

Décomposition en Modes Dynamiques Économique pour l'identification de structures cohérentes dans des écoulements 3D

Florimond Guéniat^{1,2}, Maxime Delorme¹, Luc R. Pastur^{1,2}, & François Lusseyran²

¹ LIMSI CNRS

² Université Paris Sud

florimond.gueniat@limsi.fr

Résumé. Dans cette contribution, nous présentons une méthode pour extraire des structures spatiales cohérentes de grands jeux de données, par une variante économique de la décomposition en modes dynamiques. La convergence de l'algorithme repose sur une décimation pertinente de l'information. Une nouvelle définition de qualité des observables conservées est introduite, qui ne nécessite pas de connaître les équations d'évolutions du système.

Abstract. In this paper, we present an efficient way to use Dynamical Mode Decomposition on large dataset. This method relies on a smart decimation of the dataset. Therefore, an equation-free definition for the quality of selected observables is proposed.

1 Introduction

Un système physique ouvert possède, potentiellement, un nombre infini de degrés de liberté. Néanmoins, les écoulements sont le plus souvent organisés autour de structures cohérentes qui jouent un rôle décisif dans la dynamique. On peut par exemple penser aux grandes structures tourbillonnaires des allées de Von Karman, que l'on observe aussi bien dans des expériences de laboratoire que dans le sillage de structures à grande échelle telles que des navires ou des îles, pour lesquels les nombres de Reynolds sont très grands et la turbulence pleinement développée sur la gamme des échelles inertielles. Ces structures invitent à chercher des moyens de réduction de la dimension effective des écoulements considérés.

Une méthode récente, la Décomposition en Modes Dynamiques (DMD)[1,2], permet, sous l'hypothèse d'existence d'un opérateur d'évolution stationnaire, de décomposer toute réalisation \mathbf{u} du champ mesuré sous forme de modes spatiaux Φ dépendant de l'espace \mathbf{r} appelés modes dynamiques et de modes temporels $\alpha(t)$, qui prennent la forme d'exponentielles complexes. Cette méthode ne peut, malheureusement, pas toujours être utilisée dans le cas de grands jeux de données – par exemple en simulation numérique 3D –, résolues à la fois temporellement et spatialement, la mémoire et les puissances de calcul requis étant hors de portée pour l'application de la DMD.

Dans cet article, nous présentons une méthode efficace pour l'extraction des structures spatiales et temporelles DMD dans le cas de grosses bases de données. L'algorithme se base sur une dégradation de la résolution spatiale. Décimer les observables implique de choisir attentivement celles conservées, *ie* qu'elles portent une information pertinente relativement à la dynamique du champ. Cela rejoint la question de l'observabilité du système réduit. Construire les matrices de Kalman ou d'observabilité[5,6] n'est pas possible dans le cas de grands jeux de données, où la dimension du vecteur d'état de l'ordre du million. Afin de s'assurer qu'il n'y a aucune perte d'information critique, une méthode de vérification de qualité de l'observabilité des composantes sélectionnées est proposée.

Dans cette contribution, nous présentons la méthode en détail et l'illustrons sur un écoulement de cavité 3D entraînée par couche cisaillée.

2 Décomposition en mode dynamique

L'objectif de la décomposition en mode dynamique (DMD) est d'extraire d'un jeu de données, des bases spatiales et temporelles pertinentes, au vu de la dynamique, pour décrire les réalisations des champs observés, par exemple de vitesse. Nous rappelons rapidement ici l'algorithme. Le lecteur intéressé pourra trouver plus de détails dans [3,4,2].

2.1 Algorithme

Opérateur d'évolution Soit un système dynamique évoluant sur une variété \mathcal{D} de dimension $n_{\mathcal{D}}$, et X le vecteur d'état appartenant à \mathcal{D} . Soit encore $\phi_{\Delta t}$ le flot de \mathcal{D} dans \mathcal{D} , amenant $X(t)$ au temps t à l'état $X(t + \Delta t)$ au temps $t + \Delta t$; où Δt est le pas de temps. Soit Π la mesure de \mathcal{D} sur l'espace des observables Ω , de dimension n_p . L'observable \mathbf{u} s'écrit donc $\mathbf{u}(t) = \Pi(X(t))$. Dans la suite, par abus de notation, $\mathbf{u}(t_0 + n\Delta t)$ sera écrit comme \mathbf{u}_n .

On considère l'opérateur A défini par :

$$A \circ \Pi = \Pi \circ \phi_{\Delta t}, \quad (1)$$

où \circ est l'opérateur de composition. On peut remarquer, de l'équation (1) et de la définition du flot $\phi_{\Delta t}$, que l'on a :

$$A\mathbf{u}_n = A\mathbf{u}_{n+1}. \quad (2)$$

A est donc un opérateur d'évolution sur les observables. L'objectif de la DMD est d'extraire des vecteurs propres et des valeurs propres de cet opérateur.

Opérateur similaire En mécanique des fluides, la dimension de l'opérateur A est prohibitivement grande. L'algorithme DMD vise à chercher un opérateur similaire à A , de taille réduite, afin de pouvoir estimer une partie des valeurs et vecteurs propres de A . Cet opérateur réduit est construit à partir des informations disponibles, *ie* le jeu de données. Par soucis de simplicité, nous ne présentons que la méthode proposée par Schmid en 2008 [1], basée sur la construction d'une matrice compagnon.

Si $N + 1$ est le nombre de réalisations à disposition, et en notant K_i^j un jeu de données des réalisations, ordonnées en temps, *ie* $K_i^j = \{\mathbf{u}_i, \mathbf{u}_{i+1}, \dots, \mathbf{u}_{j-1}, \mathbf{u}_j\}$, alors :

$$\begin{aligned} AK_1^N &= \{A\mathbf{u}_1, \dots, A\mathbf{u}_N\} \\ &= \{\mathbf{u}_2, \dots, \mathbf{u}_{N+1}\} \\ &= K_2^{N+1}. \end{aligned} \quad (3)$$

On suppose que la dernière réalisation est linéairement dépendante des N premières. Cette hypothèse est pertinente dès lors que l'on ne rajoute pas d'informations dynamique nouvelle en rajoutant des réalisations à la base de données. On a alors :

$$\mathbf{u}_{N+1} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \dots + c_N\mathbf{u}_N + \mathbf{r}, \quad (4)$$

où \mathbf{r} est un résidu. Les éléments c_i peuvent être identifiés par minimisation de la norme :

$$\mathbf{c}_{opt} = \min_{\mathbf{c}} \left\| \mathbf{u}_{N+1} - \sum_{j=1}^N c_j \mathbf{u}_j \right\|_2. \quad (5)$$

En utilisant l'équation (3), on obtient :

$$K_2^{N+1} = K_1^N S + R. \quad (6)$$

R est un résidu. S est une matrice compagnon de taille $N \times N$. Elle permute les i premières réalisations sur les suivantes, à l'exception de la dernière, qui est développée sur la dernière colonne suivant l'équation (4). Si l'on néglige R alors la matrice compagnon est similaire à A , et donc les valeurs propres de S sont des valeurs propres de A . De plus, pour tout vecteur propre ζ de S , $\Phi \equiv K_1^N \times \zeta$ est un vecteur propre de A .

3 Décomposition en mode dynamique économique

La DMD s'est montrée très efficace pour le traitement de données et l'identification de structures à partir de données parcellaires ou faiblement résolues[4]. Néanmoins, dans le cadre de simulations numériques, où le nombre de points par réalisation est typiquement de l'ordre du million pour des champs 3D, manipuler les jeux de données de plusieurs dizaines de Go peut s'avérer problématique.

On peut reformuler la décomposition en mode dynamique, en utilisant ses propriétés spectrales :

$$K_1^N = MV + R \approx MV, \quad (7)$$

R est une matrice de résidu. $V \in \mathbb{C}^{n_p \times N}$ est matrice de Vandermonde généralisée suivante :

$$V := \begin{pmatrix} \lambda_1^{t_1} & \lambda_1^{t_2} & \dots & \lambda_1^{t_N} \\ \lambda_2^{t_1} & \lambda_2^{t_2} & \dots & \lambda_2^{t_N} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{N_{\text{md}}}^{t_1} & \lambda_{N_{\text{md}}}^{t_2} & \dots & \lambda_{N_{\text{md}}}^{t_N} \end{pmatrix}, \quad (8)$$

et $M \equiv \{\Phi_1 \dots \Phi_{N_{\text{md}}}\} \in \mathbb{C}^{n_p \times N_{\text{md}}}$ contient les modes spatiaux.

En fait, il est connu [7,8] que les matrices de Vandermonde diagonalisent les matrices compagnons. On peut alors réécrire :

$$M \approx KV^+$$

avec, en général, $N_{\text{md}} = N$. On a alors $V^+ = V^{-1}$. Dans le cas contraire, V^+ est la pseudo inverse de Moonrose de la matrice de Vandermonde V .

Identifier les valeurs propres permet donc de reconstruire les modes. Une idée naturelle pour extraire les valeurs propres de l'opérateur est d'utiliser la DMD pour capter, sur des ensembles dégradés de données, les valeurs propres indispensables à la construction de la matrice de Vandermonde.

La question du choix des observables conservées lors de la dégradation spatiale se pose alors, la qualité de l'identification des fréquences étant cruciale.

3.1 Qualité des observables

On cherche à dégrader le champ, c'est-à-dire à ne conserver qu'un vecteur $\tilde{\mathbf{u}}$, de taille $\tilde{n}_p < n_p$, issu de \mathbf{u} . $\tilde{\mathbf{u}}$ appartient à l'espace réduit $\tilde{\Omega} \subset \Omega$. On obtient le système suivant :

$$\begin{cases} \mathbf{u}_{n+1} = A\mathbf{u}_n \\ \tilde{\mathbf{u}}_n = C\mathbf{u}_n \end{cases},$$

C est la matrice de passage de Ω à $\tilde{\Omega}$. Par conséquent, C est une matrice de taille $\tilde{n}_p \times n_p$, composée de 1 et 0. La i ème ligne de C est en pratique emplie de 0 sauf pour un élément c_{ij} , tel que $\tilde{u}_i = u_j$.

L'observabilité d'un système est habituellement déterminée par l'estimation du rang de la matrice d'observabilité de Kalman $\mathcal{K} \in \mathcal{M}_{n_p \times \tilde{n}_p, n_p}$:

$$\mathcal{K} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n_p-1} \end{pmatrix}.$$

Si le rang de \mathcal{K} est n_p , alors $\tilde{\mathbf{u}}$ est observable [6].

En mécanique des fluides, l'opérateur A n'est en général pas accessible, mais la DMD permet une estimation de l'opérateur en utilisant l'équation (6) :

$$A \approx K_1^N S K_1^N.$$

En pratique, calculer A^i , pour $i > 1$, est risqué. L'opérateur est seulement estimé, et le module des valeurs propres de A peut être suffisamment supérieur à 1 pour que le conditionnement de A^i s'effondre. De plus, la dimension de \mathcal{K} peut rapidement mettre hors de portée un calcul de rang.

Lors de la propagation de \mathbf{u} :

$$\mathbf{u}_{n+1} = A\mathbf{u}_n,$$

la i ème composante de \mathbf{u}_{n+1} est donné par $\sum a_{ij}u_n^j$. Si de nombreuses composantes u_n^j sont nécessaires pour obtenir u_{n+1}^i , c'est-à-dire si de nombreuses composantes a_{ij} sont significativement non nulles sur la i ème ligne de A , alors la sortie u_{n+1}^i prend de l'information dans une grande partie du vecteur d'état, \mathbf{u} . Donc, une telle composante est probablement à même d'être considérée dans la recherche d'une bonne observable pour le champ réduit, $\tilde{\mathbf{u}}$.

De la même façon, une minorité de composantes a_{ji} significativement non nulles sur la i ème colonne de A indique que la i ème composante de \mathbf{u} influe peu sur l'évolution des autres composantes $j \neq i$ du vecteur état. Sa dynamique est spatialement localisée. En suivant la définition de l'observabilité, capter toute la dynamique du système implique de devoir prendre en compte une telle composante.

Se basant sur ces deux remarques, nous proposons une nouvelle définition de la qualité d'une observable, liée à sa dépendance dynamique vis-à-vis des différents points du champ. On définit la i ème composante du vecteur $\boldsymbol{\nu}$, que l'on nomme vecteur de pertinence dynamique, en comptant les composantes significatives de la i ème ligne de la matrice A moins le nombre de composantes significatives de la i ème colonne. $\boldsymbol{\nu}$ est donc un vecteur de dimension $n_p \times 1$, comme \mathbf{u} .

Seules les composantes u_i du vecteur d'état, associées aux \tilde{n}_p composantes v_i les plus grandes de $\boldsymbol{\nu}$, seront à considérer en tant qu'observables de qualité.

Cependant, la construction de A , de taille $n_p \times n_p$, et donc l'estimation de la qualité *a priori* des observables, peut être délicate dans le cas de grands jeux de données. Il est néanmoins possible de vérifier *a posteriori* l'intérêt des observables choisies. Il suffit de construire l'opérateur \tilde{A} qui agit sur des observables réduites choisies arbitrairement, en suivant le processus d'estimation de A . Par itération, on peut rejeter et remplacer les composantes ayant de mauvaises qualités d'observabilité par d'autres, en suivant la méthode proposée, jusqu'à obtenir un panel d'observables $\tilde{\mathbf{u}}$ ayant des propriétés d'observabilité maximale.

4 Résultats sur un jeu de données 3D

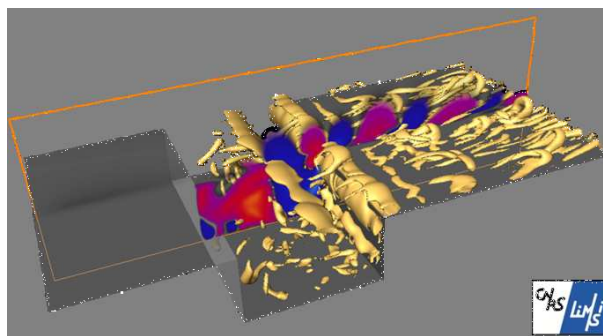


Figure 1. Vue d'un instantané de l'écoulement produit par simulation numérique 3D. L'écoulement incident vient de la gauche de l'image. Sont représentées des isosurfaces du facteur Q (zones de rotation) auxquelles se superposent le champ de vorticité, dans un plan de coupe parallèle à l'écoulement.

Pour illustrer la méthode, nous considérons un écoulement de cavité 3D entraînée par une couche cisailée *cf.* figure 1. Les écoulements de cavité sont connus pour présenter des spectres de puissance

piqués autour des fréquences d'oscillations auto-entretenues de la couche cisailée. Plus de détails peuvent être trouvés dans [9].

La sélection des points d'observations, a posteriori, a été effectuée. Nous avons sélectionné $\tilde{n}_p = 3500$ points d'observation sur un total de $n_p \approx 1200000$ points. La qualité des points sélectionnés se révèle sur les spectres de la figure 2.

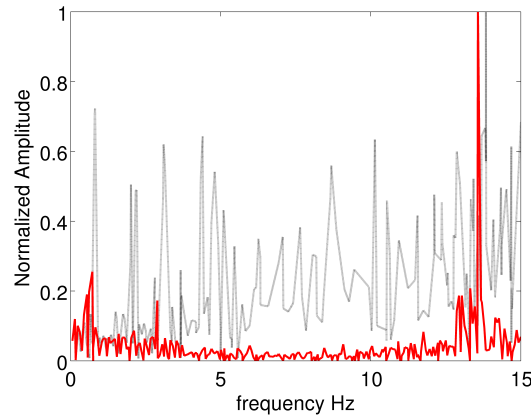


Figure2. Spectre extrait de l'opérateur \tilde{A} , en rouge pour des observables bien choisies, en noir pointillé pour des observables mal choisies. Voir texte pour le détail.

Le spectre provient de l'interprétation de l'équation (8). Pour chaque mode k , la valeur propre correspondante de l'opérateur est associée à une fréquence f_k , que l'on peut expliciter en réécrivant la valeur propre : $\lambda_k = \exp(\sqrt{-1} 2\pi f_k \Delta t)$. Le spectre est obtenu en associant à cette fréquence la norme du mode correspondant, au sens \mathcal{L}^2 .

Pour illustrer l'importance du choix des points d'observation, le spectre est tracé dans le cas d'observables bien choisies (en rouge, gris foncé), et dans le cas d'observables mal choisies, en pointille (gris clair). Comme attendu, le spectre est inutilisable dans le second cas, les observables considérées ne rendant pas bien compte de la dynamique globale.

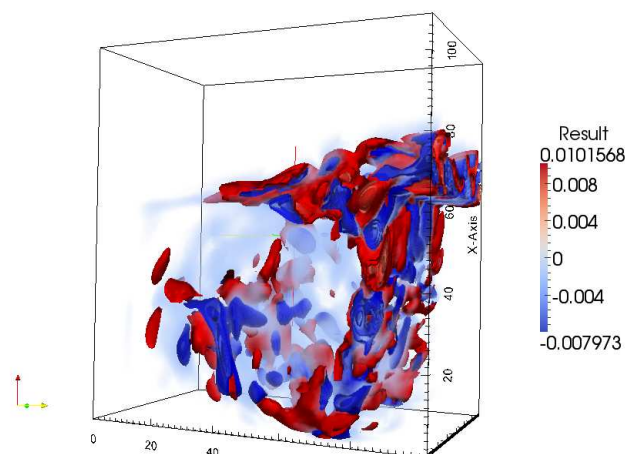


Figure3. Norme de la vitesse du mode dominant à $f = 13.57\text{Hz}$, orientée par le signe de la composante transverse de vitesse.

Le mode de couche cisailée, représenté sur la figure 3, correspond à une fréquence $f = 13.57\text{Hz}$. C'est le mode dominant, correspondant au pic principal du spectre de la figure 2. Le mode exhibe des structures tourbillonnaires dans la couche cisailée, caractéristique de l'instabilité de Kelvin-Helmoltz. On remarque que ces structures sont intrinsèquement tri-dimensionnelles, ce qui suggère l'existence d'un mode d'instabilité transverse se développant sur les tourbillons de couche cisailée avant l'impact sur le mur aval.

5 Conclusion

Dans cette contribution, on a tiré parti de la résilience de la décomposition en mode dynamique aux dégradations spatiales pour proposer une variante de l'algorithme, permettant de traiter, pour un coût raisonnable, de grands jeux de données. On peut ainsi approcher les valeurs propres de l'opérateur d'évolution calculé par la méthode classique de décomposition en modes dynamiques. Ces valeurs propres permettent la reconstruction de la matrice de Vandermonde asservissant l'évolution temporelle des modes dynamiques. L'inversion de cette matrice de Vandermonde donne, finalement, accès aux modes dynamiques 3D correspondant au jeu de données complet.

Cet algorithme requiert la qualité d'observable des points spatiaux conservés, afin de ne pas introduire d'erreurs dans l'identification des valeurs propres, et donc, de ne pas polluer les modes. La qualité d'une observable est évaluée *a posteriori*, par l'étude de l'interdépendance des composantes du vecteur d'état, au travers de l'opérateur d'évolution approximé. Cette approximation est permise par l'algorithme de la décomposition en mode dynamique. Au contraire des méthodes classiques d'évaluation de l'observabilité, basées par exemples sur les matrices de Kalman, cette méthode ne nécessite pas les équations d'évolution du système.

L'algorithme a été testé avec succès sur un écoulement de cavité ouverte 3D duquel les modes dynamiques 3D ont pu être extraits à partir de données spatialement dégradées.

Références

1. P.J. SCHMID & J. SESTERHENN, Dynamic mode decomposition of numerical and experimental data, in *61st annual meeting of the APS division of fluid dynamics*, APS, San Antonio, Texas, USA (2008)
2. C. W. ROWLEY, I. MEZIC, S. BAGHERI, P. SCHLATTER & D. S. HENNINGSON, Spectral analysis of nonlinear flows, *Journal of Fluid Mechanics* **641**, pp. 115-127 (2009)
3. F. Guéniat, J. Basley, C. Douay, L. Pastur, F. Lusseyran, Th. Faure : Opérateur de Koopman et identification de structures cohérentes : application à un écoulement de cavité, *Rencontre du Non Linéaire*, **2011**
4. P.J. SCHMID, Dynamic mode decomposition of numerical and experimental data, *Journal of Fluid Mechanics* **656**, pp. 6-28 (2010)
5. Kalman R. E. : On the General Theory of Control Systems, *Proc. 1st Int. Cong. of IFAC*, Moscow 1960, London 1961.
6. Stanley G.M. and Mah, R.S.H. : Observability and Redundancy in Process Data Estimation, *Chem. Engng. Sci.* **36**, 259 (1981)
7. Bellman, Richard : Introduction to Matrix Analysis, SIAM (1987)
8. K. Chen, J.H. Tu & C.W. Rowley, Variants of Dynamic Mode Decomposition : Boundary Condition, Koopman, and Fourier Analyses, *Journal of Nonlinear Science*, 2012, Volume 22, pp 887-915.
9. J. BASLEY Experimental investigation of global structures, in an incompressible cavity flow using time-resolved PIV *Experiments in Fluids*- (2010)