scale structures in a turbulent mixing layer. Part 1. Proper Orthogonal Decomposition. *J. Fluid Mech.*, **391**, pp. 91-122.

Fahl M. (2000): Trust-Region methods for flow control based on Reduced Order Modeling. *PhD dissertation, Trier university.* Downloadable in http://www.mathematik.uni-trier.de:8080/~fahl.

Fiedler H.E. (1998): Control of free turbulent shear flows. In *Flow Control: Fundamental and Practices*, M. Gad-el-Hak, A. Pollard and J.-P. Bonnet eds, Lecture Notes in Physics, m 53, Springer, pp. 336-429.

Golub G.H. and Van Loan C.F. (1990): Matrix computations, second edition. The Johns Hopkins University Press, Baltimore.

Gordeyev S. (1999): Investigation of coherent structure in the similarity region of the planar turbulent jet using POD and wavelet analysis. *PhD dissertation, university of Notre Dame.* Downloadable in http://www.nd.edu/~sgordeye/research.html.

Graham W.R., Peraire J. and Tang K.Y. (1997): Optimal control of vortex shedding using low order models. Part 1: Open-loop model development. *Submitted to Int. J. for Num. Meth. in Engr.*. Downloadable in http://raphael.mit.edu/pubs.html.

Herzog S. (1986): The large scale structure in the near-wall of turbulent pipe flow. *PhD Dissertation, Cornell University.*

Higham N.J. (1989): Matrix nearness problems and applications. In *Applications of matrix theory*, Glover and Barnett editors, Clarendon Press, pp. 1-27

Hinze M. (2000): Optimal and instantaneous control of the instationary Navier-Stokes equations. *Accreditation to supervise research dissertation, Berlin University.* Downloadable in http://www.math.tu-berlin.de/~ hinze/.

Holmes P., Lumley J.L. & Berkooz G. (1996): Turbulence, Coherent Structures, Dynamical Systems and Symmetry. *Cambridge Monographs on Mechanics.*

Hotelling H. (1933): Analysis of a complex statistical variables into principal components. *Journal of Educational Psychology*, **24**, pp. 417-441.

Hubert L., Meuleman J. and Heiser W. (2000): Two purposes for matrix factorization: a historical appraisal. *SIAM Review*, **42**, pp. 68-82.

Iollo A. (1997): Remarks on the approximation of the Euler equations by a low order model. *INRIA research report n° 3329.* Downloadable in http://www.inria.fr/rrrt/rr-3329.html.

Iollo A., Lanteri S. and Désidéri J.-A. (1998): Stability properties of POD-Galerkin approximations for the compressible Navier-Stokes equations. *INRIA research report n° 3589.* Downloadable in http://www.inria.fr/rrrt/rr-3589.html.

Joliffe I.T. (1986): Principal Component Analysis. *Springer-Verlag*, New-York.

Karhunen K. (1946): Zur spektral theorie stochasticher prozesse. *Ann. Acad. Sci. Fennicae*, Ser. A1, **34**.

Kirby M. and Sirovich L. (1990): Application of the Karhunen-Loeve procedure for the characterization of human faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **12**, N^o1, pp. 103-108.

Kosambi D.D. (1943): Statistics in function space. *J. Indian Math. Soc.*, **7**, pp. 76-88.

Loève M. (1945): Fonctions aléatoires du second ordre. *Compte Rend. Acad. Sci.*, Paris, **220**.

Loève M. (1955): Probability Theory. *Van Nostrand.*

Lumley J.L. (1967): The structure of inhomogeneous turbulence. *Atmospheric Turbulence and Wave Propagation*, ed. A.M. Yaglom & V.I. Tatarski, pp. 166-178.

Lumley J.L. (1970): Stochastic Tools in Turbulence. Academic Press, New York.

Moin P. and Moser R.D. (1989): Characteristic-eddy decomposition of turbulence in a channel. *J. Fluid Mech.*, **200**, pp. 471-509.

Obukhov A.M. (1941): Energy distribution in the spectrum of a turbulent flow. *Izvestya Aked. Nauk. SSSR, Ser. Geogr. Geophys.* , **4-5**, pp. 453-466.

Obukhov A.M. (1954): Statistical description of continuous fields. *Trudy Geophys. Int. Aked. Nauk. SSSR*, **24**, pp. 3-42.

Papoulis A. (1965): Probability, Random variables, and Stochastic Processes. *Mc Graw-Hill, New-York.*

Pougachev V.S. (1953): General theory of the correlations od random functions. *Izv. Akad. Nauk. SSSR*, Ser. Mat. **17**, pp. 401-402.

Rathinam M. and Petzold L.R. (2001): A new look at proper orthogonal decomposition. Submitted to SIAM, Journal on Numerical Analysis. Downloadable in http://www.engineering.ucsb.edu/~cse.

Ravindran S.S. (2000a): Reduced-order adaptive controllers for fluid flows using POD. *J. of Scientific Computing*, **15**, N° 4, pp. 457-478.

Ravindran S.S. (2000b): A reduced-order approach for optimal control of fluids using proper orthogonal decomposition. *int. J. Numer. Meth. Fluids*, **34**, pp. 425-448.

Rempfer D. and Fasel H.F. (1994): Evolution of three-dimensional Coherent Structures in a flat-plate boundary layer. *J. Fluid Mech.*, **260**, pp. 351-375.

Riesz F. and Nagy B.S. (1955): Functionnal Analysis. Ungar, N.Y.

Rivlin T.J. (1981): An introduction to the approximation of functions. *Dover.*

Rowley C.W (2002): Modeling, simulation and control of cavity flow oscillations. *PhD dissertation, California Institue of technology.* Downloadable in http://www.princeton.edu/~cwrowley.

Sanghi S. (1991): Mode interaction models in near-wall turbulence. *PhD dissertation, Cornell university.*

Sirovich L. (1987a): Turbulence and the dynamics of coherent structures. Part 1 :Coherent structures. Quarterly of Applied Mathematics **XLV**, N° 3, pp. 561–571.

Sirovich L. (1987b): Turbulence and the dynamics of coherent structures. Part 2 : Symmetries and transformations. Quarterly of Applied Mathematics **XLV**, N° 3, pp. 573–582.

Sirovich L. (1987c): Turbulence and the dynamics of coherent structures. Part 3 : Dynamics and scaling. Quarterly of Applied Mathematics **XLV**, N° 3, pp. 583–590.

Temam R. (1988): Infinite-dimensional dynamical systems in Mechanics and Physics. *Springer-Verlag, New York.*

Townsend A.A. (1976): The structure of turbulent shear flow. Second Edition, *Cambridge University Press.*

Ukeiley L., Cordier L., Manceau R., Delville J., Glauser M. and Bonnet J.-P. (2001): Examination of large-scale structures in a turbulent mixing layer. Part 2. Dynamical systems model. *J. Fluid Mech.*, **441**, pp. 67-108.

Volkwein S. (1999): Proper Orthogonal Decomposition and Singular Value Decomposition. *Technical report Institut für Mathematik n° 153* , Graz university. Downloadable in http://www.kfunigraz.ac.at/imawww/volkwein/publist.html.

Volkwein S. (2001): Optimal and suboptimal control of Partial Differential Equations: augmented Lagrange-SQP methods and Reduced Order Modeling with Proper Orthogonal Decomposition. *Accreditation to supervise research dissertation, Graz university.* Downloadable in http://www.kfunigraz.ac.at/imawww/volkwein/publist.html.

CISM/ERCOFTAC Advanced Course: Eddy Structure Identification Techniques in Free Turbulent Shear Flows. R. Adrian, J.P. Bonnet, J. Delville, F. Hussain, J. Lumley, O. Metais, C. Vassilicos. Udine, Italy, 23-27 June 1994. Springer-Verlag 1996.