



**HAL**  
open science

# Construction de modèles d'ordre réduit non-linéaires basés sur la décomposition orthogonale propre pour l'aéroélasticité

Antoine Placzek

► **To cite this version:**

Antoine Placzek. Construction de modèles d'ordre réduit non-linéaires basés sur la décomposition orthogonale propre pour l'aéroélasticité. Mécanique [physics.med-ph]. Conservatoire national des arts et métiers - CNAM, 2009. Français. NNT: . tel-00461691

**HAL Id: tel-00461691**

**<https://theses.hal.science/tel-00461691v1>**

Submitted on 5 Mar 2010

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

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

# THÈSE

présentée par

Antoine PLACZEK

en vue de l'obtention du grade de  
Docteur du Conservatoire National des Arts et Métiers

Discipline : Mécanique  
École Doctorale ED415

---

CONSTRUCTION DE MODÈLES D'ORDRE RÉDUIT  
NON-LINÉAIRES BASÉS SUR LA DÉCOMPOSITION  
ORTHOGONALE PROPRE POUR L'AÉROÉLASTICITÉ.

---

*Directeur de Thèse : Roger OHAYON*  
*Encadrement : Duc-Minh TRAN*

Thèse soutenue le 16 décembre 2009 devant le jury composé de :

Charbel FARHAT	Université de Stanford	<i>Président du jury</i>
Aziz HAMDOUNI	Université de La Rochelle	<i>Rapporteur</i>
Angello IOLLO	Université de Bordeaux	<i>Rapporteur</i>
Jean-Pierre GRISVAL	ONERA Châtillon	<i>Examineur</i>
Roger OHAYON	CNAM Paris	<i>Examineur</i>
Duc-Minh TRAN	ONERA Châtillon	<i>Examineur</i>
Christophe LE GARREC	Airbus Toulouse	<i>Invité</i>



---

# Remerciements

---

Le présent mémoire de thèse résume trois années de recherches qui n'auraient sans doute jamais vu le jour sans la contribution ni le soutien d'un certain nombre de personnes que je tiens à remercier ici.

Je souhaite tout d'abord remercier M. le Professeur Roger OHAYON pour la confiance qu'il m'a accordée en me proposant ce sujet de recherche, et la liberté dont j'ai joui pour le mener à bien. Je remercie les rapporteurs, MM. les Professeurs Aziz HAMDOUNI et Angello IOLLO, pour avoir accepté de relire ce (long) document. Leurs remarques avisées m'auront permis d'aborder certaines problématiques avec un œil nouveau. Je remercie aussi M. le Professeur Charbel FARHAT de s'être déplacé depuis Stanford et d'avoir accepté de présider le jury au cours de la soutenance. Enfin, je tenais à remercier M. Jean-Pierre GRISVAL pour m'avoir accueilli au cours de ces trois années au sein du département DADS de l'Onera, et pour avoir renouvelé sa confiance pour les prochaines années.

Je tenais aussi à remercier Minh qui m'a encadré au cours de cette thèse et plus généralement l'ensemble des collègues du département DADS pour l'ambiance qui y règne. Merci aussi à Alain pour son aide et ses conseils à propos de l'utilisation du code *elsA* et plus généralement des aspects concernant la simulation numérique en aéroélasticité pour les turbomachines.

Merci également à la Team TDM (Fabrice, Jean-Sébastien, Mahdi, Minh, Naka, Sylvie, ainsi que ceux qui s'y sont greffés plus ou moins régulièrement) avec qui la pause café a toujours été l'occasion de passer des bons moments à discuter de tout et n'importe quoi, à essayer de suivre d'une oreille plus ou moins attentive les cours de (manga) japonais, ou les poèmes chinois. Merci aussi aux volleyeurs de l'ASCO avec qui les matchs ont toujours été l'occasion de se changer les idées, de se défouler un peu, et même de gagner par deux fois le ZIPEC !

Je tiens aussi à remercier ceux qui sont à l'origine de ma motivation pour le travail de recherche et que j'ai rencontré à La Rochelle (Dondon, Erwan, Sylvain, Aziz,...) ainsi qu'à Nantes (Jean-François, Cédric).

Enfin un grand merci à ceux qui ne sont certainement pas les derniers à devoir être remerciés, à savoir la bande des compiégnais qui est toujours là depuis toutes ces années, à ma famille qui m'a toujours supporté et qui m'a permis d'arriver jusqu'ici, et ce, en ayant toujours le choix, et enfin tout particulièrement à Stéphanie, grâce à qui les moments difficiles de fin de thèse sont restés malgré tout fort agréables.



---

# Résumé

---

La prédiction numérique de la réponse de systèmes aéroélastiques requiert la résolution d'un nombre important d'équations lorsque les techniques classiques de discrétisation éléments finis ou volumes finis sont employées. Le coût de calcul associé devient rapidement rédhibitoire dès lors qu'il s'agit d'effectuer une analyse paramétrique du système afin de déterminer ses limites de stabilité, d'appliquer un contrôle actif ou encore d'optimiser certains paramètres.

Le recours à un modèle d'ordre réduit le plus représentatif qui soit du système dynamique initial devient alors indispensable pour réduire les coûts de calcul et rendre ainsi possible les études paramétriques. Parmi les multiples méthodes de réduction développées par le passé, la projection de Galerkin sur une base de vecteurs issus de la décomposition orthogonale propre (POD) d'un ensemble de réponses du système s'est imposée en mécanique des fluides.

La difficulté de la construction d'un tel modèle d'ordre réduit pour les systèmes aéroélastiques est triple : tout d'abord les équations de Navier-Stokes d'un fluide compressible s'accordent difficilement avec la projection de Galerkin lorsqu'il s'agit de traiter les non-linéarités, ensuite le mouvement ou la déformation du domaine de calcul pour les applications en aéroélasticité entraîne une incohérence dans la définition des modes POD, et enfin le système réduit demande à être stabilisé pour pallier les diverses sources d'erreur introduites au cours de la construction du modèle d'ordre réduit.

Dans ce mémoire, le principe de construction d'un modèle d'ordre réduit reposant sur l'utilisation des modes POD est d'abord rappelé dans une première partie puis appliqué à un système dynamique linéaire pour lequel plusieurs formulations sont développées puis comparées pour tenir compte de conditions aux limites spécifiques. La seconde partie concerne le développement d'un modèle d'ordre réduit des équations de Navier-Stokes pour un fluide compressible visqueux. Un premier modèle d'ordre réduit développé pour un domaine de frontières fixes est corrigé au moyen de diverses techniques puis validé sur l'exemple d'un profil NACA0012 positionné de manière à provoquer le développement d'une allée instationnaire de Von Kármán dans le sillage. Dans un second temps, le modèle d'ordre réduit est étendu au cas d'un domaine mobile en faisant l'hypothèse d'un mouvement de corps rigide. Ainsi, la formulation adoptée permet de se ramener à un domaine fixe afin d'éviter le problème de définition des modes POD. Ce second modèle d'ordre réduit est alors appliqué à la reproduction de l'écoulement transsonique autour d'un profil NACA0064 animé d'un mouvement d'oscillation harmonique autour d'une position d'équilibre. Plutôt que de calculer explicitement les coefficients du modèle d'ordre réduit à partir des expressions analytiques, une identification des termes est pratiquée afin d'une part d'améliorer la précision des résultats et de diminuer considérablement le temps de calcul pour la construction du modèle réduit.



---

# Table des matières

---

<b>Introduction</b>	<b>1</b>
<b>I Choix du modèle d'ordre réduit et application à un système dynamique linéaire</b>	<b>5</b>
<b>1 État de l'art des méthodes de réduction de modèle pour l'aérodynamique</b>	<b>7</b>
1.1 Introduction	8
1.2 Principe de la réduction de modèles et classification	9
1.3 Description des principaux modèles d'ordre réduit physiques	13
1.3.1 Projection sur la base des modes propres et sous-structuration dynamique	13
1.3.2 Approximation de la fonction de transfert	15
1.3.3 Identification du système réduit	18
1.3.4 Décomposition Orthogonale Propre	21
1.3.5 Modes d'oscillation et d'interaction principaux	25
1.3.6 Équilibrage harmonique	27
1.3.7 Réduction sur une base construite <i>a priori</i>	28
1.3.8 Méthodes de réduction alternatives et dépendance avec un paramètre	29
1.4 Synthèse et justification du choix de la méthode POD	32
<b>2 Principe de construction d'un modèle d'ordre réduit POD-Galerkin</b>	<b>35</b>
2.1 Introduction	36
2.2 Description de la décomposition orthogonale propre pour un système déterministe	37
2.2.1 Décomposition orthogonale propre d'un ensemble de données	37
2.2.2 Détermination des modes POD par la méthode des clichés	41
2.2.3 Propriétés de la décomposition orthogonale propre	43
2.3 Un exemple d'application de la POD en analyse de données	47
2.4 Principe de construction du modèle d'ordre réduit à partir de la base POD	50
2.4.1 Construction d'un modèle d'ordre réduit linéaire par projection discrète	50
2.4.2 Construction d'un modèle d'ordre réduit implicite par projection sur le sous-espace	51
2.4.3 Construction d'un modèle d'ordre réduit polynomial par projection de Galerkin	52
2.4.4 Construction d'un modèle d'ordre réduit par identification des coefficients	53
2.5 Décomposition orthogonale propre des systèmes multivariés	56
2.5.1 Description de l'approche POD scalaire	56
2.5.2 Description de l'approche POD vectorielle	58
2.6 Synthèse du chapitre 2	61
<b>3 Approche POD-Galerkin pour les systèmes dynamiques linéaires non-autonomes</b>	<b>63</b>
3.1 Introduction	64
3.2 Modèles d'ordre réduit pour les systèmes dynamiques linéaires autonomes	65
3.2.1 Formulation analytique du modèle d'ordre réduit	65
3.2.2 Formulation discrète du modèle d'ordre réduit	68
3.2.3 Formulation POD-Galerkin du modèle d'ordre réduit	70
3.3 Évaluation du modèle d'ordre réduit POD-Galerkin du système linéaire autonome	74



TABLE DES MATIÈRES

3.3.1	Évaluation du modèle d'ordre réduit POD-Galerkin . . . . .	74
3.3.2	Choix de la base des clichés et du nombre de modes POD . . . . .	77
3.4	Modèles d'ordre réduit pour les systèmes dynamiques linéaires non-autonomes . . . . .	82
3.4.1	Adaptation de la méthode des pénalités au modèle d'ordre réduit . . . . .	83
3.4.2	Adaptation de la méthode des multiplicateurs de Lagrange au modèle d'ordre réduit . . . . .	84
3.4.3	Méthode de la fonction de contrôle . . . . .	85
3.5	Évaluation du modèle d'ordre réduit POD-Galerkin du système linéaire non-autonome . . . . .	87
3.5.1	Aspect des modes POD non-homogènes et homogénéisés . . . . .	87
3.5.2	Réponse du modèle d'ordre réduit non-autonome . . . . .	87
3.6	Synthèse du chapitre 3 . . . . .	93
 <b>II Développement de modèles d'ordre réduit non-linéaires pour l'aéroélasticité</b>		<b>95</b>
<b>4</b>	<b>Modélisation du fluide en vue de la construction d'un modèle d'ordre réduit</b>	<b>97</b>
4.1	Introduction . . . . .	98
4.2	Modélisation d'un fluide compressible visqueux avec les variables conservatives . . . . .	99
4.2.1	Détermination des hypothèses relatives au modèle fluide . . . . .	99
4.2.2	Caractérisation d'un fluide newtonien homogène . . . . .	100
4.2.3	Équations de Navier-Stokes en formulation ALE avec les variables conservatives . . . . .	105
4.3	Problématique de l'aéroélasticité et méthodologie de résolution . . . . .	108
4.3.1	Principe de l'aéroélasticité . . . . .	108
4.3.2	Formulations du problème de couplage . . . . .	109
4.3.3	Définition du problème aéroélastique général . . . . .	110
4.4	Modélisation du fluide avec les variables primitives modifiées . . . . .	114
4.4.1	Variables aérodynamiques pour l'obtention d'un système quadratique . . . . .	114
4.4.2	Définition du produit scalaire . . . . .	115
4.4.3	Équations de Navier-Stokes en formulation ALE avec les variables primitives modifiées . . . . .	117
4.5	Synthèse du chapitre 4 . . . . .	120
<b>5</b>	<b>Construction d'un modèle d'ordre réduit POD-Galerkin pour l'aérodynamique</b>	<b>121</b>
5.1	Introduction . . . . .	122
5.2	Construction du modèle d'ordre réduit POD-Galerkin . . . . .	123
5.2.1	Projection de Galerkin des équations de Navier-Stokes d'un fluide compressible . . . . .	123
5.2.2	Calcul des coefficients du modèle d'ordre réduit . . . . .	125
5.2.3	Intégration du modèle d'ordre réduit en temps . . . . .	129
5.3	Correction des instabilités des modèles d'ordre réduit POD-Galerkin . . . . .	130
5.3.1	Origine des instabilités et principe de la correction . . . . .	130
5.3.2	Techniques de correction a priori des modèles d'ordre réduit POD-Galerkin . . . . .	131
5.3.3	Techniques de correction a posteriori des modèles d'ordre réduit POD-Galerkin . . . . .	134
5.3.4	Autres techniques de correction des modèles d'ordre réduit POD-Galerkin . . . . .	138
5.4	Calcul des modes POD de l'écoulement autour d'un profil NACA0012 fixe en forte incidence . . . . .	139
5.4.1	Simulation numérique de l'écoulement autour du profil NACA0012 . . . . .	139
5.4.2	Calcul des modes POD de l'écoulement autour du profil NACA0012 . . . . .	142
5.5	Réponse du modèle d'ordre réduit de l'écoulement autour du profil NACA0012 . . . . .	147
5.5.1	Réponse du modèle d'ordre réduit non corrigé . . . . .	147
5.5.2	Introduction de termes correctifs dans le modèle d'ordre réduit . . . . .	149
5.5.3	Calibrage des paramètres du modèle d'ordre réduit . . . . .	154
5.5.4	Comparaison des réponses corrigées . . . . .	158
5.6	Synthèse du chapitre 5 . . . . .	160

<b>6</b>	<b>Construction d'un modèle d'ordre réduit POD-Galerkin pour l'aéroélasticité</b>	<b>161</b>
6.1	Introduction	162
6.2	Simplification des équations de Navier-Stokes pour un mouvement de corps rigide	163
6.2.1	Applicabilité de la POD en aéroélasticité	163
6.2.2	Simplification des équations de Navier-Stokes	167
6.3	Construction du modèle d'ordre réduit POD-Galerkin pour un mouvement de corps rigide	169
6.3.1	Développement d'un modèle d'ordre réduit basé sur la formulation en vitesse absolue	169
6.3.2	Développement d'un modèle d'ordre réduit basé sur la formulation en vitesse relative	175
6.3.3	Construction de modèles d'ordre réduit par identification des coefficients	178
6.3.4	Résumé et comparaison qualitative des différents modèles d'ordre réduit	180
6.4	Calcul des modes POD de l'écoulement autour d'un profil NACA0064 oscillant	182
6.4.1	Simulation numérique de l'écoulement transsonique autour du profil NACA0064	182
6.4.2	Simplification des coefficients du modèle d'ordre réduit pour les équations d'Euler	186
6.4.3	Calcul de la base des modes POD de l'écoulement autour du profil NACA0064	187
6.5	Réponse du modèle d'ordre réduit de l'écoulement autour du profil NACA0064	191
6.5.1	Réponse du modèle d'ordre réduit POD-Galerkin corrigé	191
6.5.2	Réponse du modèle d'ordre réduit POD-Galerkin identifié	198
6.6	Synthèse du chapitre 6	204
	<b>Conclusion</b>	<b>207</b>
<b>III</b>	<b>Annexes</b>	<b>209</b>
<b>A</b>	<b>Éléments d'algèbre linéaire et matriciel</b>	<b>211</b>
A.1	Matrice d'applications linéaires et changement de base	211
A.2	Décomposition en valeurs singulières	213
A.3	Résolution approchée de problèmes linéaires	214
<b>B</b>	<b>Compléments théoriques sur la décomposition orthogonale propre</b>	<b>217</b>
B.1	Expression de l'opérateur $R$ dans une base canonique	217
B.2	Décomposition spectrale de l'opérateur $R$	218
B.2.1	Détermination du rang de l'opérateur $R$	218
B.2.2	Décomposition spectrale des opérateurs de Hilbert-Schmidt	220
B.3	Équivalence entre la méthode directe et la méthode des clichés en dimension finie	222
<b>C</b>	<b>Compléments d'analyse du modèle d'ordre réduit du système linéaire autonome</b>	<b>227</b>
C.1	Influence de l'échantillonnage des clichés d'un système périodique	227
C.2	Convergence des modes POD vers les modes propres	232
C.2.1	Un résultat préliminaire	232
C.2.2	Convergence des vecteurs propres pour la méthode directe	233
C.2.3	Convergence des vecteurs propres pour la méthode des clichés	235
<b>D</b>	<b>Formulaire de calcul tensoriel</b>	<b>237</b>
D.1	Tenseurs et produits	237
D.1.1	Notations tensorielle et indicielle	237
D.1.2	Tenseurs particuliers	237
D.1.3	Produits de tenseurs	238
D.2	Opérateurs différentiels	239
D.3	Relations usuelles entre les opérateurs	241
D.4	Théorèmes de Green-Ostrogradski et d'intégration par parties	242

TABLE DES MATIÈRES

---

<b>E</b>	<b>Changements de repères et de variables pour les équations de Navier-Stokes</b>	<b>243</b>
E.1	Écriture des équations de Navier-Stokes dans un repère non galiléen . . . . .	244
E.1.1	Changement de base et repérage d'un point dans un repère entraîné . . . . .	244
E.1.2	Équations de Navier-Stokes en repère entraîné avec les composantes de la vitesse absolue . . . . .	249
E.1.3	Équations de Navier-Stokes en repère entraîné avec les composantes de la vitesse relative . . . . .	251
E.2	Impact du jeu de variables sur la forme des équations de Navier-Stokes . . . . .	258
E.2.1	Écriture générique des équations du fluide . . . . .	258
E.2.2	Variables quasi-conservatives $\mathbf{q} = [\rho, \mathbf{u}, e_I]$ et $\mathbf{q} = [\rho, \mathbf{u}, e_T]$ . . . . .	259
E.2.3	Variables thermiques $\mathbf{q} = [\rho, \mathbf{u}, \theta]$ . . . . .	260
E.2.4	Variables primitives $\mathbf{q} = [\rho, \mathbf{u}, p]$ . . . . .	261
E.2.5	Variables primitives modifiées $\mathbf{q} = [\vartheta, \mathbf{u}, p]$ . . . . .	261
E.3	Discrétisation par la méthode des Volumes Finis . . . . .	262
E.3.1	Discrétisation des équations de Navier-Stokes . . . . .	262
E.3.2	Approximation du gradient en Volumes Finis . . . . .	265
<b>F</b>	<b>Analyse des contributions des coefficients des modèles d'ordre réduit</b>	<b>267</b>
F.1	Analyse des coefficients du modèle d'ordre réduit autonome . . . . .	267
F.1.1	Comparaison des contributions convectives et diffusives . . . . .	267
F.1.2	Comparaison des contributions surfaciques et volumiques . . . . .	268
F.2	Analyse des coefficients du modèle d'ordre réduit non-autonome . . . . .	271
F.2.1	Comparaison des coefficients analytiques et identifiés . . . . .	271
F.2.2	Comparaison des contributions autonomes et non-autonomes . . . . .	274
	<b>Bibliographie</b>	<b>277</b>

---

# Nomenclature

---

## 1. Système dynamique

$\Sigma$	: ensemble d'équations gouvernant le système dynamique
$x, \mathbf{x}$	: vecteur d'état (dans un espace vectoriel quelconque et discret)
$y$	: vecteur de sortie du système dynamique
$\mu$	: vecteur de paramètres du système dynamique
$p$	: dimension du vecteur de paramètres
$t_0, t_f$	: instant initial et instant final pour l'étude du système dynamique
$n_v$	: nombre de variables indépendantes décrivant le système dynamique
$N$	: nombre d'éléments pour la discrétisation du système dynamique
$N_v$	: nombre de degrés de liberté du système dynamique discrétisé
$\mathbf{q}$	: vecteur des $n_v$ variables décrivant le système dynamique

## 2. Ensembles mathématiques et domaines

$\mathcal{V}$	: espace sur lequel est défini le vecteur d'état du système dynamique
$\mathcal{M}(n, m)$	: ensemble des matrices réelles à $n$ lignes et $m$ colonnes
$\mathcal{K}_n(\mathbf{M}, \mathbf{d})$	: espace de Krylov de dimension $n$ engendré par la matrice $\mathbf{M}$ et le vecteur $\mathbf{d}$
$\mathbb{R}^n$	: espace vectoriel euclidien de dimension $n$
$\mathbb{R}_+$	: droite des réels positifs
$I$	: intervalle de temps considéré pour l'étude du système dynamique
$I_e$	: intervalle de temps pour l'échantillonnage des clichés
$\Omega$	: domaine spatial de calcul pour la description du système dynamique
$\partial\Omega$	: frontière du domaine $\Omega$
$\Gamma$	: interface fluide-structure
$L^2(\Omega)$	: espace de Lebesgue des fonctions de carré intégrable sur un domaine $\Omega$
$d$	: dimension du domaine spatial $\Omega \subset \mathbb{R}^d$
$\mathbf{n}$	: normale extérieure au domaine $\Omega$ définie sur $\partial\Omega$

## 3. Décomposition Orthogonale Propre

$S$	: sous-espace de dimension réduite engendré par les modes POD
$q$	: dimension du sous-espace de projection
$X$	: ensemble des clichés
$M$	: nombre de clichés contenus dans $X$
$x^{(m)}$	: cliché solution du système dynamique à l'instant $t_m$
$s^{(m)}$	: cliché centré (de moyenne nulle)
$\varphi^{(j)}, \Phi$	: mode POD d'indice $j$ et matrice des modes POD
$a^{(m)}_j$	: amplitude issue de la projection du cliché à l'instant $t_m$ sur le mode POD $\varphi^{(j)}$
$a_j(t), \mathbf{a}$	: amplitude associée au mode POD $\varphi^{(j)}$ et vecteur des amplitudes à l'instant $t$
$\mathcal{P}_S$	: projecteur orthogonal sur le sous-espace $S$
$E[\cdot], \bar{\cdot}$	: opérateur de moyenne temporelle
$\alpha_m$	: poids de pondération associés à l'opérateur de moyenne pour chaque instant $t_m$
$\lambda_j$	: valeur propre associée au mode POD $\varphi^{(j)}$
$R, \mathbf{R}$	: opérateur aux valeurs propres de la méthode directe
$R^*, \mathbf{R}^*$	: opérateur aux valeurs propres de la méthode des clichés

$\mathcal{J}$	: fonctionnelle du problème d'optimisation définissant les modes POD
$r$	: rang de l'opérateur aux valeurs propres
$c^{(m)}_j$	: coefficient de la décomposition du mode POD $\varphi^{(j)}$ sur les clichés $x^{(m)}$
$\mathbf{d}^{(j)}$	: vecteur propre associé à l'opérateur aux valeurs propres de la méthode des clichés
$\epsilon_q$	: erreur de troncature de la décomposition orthogonale propre
$\eta_q$	: pourcentage d'énergie captée par les $q$ premières valeurs propres
$\tilde{\eta}_i$	: pourcentage d'énergie captée par la valeur propre $\lambda_i$ associée au mode POD $\varphi^{(i)}$

#### 4. Modèle d'ordre réduit

$n_p$	: nombre de périodes d'oscillation correspondant à la durée d'échantillonnage
$\mathcal{A}$	: matrice des corrélations temporelles intervenant dans le problème d'identification
$\mathcal{X}_i$	: vecteur des coefficients à identifier pour l'équation relative à l'amplitude modale $a_i(t)$
$p$	: nombre de valeurs singulières conservées dans l'approximation du pseudo-inverse
$N_c, N_{c,i}$	: nombre de coefficients pour toutes les (une seule) équation(s) du modèle d'ordre réduit
$N_c^a, N_c^{na}$	: nombre de coefficients autonomes et non-autonomes du modèle d'ordre réduit
$\mathcal{K}, \mathcal{L}, \mathcal{Q}$	: opérateurs autonomes constant, linéaire et quadratique du modèle d'ordre réduit
$\mathcal{S}(t)$	: terme source non-autonome du modèle d'ordre réduit
$\mathcal{D}$	: opérateur d'amortissement pour la correction du modèle d'ordre réduit
$\epsilon_\infty$	: erreur relative de reconstruction en norme infinie
$\epsilon_{L^2, x}$	: erreur spatiale relative en norme $L^2$ pour la quantité $x$
$\overline{\epsilon}_{a_j}$	: erreur moyenne en temps pour l'amplitude modale $a_j$
$\delta_{\log}$	: décrétement logarithmique
$\hat{\mathbf{f}}(\mathbf{a})$	: second membre polynomial du modèle d'ordre réduit
$p_c$	: nombre de paramètres calibrés pour la correction du modèle d'ordre réduit
$\mathbf{w}$	: vecteur des paramètres calibrés de dimension $p_c$

#### 5. Matrices et opérateurs

$\mathbf{I}_n, \mathbf{I}_{n,1}$	: matrice identité de dimension $n$ et vecteur de la diagonale de la matrice identité
$\mathbf{0}_n, \mathbf{0}_{n,m}$	: matrice nulle de dimension $n \times n$ et $n \times m$
$\delta_{ij}$	: symbole de Kronecker
$\langle \cdot, \cdot \rangle_H$	: produit scalaire entre deux vecteurs de l'espace de Hilbert $H$
$\  \cdot \ _H$	: norme induite par le produit scalaire sur l'espace de Hilbert $H$

#### 6. Indices et exposants

$\cdot_h$	: indice relatif à la discrétisation
$\hat{\cdot}$	: grandeurs relatives au système dynamique réduit
$\cdot^T$	: opérateur de transposition d'une matrice
$\cdot^+$	: pseudo-inverse d'une matrice
$\cdot_A, \cdot_E$	: grandeur vectorielle exprimée dans la base du repère absolu ou du repère entraîné

#### 7. Repères et descriptions cinématiques

$\Omega_M, \Omega_C, \Omega_A$	: domaine matériel, de contrôle et arbitraire pour les descriptions cinématiques
$\mathbf{X}, \mathbf{x}, \boldsymbol{\chi}$	: coordonnées matérielles, spatiales et arbitraires
$\mathcal{R}_A, \mathcal{R}_E$	: repère absolu et entraîné (relatif)
$\mathbf{s}$	: vitesse de déplacement des points du domaine arbitraire
$\mathbf{s}_e, \mathbf{s}_d$	: vitesse d'entraînement et vitesse de déformation
$\boldsymbol{\omega}$	: vecteur vitesse angulaire de rotation du repère entraîné
$\tilde{\mathbf{x}}$	: position relative d'un point dans le repère entraîné
$x, y, z, x_i$	: directions associées aux vecteurs de la base spatiale

#### 8. Variables aérodynamiques

$\rho$	: masse volumique du fluide
$\vartheta$	: covolume (volume spécifique)
$\mathbf{u}$	: vecteur vitesse absolue du fluide
$\mathbf{v}$	: vecteur vitesse relative du fluide
$\boldsymbol{\sigma}$	: tenseur des contraintes de Cauchy

---

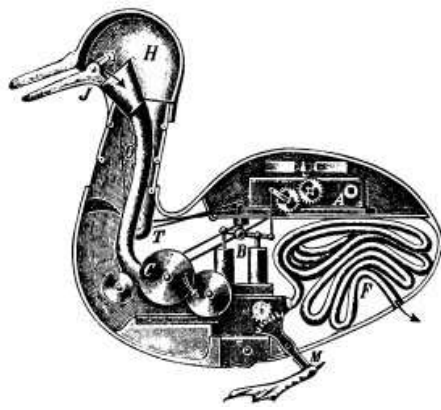
$\mathcal{E}_I, e_I$	: énergie interne et énergie interne volumique
$\mathcal{E}_C, e_C$	: énergie cinétique et énergie cinétique volumique
$\mathcal{E}_T, e_T$	: énergie totale et énergie totale volumique
$\mathbf{q}_\theta$	: vecteur flux de chaleur
$k_\theta$	: constante de conductivité thermique
$\theta$	: température
$p$	: pression
$\boldsymbol{\tau}$	: tenseur déviatorique (tenseur des contraintes visqueuses)
$\lambda, \mu$	: coefficients de viscosité du fluide
$\mathbf{D}$	: tenseur des taux de déformation
$R_s$	: constante spécifique des gaz parfaits
$c_p, c_v$	: coefficients de chaleur spécifique à pression et volume constant
$\gamma$	: rapport des coefficients de chaleur spécifique
Re, Pr, Ma	: nombres adimensionnels de Reynolds, de Prandtl et de Mach
$C_L, C_D$	: coefficients de portance et de traîné
$\mathbf{f}_{\text{Cor}}, \mathbf{f}_{\text{cen}}$	: forces volumiques de Coriolis et centrifuge

### 9. Abréviations

POD	: « Proper Orthogonal Decomposition »
SVD	: « Singular Value Decomposition »
MP	: Méthode des Pénalités
MML	: Méthode des Multiplicateurs de Lagrange
MFC	: Méthode des Fonctions de Contrôle
ALE	: « Arbitrary Lagrangian-Eulerian »
FALE-RAUA	: Formulation ALE des équations de Navier-Stokes dans le repère absolu avec les composantes de la vitesse absolue (exprimée dans le repère absolu)
FALE-REUE	: Formulation ALE des équations de Navier-Stokes dans le repère entraîné avec les composantes de la vitesse absolue (exprimée dans le repère entraîné)
FALE-REVE	: Formulation ALE des équations de Navier-Stokes dans le repère entraîné avec les composantes de la vitesse relative (exprimée dans le repère entraîné)

*Note* : Les vecteurs et les matrices seront représentés par des lettres grasses, avec des lettres minuscules pour les premiers et majuscules pour les seconds dans la mesure du possible. Par ailleurs, les quantités vectorielles seront notées en italique lorsque aucune référence à une base particulière n'est faite. Dans le cas contraire, la notation droite est employée avec éventuellement un indice précisant la base dans laquelle la quantité est exprimée.









---

# Introduction

---

Dans les moteurs d'avions ou plus généralement dans les turbomachines, certains éléments du compresseur et de la turbine sont confrontés à des problèmes de stabilité vibratoire importants qui limitent leur domaine de fonctionnement. Il est donc indispensable de prédire d'une manière ou d'une autre les frontières du domaine de stabilité de ces pièces mécaniques regroupées sous l'appellation générique de roues aubagées. La prédiction des caractéristiques vibratoires se fait généralement soit au moyen d'essais expérimentaux onéreux, soit par ordinateur à l'aide de codes de calcul numériques.

La réalisation d'essais expérimentaux n'est toutefois pas envisageable au cours de la phase de conception et le recours aux méthodes numériques s'impose donc comme une méthode alternative inévitable. L'objectif principal est d'évaluer le plus précisément possible les fréquences de vibration et surtout l'amortissement qui caractérisent la stabilité du système mécanique, afin de réduire les marges de sécurité introduites pour assurer le fonctionnement du système dans un domaine stable.

La principale difficulté pour l'estimation précise des caractéristiques vibratoires est la nécessité de travailler avec des modèles physiques très élaborés puisque l'interaction entre l'écoulement aérodynamique à travers les étages de la turbomachine et la structure déformable des aubes donne naissance à des phénomènes complexes qualifiés d'*aéroélastiques* puisqu'ils résultent du couplage entre les deux milieux. Les simulations numériques doivent alors être capables de reproduire le couplage entre le mouvement de la structure élastique et l'écoulement exerçant les charges aérodynamiques sur la paroi des aubes.

La tâche n'est cependant pas aisée car — même en considérant une structure rigide — l'écoulement est le siège de phénomènes aérodynamiques particulièrement difficiles à modéliser en raison des vitesses et des pressions extrêmes qui règnent, ainsi que du mouvement relatif des différents étages de la turbomachine. Le développement de modèles extrêmement élaborés du fluide a néanmoins permis de reproduire des phénomènes impliquant le caractère compressible, visqueux et turbulent des écoulements dans les turbomachines où les aspects non-linéaires jouent un rôle important [40, 133]. Simultanément, l'augmentation progressive de la puissance des processeurs et leur parallélisation dans les supercalculateurs a rendu possible l'obtention de solutions numériques, bien que les coûts de calcul pour y parvenir restent encore la plupart du temps considérables.

Quoique réalisables, les simulations *aéroélastiques* dans un système aussi compliqué qu'un (voire plusieurs) étage(s) de turbomachine deviennent très coûteuses en temps de calcul. Le coût de calcul est surtout tributaire de la modélisation du fluide puisque la structure peut être — en première approche — raisonnablement modélisée par un matériau élastique linéaire. En revanche, le modèle du fluide doit tenir compte du caractère non-linéaire, compressible, visqueux et turbulent de l'écoulement pour espérer obtenir une estimation précise des caractéristiques vibratoires du système *aéroélastique*. Bien que les simulations soient une solution attractive, les coûts de calcul rédhibitoires doivent être réduits afin que l'outil de prédiction devienne utilisable en pratique dans une phase de conception. La réduction du coût de calcul est d'autant plus indispensable qu'un nombre important de simulations doit être réalisé afin d'obtenir une cartographie du domaine de stabilité des turbomachines sur une certaine gamme de paramètres. En effet, le domaine de fonctionnement dépend en particulier du nombre de Mach, de la vitesse de rotation des roues aubagées, des déphasages inter-aubes,... Toutes ces combinaisons de paramètres entraînent par conséquent un nombre considérable de simulations numériques qu'il n'est pas possible de réaliser en pratique avec précision.

Un moyen simple de réduire le temps de calcul consiste à employer pour le fluide des modèles plus simples, pour lesquels certaines hypothèses restrictives sont émises. Ainsi, la turbulence, voire même le caractère visqueux du fluide peut être négligé. En *aéroélasticité*, il est même encore courant de recourir aux modèles théoriques de surfaces portantes établis pour un fluide parfait, irrotationnel supportant des petites perturbations isentropiques. L'emploi de ces modèles dits *simplifiés* est néanmoins risqué et conduit parfois

à des réponses de mauvaise qualité puisque la prédiction de l'amortissement du système aéroélastique requiert la modélisation précise de la dissipation dans le fluide. Une seconde classe de méthodes portant le nom de *modèles d'ordre réduit*, parfois abrégés en ROM pour l'acronyme anglais « Reduced Order Model », a donc vu le jour. Le principe consiste à construire un nouveau modèle des équations du fluide comportant un très petit nombre de degrés de liberté. Ce modèle est en général élaboré au moyen d'une projection des équations du modèle complexe sur une base réduite, ou encore par l'identification des coefficients d'un système dynamique réduit à partir des réponses du modèle non réduit. L'objectif est de préserver la physique des phénomènes complexes dans le système dynamique réduit afin que les réponses produites restent de bonne qualité malgré le très faible nombre de degrés de liberté.

Par le passé, un certain nombre de travaux ont été menés pour tenter de construire des modèles d'ordre réduit des équations du fluide. Karpel [127, 128] a ainsi développé une méthode de lissage des forces aérodynamiques généralisées basée sur l'interpolation de la réponse évaluée en quelques points de fonctionnement. Cette méthode a l'avantage de pouvoir être adaptée aux différents modèles de fluides gouvernés par des équations plus ou moins complexes. Toutefois, cette méthode ne s'applique que dans le cadre d'un couplage aéroélastique faible, *i. e.* pour un couplage dans lequel un mouvement harmonique de la structure est imposé. Par ailleurs, la méthode requiert le réglage assez empirique d'un certain nombre de paramètres spécifiques au type d'écoulement modélisé et à la configuration étudiée. Plutôt que de chercher à interpoler la réponse à partir d'un certain nombre de points dans le domaine de fonctionnement, certains auteurs ont réellement cherché à construire un modèle d'ordre réduit du fluide en projetant les équations sur une base réduite comportant un faible nombre de vecteurs. Lumley [156] avait ainsi obtenu un modèle d'ordre réduit capable de reproduire les structures turbulentes et de capter ainsi la transition vers la turbulence. La base de projection employée était construite à l'aide de la *décomposition orthogonale propre* d'un ensemble de solutions calculées au préalable avec le modèle non réduit. Ce type de modèle d'ordre réduit est rapidement devenu populaire puisque des phénomènes particulièrement complexes peuvent être reproduits à très faible coût. Un des avantages primordial de la méthode est le fait que les non-linéarités peuvent être préservées.

Concernant les écoulements dans les turbomachines, les modèles d'ordre réduit construits au moyen de la décomposition orthogonale propre ont principalement été formulés pour les équations linéarisées du fluide ce qui facilite la mise en œuvre et permet d'étudier le système dans le domaine fréquentiel. C'est ainsi qu'ont procédé Hall *et al.* [104] et Epureanu *et al.* [68] par exemple. Plus récemment, Amsallem *et al.* [10] ont repris le même type de modèles et ont introduit la notion d'interpolation entre les bases de projection pour obtenir la réponse du système sur une certaine gamme de paramètres dont le nombre de Mach notamment. Dans un registre assez proche, Florea et Hall [78] ont tenté de construire de manière similaire un modèle d'ordre réduit par projection des équations du modèle non-réduit, mais en utilisant cette fois une base de modes propres calculés pour les équations linéarisées du fluide. La méthode, largement employée en mécanique des structures, a ici été étendue au cas des fluides, mais l'extraction des modes propres devient rapidement impossible lorsque le nombre de degrés de liberté du modèle initial devient important. Enfin, la méthode de la *balance harmonique* employée notamment par Hall *et al.* [103] et Thomas *et al.* [227] consiste à supposer la solution périodique et à formuler le système d'équations pour les premiers coefficients de la décomposition de Fourier. L'inconvénient majeur est la nécessité de faire l'hypothèse d'une réponse harmonique. Une bonne revue et une comparaison de différentes méthodes de réduction pour les écoulements dans les turbomachines ont été réalisées par Dowell *et al.* [60].

Les différents travaux menés dans le cadre de la réduction de modèle pour l'écoulement dans les turbomachines ont permis de décrire des phénomènes aérodynamiques de plus en plus complexes, tout en conservant un coût de calcul très faible. Néanmoins, les phénomènes qualifiés de « dynamiquement non-linéaires » par Dowell *et al.* [60] restent à l'heure actuelle impossibles à reproduire puisque la quasi-totalité des modèles d'ordre réduit pour les écoulements dans les turbomachines est formulée pour les équations linéarisées. La décomposition orthogonale propre possède toutefois le potentiel nécessaire pour la construction d'un modèle d'ordre réduit non-linéaire, d'autant plus qu'elle ne repose pas sur l'hypothèse d'un écoulement harmonique comme la méthode de balance harmonique. En dehors du cadre de l'aéroélasticité des turbomachines quelques rares travaux ont été réalisés à l'aide de la décomposition orthogonale propre dans le but de construire un modèle d'ordre réduit valable pour les fluides compressibles d'une part, ou capable de reproduire des phénomènes issus du couplage fluide-structure d'autre part. Citons ainsi la formulation non-linéaire du modèle d'ordre réduit développé par Vigo [240] à l'aide des *variables primitives modifiées*

(covolume, composantes de la vitesse et pression) pour les écoulements compressibles, ou encore le modèle de Liberge [145] construit sur un domaine fictif comprenant à la fois le fluide et la structure afin de pouvoir traiter correctement le problème du couplage fluide-structure.

Ces deux derniers modèles présentent chacun un intérêt substantiel pour la tâche qui nous incombe, bien qu'aucune des formulations ne réponde totalement à notre problématique. Il s'agit en effet dans le cadre de cette thèse de parvenir à construire un modèle d'ordre réduit non-linéaire valable pour un écoulement compressible, visqueux et instationnaire et capable de tenir compte du mouvement d'une structure à l'origine de phénomènes aéroélastiques. Pour cela, on se propose de s'inspirer de la formulation développée par Vigo [240]. L'originalité du travail réalisé ici réside dans les méthodes de correction et d'identification employées pour stabiliser le modèle d'ordre réduit, et surtout dans l'extension des équations pour la prise en compte du mouvement d'une structure. Les développements présentés dans ce mémoire concernent un profil isolé mais la méthodologie reste valable pour traiter les cascades de profils caractéristiques des roues aubagées présentes dans les turbomachines.

La démarche suivie ici se divise en deux grandes parties : la première concerne le choix de la méthode de réduction et son application à un système dynamique linéaire, tandis que la seconde concerne l'extension à un système non-linéaire représentatif de l'écoulement d'un fluide compressible visqueux autour d'une structure animée d'un mouvement de corps rigide. Le premier chapitre constitue un état de l'art des différentes méthodes de réduction envisageables pour construire le modèle de substitution. À l'issue de ce chapitre, la décomposition orthogonale propre est choisie de manière objective selon des critères relatifs aux phénomènes devant pouvoir être reproduits et les principales propriétés de la méthode sont rappelées dans le chapitre 2. La première partie s'achève dans le chapitre 3 par l'évaluation de la formulation POD-Galerkin qui est d'abord comparée à une formulation analytique puis discrète du modèle d'ordre réduit dans le cas d'un système linéaire autonome. Trois techniques sont alors développées pour tenir compte avec l'approche POD-Galerkin d'une condition limite non-homogène de type Dirichlet. Le modèle d'ordre réduit non-autonome obtenu dans ce cas représente alors l'archétype du problème d'aéroélasticité dans lequel l'instationnarité est provoquée par une condition limite de ce type. Les travaux du chapitre 3 sont résumés dans les actes du congrès ISMA 08 [180] ainsi que dans l'article [181] publié dans le *Journal of Sound and Vibration*. Dans le chapitre 4 de la seconde partie, un rappel des équations de Navier-Stokes écrites dans un repère absolu puis entraîné avec les variables conservatives pour un fluide supposé compressible et visqueux sera effectué. La formulation de Vigo [240] écrite à l'aide des variables primitives modifiées sera alors étendue afin de tenir compte du mouvement d'une structure dans un repère entraîné. Les équations ainsi obtenues serviront par la suite pour la projection de Galerkin intervenant dans la construction du modèle d'ordre réduit. Un premier modèle d'ordre réduit non-linéaire qualifié d'*aérodynamique* valable dans un domaine de frontières fixes est alors développé au chapitre 5 pour reproduire le phénomène de détachement de vortex dans le sillage d'un profil NACA0012 fixe fortement incliné. Un certain nombre de méthodes de correction seront alors mises en œuvre pour tenter de reproduire au mieux la réponse du système non-réduit. Enfin, un dernier modèle d'ordre réduit qualifié d'*aéroélastique* par abus de langage est construit au chapitre 6 pour tenir compte du mouvement d'une structure supposée rigide. Celui-ci est mis en œuvre pour reproduire l'écoulement autour d'un profil NACA0064 auquel un mouvement d'oscillation harmonique est imposé, ce qui provoque l'oscillation du choc sur les parois. Une méthode d'identification des coefficients du modèle d'ordre réduit est alors employée pour reproduire correctement la réponse et réduire le coût de construction du modèle d'ordre réduit. La formulation ainsi que quelques résultats relatifs aux chapitres 5 et 6 sont résumés dans les actes du congrès IFASD 09 [182]. Cette dernière étape ouvre donc la voie à l'utilisation du modèle d'ordre réduit aéroélastique pour reproduire la réponse du système sur une certaine gamme de fréquence ou d'amplitude de mouvement de la structure.



## **Première partie**

# **Choix du modèle d'ordre réduit et application à un système dynamique linéaire**



# État de l'art des méthodes de réduction de modèle pour l'aérodynamique

---

## Sommaire

---

<b>1.1</b>	<b>Introduction</b> . . . . .	<b>8</b>
<b>1.2</b>	<b>Principe de la réduction de modèles et classification</b> . . . . .	<b>9</b>
<b>1.3</b>	<b>Description des principaux modèles d'ordre réduit physiques</b> . . . . .	<b>13</b>
1.3.1	Projection sur la base des modes propres et sous-structuration dynamique . . . . .	13
1.3.2	Approximation de la fonction de transfert . . . . .	15
1.3.2.1	Balanced Truncation . . . . .	16
1.3.2.2	Méthodes d'approximation de Krylov . . . . .	17
1.3.3	Identification du système réduit . . . . .	18
1.3.4	Décomposition Orthogonale Propre . . . . .	21
1.3.5	Modes d'oscillation et d'interaction principaux . . . . .	25
1.3.6	Équilibrage harmonique . . . . .	27
1.3.7	Réduction sur une base construite <i>a priori</i> . . . . .	28
1.3.8	Méthodes de réduction alternatives et dépendance avec un paramètre . . . . .	29
<b>1.4</b>	<b>Synthèse et justification du choix de la méthode POD</b> . . . . .	<b>32</b>

---



## 1.1 Introduction

Dans ce chapitre, on considère un système dynamique représenté par un ensemble d'équations dont la solution, caractérisée par un certain nombre de variables indépendantes, fournit une description du phénomène étudié. Le système dynamique correspond aux équations du fluide dans le cadre de cette thèse, mais peut, d'une manière plus générale, traduire le comportement d'un système mécanique ou physique quelconque. Pour commencer, le principe de la réduction de modèle et les différentes approches envisageables sont rappelées d'un point de vue formel afin d'introduire la classification employée par la suite pour présenter les différentes catégories. On choisit ici de distinguer les modèles d'ordre réduit de type *surface de réponse*, dont la construction est indépendante des équations gouvernant le système dynamique, et les modèles d'ordre réduit *physiques*, qui, comme leur nom l'indique, tentent de conserver les propriétés physiques en utilisant les équations du système dynamique. À l'intérieur de ces deux catégories, plusieurs classes de modèles d'ordre réduit peuvent encore être isolées. Dans tous les cas, le principe de la méthode de réduction est brièvement décrit et les principaux avantages et inconvénients sont signalés. À l'issue de cet état de l'art des méthodes de réduction de modèles, un choix objectif de la méthode employée sera effectué dans l'optique de la construction d'un modèle d'ordre réduit non-linéaire capable de reproduire l'écoulement d'un fluide compressible.

## 1.2 Principe de la réduction de modèles et classification

La modélisation de phénomènes physiques au moyen de simulations numériques requiert la plupart du temps la résolution d'un *système dynamique*, pour peu que le phénomène étudié soit instationnaire. On supposera ici que ce système est composé d'un ensemble d'équations aux dérivées partielles ou ordinaires explicites, dont la résolution fournit l'état  $x$  du système en réponse à une *entrée*  $\mu$  supposée paramétrée par  $p$  variables réelles. Le système dynamique peut alors s'écrire sous la forme <sup>1</sup>

$$(\Sigma) : \dot{x} = f(x, \mu). \quad (1.1)$$

L'état  $x$  du système est supposé évoluer dans un espace de dimension infinie  $\mathcal{V}$  et sur un intervalle de temps fini  $I = [t_0; t_0 + t_f] \subset \mathbb{R}_+$  où  $t_0$  est l'instant initial et  $t_f$  l'instant final de la simulation. L'opérateur  $f$  est une application de  $\mathcal{V} \times \mathbb{R}^p$  dans  $\mathcal{W}$ , et décrit la réponse du système au moyen d'opérateurs spatiaux éventuellement différentiels. Pour un grand nombre d'applications en mécanique, l'espace des fonctions de carré intégrable  $\mathcal{V} = L^2(\Omega)$  ou encore l'espace de Sobolev  $H^1(\Omega)$  d'ordre 1 fournissent un cadre idéal pour la description des milieux continus dans un domaine spatial  $\Omega \subset \mathbb{R}^d$ ,  $d = 1, 2$  ou  $3$ . Bien souvent, l'état  $x$  du système est décrit par  $n_v$  variables indépendantes pouvant appartenir chacune à des espaces  $\mathcal{V}$  différents.

La connaissance de l'ensemble des solutions de l'équation (1.1) définit le *flot* associé au système dynamique. En pratique on cherche plutôt une solution particulière associée à une condition initiale  $x_0 = x(t = t_0)$  pour laquelle le problème de Cauchy possède une solution unique. La détermination de l'état  $x$  du système n'est généralement pas une fin en soi et sert seulement d'intermédiaire pour en déduire une *sortie*  $y \in \mathcal{V}_y$  qui correspond à une grandeur d'intérêt souvent macroscopique, telle que  $y = g(x, \mu)$ . Dans le cadre de cette thèse, le système dynamique (1.1) représente les équations du fluide et l'état  $x$  correspond, par exemple, aux  $n_v$  variables conservatives tandis que la sortie  $y$  représente les forces aérodynamiques s'exerçant sur la paroi d'un profil, ou encore la fréquence ou l'amortissement de ces forces. Le paramètre  $\mu$  peut correspondre à certaines propriétés physiques du fluide, à certaines conditions limites, à un terme source ou encore aux paramètres géométriques du domaine spatial sur lequel les équations sont résolues.

En dehors des cas académiques, la résolution analytique du système  $(\Sigma)$  est impossible et c'est la raison pour laquelle les simulations numériques reposent sur une approximation du système d'équations basée sur la *discrétisation* du domaine de calcul  $\Omega$ . Le principe consiste alors à rechercher une solution approchée  $x_h$  dans un nouvel espace  $\mathcal{V}^h \subset \mathcal{V}$  de dimension finie  $N_v = \mathcal{O}(n_v/h)$  où  $h$  est le paramètre de discrétisation spatiale. L'application discrète  $f_h$  est alors définie de  $\mathcal{V}^h \times \mathbb{R}^p$  dans  $\mathcal{W}^h$ . La sortie  $y$  du système est éventuellement discrétisée afin d'être exprimée à partir de l'état discret  $x_h$  : dans ce cas, la sortie approchée  $y_h \in \mathcal{V}_y^h \subset \mathcal{V}_y$  découle de la relation approchée  $y_h = g_h(x_h, \mu)$ . La solution approchée  $x_h$  du système dynamique discrétisé correspond alors à la solution du problème approché

$$(\Sigma_h) : \dot{x}_h = f_h(x_h, \mu). \quad (1.2)$$

Les systèmes dynamiques rencontrés en mécanique sont couramment discrétisés par des méthodes Éléments Finis, Volumes Finis, ou encore des méthodes spectrales et sont capables de fournir une description d'autant plus précise et fine que le pas de discrétisation  $h$  est petit. Le nombre  $N_v$  des *degrés de liberté*, *i. e.* les composantes du vecteur  $x_h$  dans l'espace discret  $\mathcal{V}_h$ , augmente par conséquent lorsque le pas de discrétisation diminue. Lorsque le domaine spatial  $\Omega$  est discrétisé en  $N$  éléments et que l'état du système est décrit par  $n_v$  variables évaluées en un unique point par élément de discrétisation, le nombre de degrés de liberté est  $N_v = N \times n_v$ .

Ainsi, lorsque les phénomènes sont régis par un système dynamique complexe ou sont caractérisés par des échelles de temps (ou d'espace) très différentes, la simulation numérique devient très coûteuse en temps de calcul et seul un nombre très limité de calculs peut être mené. On parlera alors de *modèles d'ordre élevé* pour désigner les modèles caractérisés par un nombre important de degrés de liberté nécessaires pour accéder aux détails de la réponse du système. Les équations de Navier-Stokes peuvent ainsi être considérées comme un modèle d'ordre élevé, notamment lorsque l'écoulement d'un fluide turbulent en régime transsonique est étudié. De ce fait, bien que des modèles numériques précis existent et soient en mesure de

<sup>1</sup>Les systèmes d'ordre deux ou plus, *i. e.* impliquant des dérivées d'ordre supérieur à un en temps, peuvent se ramener à un système d'ordre un en temps en introduisant simplement une variable supplémentaire telle que si le système est  $\ddot{x} = f(x, \dot{x}, \mu)$  par exemple, alors le changement de variable  $z = \dot{x}$  permet de transformer le système en  $\dot{z} = f_z(x, z, \mu)$  et le système dynamique initial d'ordre deux en temps est équivalent à un système de deux équations d'ordre un en temps.

fournir la solution détaillée du système dynamique, leur emploi reste généralement confiné à la validation du comportement du système dynamique sur un nombre très restreint de cas d'étude.

Par conséquent, il est nécessaire de travailler avec des modèles numériques plus légers comportant un nombre de degrés de liberté moins important. L'idée n'est pas récente et un exemple célèbre est le système dynamique de Lorenz [151] qui tente de décrire le couplage entre l'atmosphère et l'océan par un système à trois degrés de liberté. Dans ce cas, le développement d'un tel modèle se justifiait par l'impossibilité de résoudre numériquement des systèmes d'équations de taille importante. Le recours à ce type de réduction reste néanmoins d'actualité lorsque l'objectif est de disposer de la réponse du système pour un grand nombre de paramètres comme c'est le cas au cours des procédures d'optimisation ou encore lorsqu'il est nécessaire d'accéder à la réponse en temps réel pour alimenter les boucles de contrôle actif introduites dans les systèmes asservis. Un autre intérêt de la réduction de modèle concerne l'analyse des systèmes complexes. En effet, les modèles d'ordre élevé produisent des solutions extrêmement détaillées contenant une masse d'information considérable dont il est parfois difficile d'extraire les informations pertinentes. Des méthodes issues par exemple du domaine de la statistique permettent alors de mettre en évidence les structures intéressantes et d'analyser ainsi plus facilement le système. Ce type d'approche s'inscrit cependant plutôt dans le registre de la compression de données car aucun système dynamique réduit n'est réellement formé. Toutefois les structures identifiées peuvent ensuite servir dans certains cas à la construction d'un modèle d'ordre réduit et c'est la raison pour laquelle on mentionne ici ce type d'approche.

À ce stade, il faut éclaircir la notion de réduction de modèle puisqu'il existe une grande diversité dans la manière de procéder. La première distinction opère au niveau du choix des hypothèses émises pour la construction du modèle d'ordre réduit :

- soit le modèle d'ordre élevé est abandonné au profit d'un modèle mathématique *simplifié* par des hypothèses plus restrictives ;
- soit le modèle d'ordre élevé sert de base à la construction d'un *modèle d'ordre réduit* capable de retranscrire avec un niveau de détail pratiquement équivalent la réponse du système.

Dans le premier cas, la réduction du nombre de degrés de liberté provient par exemple d'hypothèses simplificatrices : ainsi, si le problème est de simuler l'écoulement au travers d'une roue aubagée, l'hypothèse d'un fluide parfait au lieu d'un fluide visqueux permet de réduire le nombre de degrés de liberté dans la couche limite, de même que l'hypothèse de symétrie cyclique permet de réduire le domaine spatial à un seul secteur compris entre deux aubes successives. La réduction du nombre de degrés de liberté s'obtient néanmoins au prix d'hypothèses de base qui deviennent parfois trop restrictives pour que la réponse obtenue soit satisfaisante. Dans ce cas il ne s'agit pas à proprement parler d'une réduction de modèle mais plutôt d'un *changement* de modèle qui s'accompagne d'une perte au niveau des phénomènes modélisés. Les deux hypothèses proposées par exemple dans le cas du fluide conduisent dans le premier cas à négliger les effets visqueux, et dans le second cas à supposer que les aubes sont accordées.

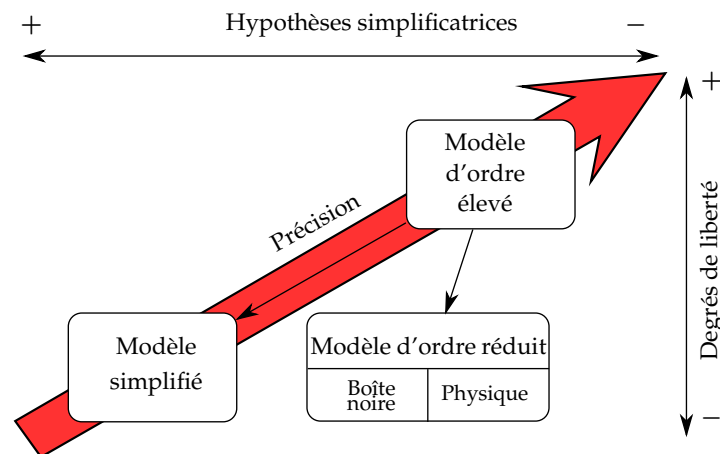
Dans le second cas, l'idée consiste à élaborer un nouveau modèle reposant sur les mêmes hypothèses de sorte que les mêmes phénomènes puissent a priori être reproduits. La plupart de ces modèles tirent parti d'un certain nombre de réponses calculées avec le modèle d'ordre élevé ou obtenues expérimentalement et qui constituent en quelque sorte une base de données représentative du système. Puisque le modèle d'ordre élevé a été employé pour obtenir les réponses, les informations contenues dans la base de données présentent un niveau de finesse élevé. Par conséquent, l'utilisation intelligente de ces données devrait pouvoir donner naissance à un modèle d'ordre réduit ne contenant qu'un faible nombre de degrés de liberté mais capable de reproduire la réponse du système avec un niveau de détails pratiquement équivalent. Ce procédé de construction présente néanmoins deux inconvénients majeurs : tout d'abord il est nécessaire de générer une base de données représentative, et ensuite le modèle d'ordre réduit obtenu est dépendant des données de par sa construction.

Le contexte de cette thèse conduit naturellement vers le développement d'un modèle d'ordre réduit puisque notre objectif est ici d'obtenir un modèle représentatif d'un écoulement complexe dans lequel les phénomènes de couplage aéroélastique ne peuvent pas être correctement captés par un modèle simplifié. Le terme de modèle d'ordre réduit regroupe encore un grand nombre de méthodes qui peuvent être classées en deux grandes catégories.

La première d'entre elles consiste à reconstruire la sortie du modèle d'ordre élevé sans information physique particulière en dehors d'un ensemble d'échantillons censé être représentatif du comportement du

système dynamique. Le modèle d'ordre élevé est donc supposé pouvoir être décrit par une fonction cible telle que la sortie  $y$  s'exprime directement en fonction du paramètre  $\mu$  sans qu'aucun système dynamique réduit ne soit construit. C'est pourquoi ces modèles sont qualifiés de *boîtes noires* ou de *modèles de type surface de réponse* puisque la fonction présupposée n'est pas directement fonction de l'état  $x$ . La plupart du temps, la détermination de la fonction cible s'effectue au moyen de processus **régressifs**. Parmi les méthodes couramment employées, on peut citer les méthodes d'estimation linéaire stochastique [1], les réseaux de neurones [61], les fonctions à base radiale, la régression par vecteurs supports [235] ou encore le Kriging [134, 163]. Une revue assez intéressante des méthodes les plus répandues pour les applications en optimisation est fournie par Forrester et Keane [79]. On ne développera pas plus en détails ces méthodes pour les raisons suivantes : tout d'abord, il paraît difficile de pouvoir reproduire avec de tels modèles le comportement instationnaire d'un système dynamique non-linéaire étant donné qu'aucune information sur la physique du problème n'est prise en compte dans la construction du modèle d'ordre réduit ; par ailleurs, l'objectif final est de parvenir à un système dynamique réduit qui puisse être couplé aux équations de la structure afin de former un modèle d'ordre réduit aéroélastique, et le couplage risque d'être difficile à mettre en œuvre si le modèle d'ordre réduit ne fournit que la sortie  $y$  du système dynamique sans information sur l'état  $x$ .

La seconde catégorie de méthodes repose, pour la grande majorité, sur des techniques de projection des équations du modèle d'ordre élevé sur une base de dimension réduite convenablement choisie. Par opposition aux méthodes précédentes, on parle de *modèles d'ordre réduit physiques* puisque les équations du système réduit sont souvent obtenues par des processus **projectifs**. Le système réduit conserve alors la même forme que celle des équations du modèle d'ordre élevé si bien qu'une partie des propriétés du système initial est conservée. Les méthodes les plus répandues sont les méthodes de sous-structuration dynamique, les méthodes d'identification de la fonction de transfert, les méthodes de projection sur la base fournie par une décomposition orthogonale propre, ou encore les méthodes de réduction sur une base construite a priori. Quelques unes de ces méthodes seront brièvement présentées dans le paragraphe suivant.



**Figure 1.1 :** Schématisation des différentes familles de modèles d'ordre réduit en fonction des hypothèses émises sur le système dynamique et du nombre de degrés de liberté impliqués.

Le classement des différentes méthodes de réduction est schématisé sur la figure 1.1 en fonction des hypothèses émises sur les modèles employés et du nombre de degrés de liberté impliqués. Les modèles d'ordre réduit ont été placés dans une zone correspondant à des hypothèses légèrement plus fortes que celles du modèle d'ordre élevé car, même si il ne s'agit pas d'une règle générale, il est courant d'émettre quelques simplifications pour faciliter la construction du modèle d'ordre réduit.

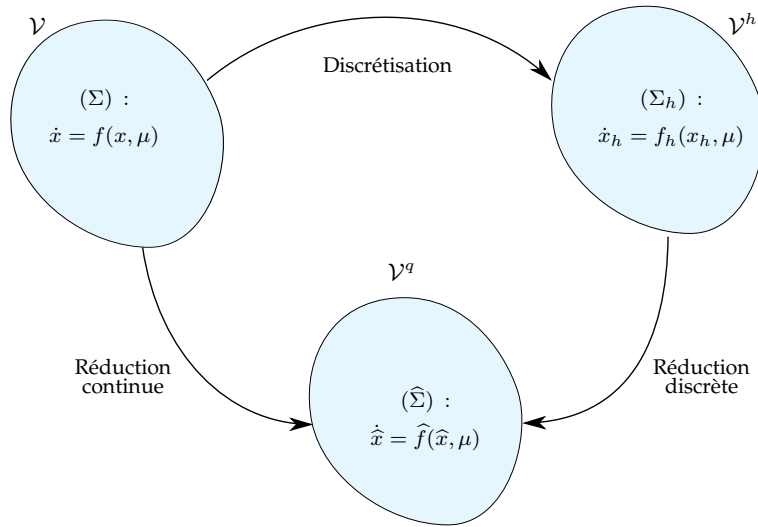
La construction d'un modèle d'ordre réduit physique s'effectue généralement en deux temps : la première étape consiste à générer la base de données représentative du système dynamique<sup>2</sup> puis à construire la base de projection, et la seconde étape est dédiée à la projection des équations du modèle d'ordre élevé sur cette base de dimension réduite.

<sup>2</sup>La production d'une base de données n'est pas toujours indispensable en fonction de la méthode de réduction adoptée.

Une fois la base de projection déterminée, le modèle d'ordre réduit est construit à partir du système dynamique continu  $(\Sigma)$  ou discrétisé  $(\Sigma_h)$  de telle manière qu'il soit gouverné par un très petit nombre de degrés de liberté. Il s'agit alors de reformuler le système afin de déterminer l'état approché  $\hat{x}$  dans un espace  $\mathcal{V}^q$  de dimension finie  $q$  telle que  $q \ll N_v$ . L'application  $f$  est alors une nouvelle fois redéfinie sur  $\mathcal{V}^q \times \mathbb{R}^p$  à valeurs dans  $\mathcal{W}^q$  et le système dynamique réduit est

$$(\hat{\Sigma}) : \dot{\hat{x}} = \hat{f}(\hat{x}, \mu). \quad (1.3)$$

La relation qui définit la sortie  $y$  en fonction de l'état  $x$  et de l'entrée  $\mu$  peut aussi être redéfinie sur des espaces de dimension réduite mais cela ne présente qu'un intérêt relatif puisque dans la majorité des cas c'est la résolution du système dynamique qui est coûteuse, et non la détermination de la grandeur d'intérêt  $y$ . Toutefois si cela est nécessaire il est toujours possible de définir une nouvelle relation  $\hat{y} = \hat{g}(\hat{x}, \mu)$ .



**Figure 1.2** : Schéma de principe de la réduction de modèle : le système réduit  $(\hat{\Sigma})$  défini dans l'espace  $\mathcal{V}^q$  se déduit soit de la formulation continue du système  $(\Sigma)$  défini dans  $\mathcal{V}$ , soit de la formulation discrète du système  $(\Sigma_h)$  défini dans  $\mathcal{V}^h$ .

Le principe de la réduction du système  $(\Sigma)$  est schématisé sur la figure 1.2 où le système discrétisé  $(\Sigma_h)$  ainsi que le système réduit  $(\hat{\Sigma})$  apparaissent. Ce dernier peut s'obtenir soit à partir de la formulation continue des équations du système, soit directement à partir de la formulation discrétisée.

Dans les paragraphes suivants, les méthodes de réduction physiques les plus courantes sont présentées brièvement en commençant par décrire celles dévolues aux systèmes linéaires puis celles applicables à n'importe quel type de système dynamique. L'accent est mis sur les applications en aérodynamique puisque notre objectif est de réaliser un modèle d'ordre réduit pour les systèmes aéroélastiques dans lesquels la simulation de la partie structure est peu pénalisante en terme de coût de calcul.

## 1.3 Description des principaux modèles d'ordre réduit physiques

L'état de l'art des méthodes de réduction de modèle est ici restreint à la description des principaux modèles dits *physiques* qui ont déjà été employés par le passé en mécanique des fluides ou qui présentent un certain potentiel pour ce type d'applications. Il existe une grande diversité de modèles d'ordre réduit physiques dont la plupart a à l'origine été développée pour des applications ou dans des domaines spécifiques, ou qui utilisent des propriétés intrinsèques au système dynamique — telles que la linéarité, l'existence d'une base modale associée,... La classification des modèles d'ordre réduit physiques est donc une tâche délicate puisque chacune possède ses propres caractéristiques. Dowell *et al.* [60] ont ainsi choisi de distinguer les méthodes basées sur les caractéristiques modales du système, de celles utilisant la notion de fonction de transfert aérodynamique. Antoulas [11] s'est intéressé quant à lui aux systèmes linéaires et a préféré opposer les méthodes basées sur la décomposition en valeurs singulières aux méthodes utilisant les sous-espaces de Krylov pour des questions de conservation de la stabilité des systèmes. Les méthodes sont ici présentées dans un ordre progressif, permettant de réduire des systèmes de plus en plus complexes en partant de systèmes linéaires pour aboutir à des systèmes non-linéaires.

### 1.3.1 Projection sur la base des modes propres et sous-structuration dynamique

En mécanique des structures, les *modes propres* d'un système élastique linéaire fournissent une base idéale pour la projection. Le principe consiste à chercher une solution particulière du système discret pour de petites oscillations libres autour d'une position d'équilibre. Pour cela, on suppose que tous les degrés de liberté suivent à un facteur près la même loi temporelle. En notant  $\mathbf{x} \in \mathbb{R}^{N_v}$  le vecteur des degrés de liberté du système dynamique, la solution particulière s'écrit

$$\mathbf{x} = \tilde{\boldsymbol{\psi}} a(t), \quad (1.4)$$

où le vecteur  $\tilde{\boldsymbol{\psi}}$  est indépendant du temps et  $a(t)$  correspond à la loi temporelle. Les équations de Lagrange pour le mouvement libre autour d'une position d'équilibre stable s'écrivent sous la forme d'un système d'équations différentielles ordinaires linéaires qui portent sur les degrés de liberté  $\mathbf{x}$ . Ces équations sont regroupées sous la forme matricielle

$$\mathbf{M} \ddot{\mathbf{x}} + \mathbf{K} \mathbf{x} = \mathbf{0} \quad (1.5)$$

dans laquelle les matrices  $\mathbf{M}$  et  $\mathbf{K}$  de  $\mathcal{M}(N_v, N_v)$  correspondent respectivement à la matrice de masse et de raideur de la structure. Le remplacement de la relation (1.4) dans le système d'équations (1.5) donne :  $\mathbf{M} \tilde{\boldsymbol{\psi}} \ddot{a}(t) + \mathbf{K} \tilde{\boldsymbol{\psi}} a(t) = \mathbf{0}$ . Lorsque le système a une position d'équilibre stable, les matrices  $\mathbf{M}$  et  $\mathbf{K}$  sont non singulières. Le système d'équations précédent s'écrit aussi  $\mathbf{K} \tilde{\boldsymbol{\psi}} = \lambda \mathbf{M} \tilde{\boldsymbol{\psi}}$  en posant  $\lambda = -\ddot{a}(t)/a(t)$ . On peut démontrer assez simplement [83, chap. 2] que le facteur  $\lambda$  est réel et strictement positif, de sorte qu'il est transformé en  $\lambda = \omega^2$ . Le système prend finalement la forme  $(\mathbf{K} - \omega^2 \mathbf{M}) \tilde{\boldsymbol{\psi}} = \mathbf{0}$  qui comprend  $N_v$  équations linéaires et homogènes. Les modes propres du système sont alors définis comme les solutions non triviales du problème aux valeurs propres

$$(\mathbf{K} - \omega_j^2 \mathbf{M}) \boldsymbol{\psi}^{(j)} = \mathbf{0}, \quad (1.6)$$

tel que  $\omega_j^2$  soit racine de l'équation algébrique  $\det(\mathbf{K} - \omega_j^2 \mathbf{M}) = 0$ . Puisque les valeurs propres  $\omega_j^2$  sont réelles et positives, les vecteurs propres associés qui portent le nom de modes propres sont aussi réels.

D'après la définition de  $\lambda$ , la loi d'évolution temporelle pour chaque mode propre s'écrit  $\ddot{a}_j(t) + \omega_j^2 a_j(t) = 0$  et la solution est harmonique de pulsation  $\omega_j$ . La détermination des modes propres de la structure est d'abord intéressante puisqu'elle permet de décrire le comportement fréquentiel de la structure et d'identifier les fréquences susceptibles de faire entrer le système en résonance en fonction des sollicitations imposées. Cependant, le calcul des modes propres est aussi à la base de la construction de méthodes de réduction de modèles. En effet, les vecteurs propres associés au problème (1.6) peuvent être regroupés pour former la base tronquée des modes propres  $\boldsymbol{\Psi} = [\boldsymbol{\psi}^{(1)}, \dots, \boldsymbol{\psi}^{(q)}]$  dans laquelle les  $q$  premiers modes propres associés aux fréquences propres  $\omega_j$  les plus basses ont été conservés. Par analogie avec l'équation (1.4), le vecteur  $\mathbf{x}$  des degrés de liberté du système se décompose sur la base des modes propres avec la relation

$$\mathbf{x} = \boldsymbol{\Psi} \mathbf{a}(t). \quad (1.7)$$

La solution particulière est donc approchée par un ensemble de modes propres associés aux premières fréquences de vibration de la structure. Le modèle d'ordre réduit correspondant au modèle d'ordre élevé décrit par les équations discrètes (1.5) est alors obtenu en introduisant la décomposition précédente (1.7) dans les équations du système puis en projetant le tout sur la base des modes propres. L'opération de projection est une simple multiplication matricielle par la transposée de la matrice  $\Psi$ . À l'issue de ces manipulations, le modèle d'ordre réduit se présente donc sous la forme

$$\left(\Psi^T \mathbf{M} \Psi\right) \ddot{\mathbf{a}}(t) + \left(\Psi^T \mathbf{K} \Psi\right) \mathbf{a}(t) = 0. \quad (1.8)$$

Or les modes propres sont orthonormaux par rapport à la matrice de masse et orthogonaux par rapport à la matrice de rigidité [83]. Les coefficients d'orthogonalité correspondent dans ce dernier cas aux valeurs propres  $\omega_j$ , si bien qu'avec ces propriétés le système d'équation (1.8) se réduit à

$$\ddot{a}_j(t) + \omega_j^2 a_j(t) = 0, \quad \text{pour tout } j = 1, \dots, q. \quad (1.9)$$

Il s'agit bien d'un modèle d'ordre réduit puisque la résolution du système précédent n'implique qu'un système de  $q$  équations au lieu des  $N_v$  équations du système discret initial. La base de projection servant à former le modèle d'ordre réduit est ici calculée directement à partir des opérateurs du système mécanique, sans qu'il n'y ait besoin de recourir à des simulations préalables pour générer une base de données représentative du système.

Le modèle d'ordre réduit construit de cette façon permet alors de calculer à moindre coût la réponse de la structure qui est donc supposée pouvoir être correctement représentée par ses premiers modes de vibration. Il s'avère que dans de nombreux cas la réponse de la structure est principalement gouvernée par les modes de basse fréquence et par conséquent cette technique de réduction s'avère efficace. Elle peut en particulier être employée dans la résolution de problèmes couplés afin de réduire au maximum les coûts de calcul, comme l'ont suggéré Dowell *et al.* [60] et Kamakoti et Shyy [125]. Carstens *et al.* [38] ont ainsi projeté les équations discrétisées d'un étage complet de turbomachine sur la base des modes propres. Pour ce type de configuration dans laquelle une symétrie cyclique de la géométrie est présente, le nombre de degrés de liberté peut être encore nettement réduit au moyen de techniques de sous-structuration. C'est ainsi qu'ont procédé Tran *et al.* [231] pour prédire l'écoulement dans un étage de turbomachine en ne simulant le système que sur un seul secteur compris entre deux aubes. L'hypothèse de symétrie cyclique est alors combinée aux méthodes de synthèse modale inspirées de celles proposées par Hurty [114] et Craig et Bampton [53] et à partir desquelles de nombreux développements ont depuis été réalisés [230, 232].

Tous ces travaux ne concernent néanmoins que la réduction de la partie structure. L'application de la méthode à un système aérodynamique s'avère particulièrement intéressante puisque dans ce cas les milieux fluide et structure pourraient être représentés par des modèles d'ordre réduit reposant sur un formalisme identique, idéal pour traiter un problème couplé aéroélastique. Lorsque le fluide est supposé incompressible, parfait et homogène, les équations d'un fluide potentiel, linéarisées pour de petites perturbations suffisent à le modéliser correctement. Dans ce cas, il est possible de définir des caractéristiques « vibratoires » du fluide décrivant les fluctuations autour d'un état d'équilibre. Il devient alors possible de définir des modes de ballonnement en basse fréquence, ou encore des modes acoustiques en haute fréquence. Ainsi, le système fluide peut être réduit au moyen de la base modale au même titre que la structure. Le lecteur pourra consulter en particulier les travaux de Morand et Ohayon [168] ainsi que les références [204, 211] pour plus de détails sur le calcul des modes propres dans ce contexte.

La difficulté apparaît lorsque le fluide évolue dans un domaine infini sur lequel il est modélisé par les équations d'Euler ou de Navier-Stokes, puisque dans ce cas, le système dynamique est non-linéaire. Dowell *et al.* [60] sont néanmoins parvenus à définir des modes propres sous l'hypothèse que l'écoulement est *dynamiquement linéaire* : la partie stationnaire de l'écoulement est supposée non-linéaire si nécessaire mais le caractère instationnaire du fluide est uniquement dû à des petites perturbations linéaires autour de l'état d'équilibre. De cette manière, le système dynamique du fluide est linéarisé et en l'absence d'excitation externe, les équations se ramènent à un problème aux valeurs propres. À la différence des matrices de masse et de raideur de la structure, l'opérateur aérodynamique linéarisé n'est pas symétrique. Les valeurs propres obtenues sont donc complexes et se voient associées des vecteurs propres à droite et à gauche. Le principe est expliqué en détail dans [59] et est résumé de façon générique pour l'ensemble des modèles fluide dans [60]. La méthode est toutefois confrontée à deux inconvénients majeurs. Le premier, purement numérique,

vient du fait que l'extraction des valeurs et vecteurs propres devient rapidement fastidieuse lorsque le nombre de degrés de liberté du modèle augmente. Des techniques particulières ont par conséquent dû être spécifiquement développées pour contourner ce problème [190]. La seconde limitation provient de l'hypothèse d'un écoulement linéarisé qui devient nécessaire pour construire le modèle d'ordre réduit.

La méthode est néanmoins efficace pour les écoulements dynamiquement linéaires, et Dowell *et al.* [60] l'ont donc employé pour reproduire l'écoulement potentiel autour d'un profil portant isolé et ont étudié l'influence de la discrétisation du domaine et de son étendue sur les valeurs propres qui décrivent la stabilité du système. Florea et Hall [78] ont appliqué la même technique sur un exemple similaire et ont indiqué que l'introduction d'une correction statique s'avère nécessaire pour tenir compte des modes propres de haute fréquence omis lors de la troncature. La solution instationnaire est alors décomposée en deux parties qui correspondent respectivement à la réponse du système à une perturbation quasi-statique et à la partie dynamique. Le calcul de la correction à chaque pas de temps alourdit toutefois le calcul. En employant une formulation basée sur la méthode des Éléments de Frontière, Shahverdi *et al.* [207] sont parvenus à construire un problème aux valeurs propres dans lequel les modes quasi-statiques sont éliminés, de sorte que la correction devient superflue. Enfin, Willcox [247] a comparé la réduction sur la base des modes propres à d'autres techniques dans le cas de la réponse aéroélastique d'un étage de turbomachine où le fluide est modélisé par les équations d'Euler linéarisées. Les résultats ont montré que la méthode devient rapidement fastidieuse à mettre en œuvre lorsque le modèle initial non réduit comporte un nombre important de degrés de liberté.

Les méthodes de condensation développées en mécanique des structures constituent aussi une alternative pour la construction des modèles d'ordre réduit en aérodynamique. C'est ainsi que Sarkar et Venkatraman [201] ont appliqué la méthode SEREP<sup>3</sup> pour construire un modèle d'ordre réduit capable de reproduire l'écoulement potentiel dans une cascade de profils.

La projection des équations sur la base des modes propres est donc une première possibilité, limitée néanmoins par deux restrictions majeures : d'une part la détermination des modes propres devient rapidement impossible à mettre en œuvre lorsque le modèle d'ordre élevé comporte un nombre important de degrés de liberté, et d'autre part les modes ne sont définis que pour des systèmes dynamiques linéaires.

Une perspective d'amélioration a cependant été proposée par Shaw et Pierre [209] qui ont tenté d'étendre la notion de modes propres pour construire des modes non-linéaires, définis comme des variétés invariantes tangentes aux modes propres linéaires à proximité d'un certain point. La méthode de construction proposée dans [209] ne s'applique toutefois qu'à des systèmes dynamiques impliquant des dérivées secondes en temps et un certain nombre de développements doivent encore être effectués pour espérer appliquer ce type d'approche aux équations non-linéaires de la mécanique des fluides.

### 1.3.2 Approximation de la fonction de transfert

En théorie du contrôle linéaire, les systèmes dynamiques sont considérés comme des systèmes entrée/sortie et se mettent donc sous la forme

$$(\Sigma_{\text{lin}}) : \begin{cases} \dot{\mathbf{x}}(t) &= \mathbf{A} \mathbf{x}(t) + \mathbf{B} \mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C} \mathbf{x}(t) + \mathbf{D} \mathbf{u}(t) \end{cases} . \quad (1.10)$$

La première équation correspond à l'écriture discrète d'un système dynamique linéaire ( $\Sigma$ ) tel qu'il a été défini à l'équation (1.2). Le vecteur  $\mathbf{x} \in \mathbb{R}^{N_v}$  contient l'ensemble des degrés de liberté discrets du système dynamique et le vecteur  $\mathbf{y} \in \mathbb{R}^{N_y}$  correspond à la sortie discrète qui décrit la grandeur d'intérêt du système. Enfin le vecteur  $\mathbf{u} \in \mathbb{R}^p$  est le vecteur d'entrée qui reflète l'action du paramètre  $\mu$  agissant sur le système dynamique. La version discrétisée de l'équation (1.2) prend alors la forme (1.10) dans laquelle les opérateurs du système dynamique se réduisent aux matrices  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  et  $\mathbf{D}$ . Le système est donc entièrement décrit par la matrice  $\Sigma_{\text{lin}} \in \mathcal{M}(N_v + N_y, N_v + p)$  définie par

$$\Sigma_{\text{lin}} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} . \quad (1.11)$$

<sup>3</sup>Abréviation de « System Equivalent Reduction and Expansion Process » en anglais. Une comparaison de la méthode SEREP avec d'autres méthodes de condensation telles que celle de Guyan-Irons est effectuée par Sastry *et al.* [203]. À l'origine, le principe de la méthode consiste à conserver un certain nombre de modes propres et de degrés de liberté dans un système linéaire comportant des non-linéarités locales. Le choix des modes conservés dépend du spectre de fréquence dans lequel le système fonctionne, et les degrés de liberté choisis dépendent de leur degré de participation aux modes retenus.



L'approche consiste à se concentrer sur la relation entrée/sortie qui suffit la plupart du temps à décrire le système puisque la connaissance du vecteur d'état n'est pas une finalité en soi. Ainsi, dans le cadre d'un système aéroélastique par exemple, le calcul du vecteur d'état  $\mathbf{x}$  décrivant les variables en chaque point du domaine spatial n'est qu'un intermédiaire pour évaluer une quantité intéressante telle que les efforts aérodynamiques par exemple. Lorsque c'est possible, il est donc plus intéressant de disposer directement de la relation entrée/sortie qui correspond à la *fonction de transfert*  $\mathbf{H}$  du système. En général la fonction de transfert s'écrit dans le domaine fréquentiel après avoir effectué une transformation de Laplace notée  $\mathcal{L}$  du système dynamique discret (1.10). On définit alors  $\mathbf{U}(s) = \mathcal{L}(\mathbf{u}(t))$  et de la même façon  $\mathbf{Y}(s) = \mathcal{L}(\mathbf{y}(t))$  où  $s$  est la variable complexe de Laplace. En exprimant alors l'entrée en fonction du vecteur d'état à l'aide de la première équation et en remplaçant dans la seconde, la fonction de transfert dans le domaine de Laplace s'écrit formellement

$$\mathbf{H}(s) = \mathbf{C} (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{B} + \mathbf{D}, \quad \text{tel que} \quad \mathbf{Y}(s) = \mathbf{H}(s) \mathbf{U}(s). \quad (1.12)$$

Le calcul de la fonction de transfert définie par l'expression précédente est toutefois difficile voire impossible en pratique lorsque la discrétisation des systèmes dynamiques conduit à un nombre important  $N_v$  de degrés de liberté. En effet, cela nécessite d'inverser la matrice  $\mathbf{A} \in \mathcal{M}(N_v, N_v)$ .

Ici encore, une projection du système dynamique sur une base adéquate permet de réduire considérablement la taille de la matrice à inverser et, partant, d'évaluer la fonction de transfert. Supposons en effet qu'on dispose de deux matrices  $\mathbf{V}$  et  $\mathbf{W}$  définies dans  $\mathcal{M}(N_v, q)$  telles que  $\mathbf{W}^T \mathbf{V} = \mathbf{I}_q$ . Alors la solution du système dynamique est approchée par  $\mathbf{x} = \mathbf{V} \hat{\mathbf{x}}$  et les équations du système dynamique sont projetées sur  $\mathbf{W}^T$ . La matrice  $\hat{\Sigma}_{\text{in}}$  décrivant le système dynamique se réduit alors à

$$\hat{\Sigma}_{\text{in}} = \begin{pmatrix} \mathbf{W}^T \mathbf{A} \mathbf{V} & \mathbf{W}^T \mathbf{B} \\ \mathbf{C} \mathbf{V} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_r & \mathbf{B}_r \\ \mathbf{C}_r & \mathbf{D} \end{pmatrix}. \quad (1.13)$$

La base de projection est alors choisie de façon à pouvoir correctement représenter le système dynamique avec un faible nombre  $q$  de vecteurs, de sorte que la matrice réduite  $\hat{\Sigma}_{\text{in}}$  appartienne à  $\mathcal{M}(q + N_y, q + p)$ . La fonction de transfert est alors aisément calculée puisqu'elle ne fait plus intervenir que des matrices de taille réduite dont l'inversion est rendue possible. Elle est donc définie par

$$\hat{\mathbf{H}}(s) = \mathbf{C} \mathbf{V} (s \mathbf{W}^T \mathbf{V} - \mathbf{W}^T \mathbf{A} \mathbf{V})^{-1} \mathbf{W}^T \mathbf{B} + \mathbf{D}. \quad (1.14)$$

Le point crucial de la méthode réside évidemment dans la détermination de la base de projection. Un certain nombre de méthodes ont été développées par le passé en particulier dans le domaine de l'automatique et de l'électronique. Antoulas [11] a ainsi distingué celles dites de *troncature équilibrée* et celles s'appuyant sur la construction de *sous-espaces de Krylov*.

### 1.3.2.1 Balanced Truncation

Une première manière de déterminer la base de projection a été introduite par Moore [167]. Le principe repose sur la définition d'un sous-espace dit *contrôlable* et d'un sous-espace dit *observable* auxquels on associe les matrices grammiennes de contrôlabilité et d'observabilité

$$\mathbf{W}_C = \int_0^\infty \exp(\mathbf{A}t) \mathbf{B} \mathbf{B}^T \exp(\mathbf{A}^T t) dt \quad \text{et} \quad \mathbf{W}_O = \int_0^\infty \exp(\mathbf{A}t) \mathbf{C} \mathbf{C}^T \exp(\mathbf{A}^T t) dt. \quad (1.15)$$

Ces matrices sont définies de telle manière que leurs vecteurs propres engendrent les sous-espaces contrôlable et observable respectivement. Elles sont solutions des équations de Lyapunov

$$\mathbf{A} \mathbf{W}_C + \mathbf{W}_C \mathbf{A}^T = -\mathbf{B} \mathbf{B}^T \quad \text{et} \quad \mathbf{A}^T \mathbf{W}_O + \mathbf{W}_O \mathbf{A} = -\mathbf{C}^T \mathbf{C}. \quad (1.16)$$

Le système est dit contrôlable (resp. observable) lorsque la matrice grammienne de contrôlabilité (resp. d'observabilité) est définie positive. Quand ces deux conditions sont vérifiées, le système dynamique est dit *équilibré* et il existe une transformation d'équilibre caractérisée par la matrice  $\mathbf{T}$  telle que les matrices grammiennes soient diagonalisées en

$$\mathbf{T}^{-1} \mathbf{W}_C (\mathbf{T}^{-1})^T = \mathbf{T}^T \mathbf{W}_O \mathbf{T} = \mathbf{\Sigma} \quad \text{avec} \quad \mathbf{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_{N_v}). \quad (1.17)$$

La matrice  $\mathbf{T}$  de la transformation d'équilibre découle du problème aux valeurs propres  $\mathbf{W}_C \mathbf{W}_O \mathbf{t}^{(j)} = \lambda_j \mathbf{t}^{(j)}$  où le vecteur  $\mathbf{t}^{(j)}$  correspond à la  $j$ -ième colonne de la matrice  $\mathbf{T}$ . Les termes de la matrice diagonale  $\Sigma$  sont donnés par  $\sigma_j = \sqrt{\lambda_j}$  et les valeurs propres sont organisées par ordre décroissant. Les valeurs propres de la diagonale sont les valeurs singulières de l'opérateur de Hankel qui est lié à l'opérateur de convolution. Le lien entre l'opérateur de Hankel et les matrices grammiennes est expliqué dans [11].

Finalemnt, la base de projection  $\mathbf{V}$  est constituée des  $q$  premières colonnes de la matrice  $\mathbf{T}$  qui sont associées aux valeurs singulières de Hankel les plus grandes, et la matrice  $\mathbf{W}^T$  contient les  $q$  premières lignes de  $\mathbf{T}^{-1}$ . Le principal avantage de la méthode est sa capacité à préserver la stabilité du modèle d'ordre élevé malgré la troncature effectuée pour réduire la dimension du système dans le modèle d'ordre réduit. Par ailleurs l'erreur globale commise reste bornée par la relation  $\sigma_{q+1} \leq \|\Sigma_{\text{lin}} - \widehat{\Sigma}_{\text{lin}}\|_{\infty} \leq 2 \sum_{i=q+1}^{N_v} \sigma_i$ . Toutefois, l'application de cette méthode pour les systèmes de très grande taille reste aussi problématique, notamment pour la résolution des équations de Lyapunov (1.16). Des méthodes approchées [99, 119] sont néanmoins disponibles pour parvenir à identifier les solutions. Enfin, une extension de la méthode aux systèmes non-linéaires a été proposée par Lall *et al.* [139].

### 1.3.2.2 Méthodes d'approximation de Krylov

Puisque le calcul de la base de projection devient rapidement difficile pour les systèmes de grande taille, une seconde famille de méthodes basées sur la correspondance des moments de la fonction de transfert s'est développée. Pour simplifier ici, on considère la fonction de transfert scalaire  $H(s)$  dont le développement de Taylor autour d'un point  $s_0$  s'écrit

$$H(s) = m_0 + m_1 (s - s_0) + m_2 (s - s_0)^2 + \dots \quad (1.18)$$

Lorsque le développement est pratiqué autour de  $s_0 = \infty$ , le problème est généralement qualifié de *réalisation partielle* — ou de « Partial Realization » en anglais — et les coefficients  $m_j$  sont les *paramètres de Markov*. En revanche, si le développement est effectué autour  $s_0 = 0$ , on parle plutôt d'*approximation de Padé* et les coefficients  $m_j$  sont les *moments* de la fonction de transfert autour du point  $s_0$ .

Le but est alors d'approcher la fonction de transfert du système complet  $H(s)$  par une fonction de transfert réduite  $\widehat{H}_q(s) = \widehat{m}_0 + \widehat{m}_1 (s - s_0) + \widehat{m}_2 (s - s_0)^2 + \dots$  telle que  $\widehat{m}_j = m_j$  pour  $0 \leq j \leq q$ . Un choix naturel pour approcher la fonction de transfert consiste à supposer qu'il s'agit d'une fraction rationnelle définie par

$$\widehat{H}_q(s) = \frac{P_{q-1}(s)}{Q_q(s)} = \frac{a_{q-1}s^{q-1} + \dots + a_1s + a_0}{b_qs^q + b_{q-1}s^{q-1} + \dots + b_1s + 1}. \quad (1.19)$$

La détermination des coefficients de la fraction rationnelle s'effectue au moyen de la résolution du système linéaire  $H(s)Q_q(s) = P_{q-1}(s)$  dans lequel le second membre est identifié à celui de l'équation (1.18). Le problème matriciel résultant implique une matrice de Hankel qui est en général très mal conditionnée [20]. L'évaluation de l'approximation de Padé au moyen du calcul explicite des moments est donc numériquement instable. Toutefois, des méthodes ont été développées pour remédier à ce problème en utilisant les connections entre les approximants de Padé et les procédures de Lanczos ou d'Arnoldi. Freund [80] a fourni sur ce point une description détaillée des différents algorithmes employés pour parvenir à identifier la fonction de transfert réduite.

Pour faire le lien avec le titre de ce paragraphe, il faut signaler qu'il existe une relation étroite entre les moments  $m_j$  définissant la matrice de Hankel et les sous-espaces de Krylov associés au problème. Étant donné une matrice  $\mathbf{M}$  et un vecteur de départ  $\mathbf{d}$  (resp.  $\mathbf{g}$ ), le sous-espace de Krylov de dimension  $q$  à droite (resp. à gauche) est défini par l'espace engendré par la séquence des  $q$  vecteurs

$$\mathcal{K}_q(\mathbf{M}, \mathbf{d}) = \text{vect} \{ \mathbf{d}, \mathbf{M} \mathbf{d}, \dots, \mathbf{M}^{q-1} \mathbf{d} \} \quad \left( \text{resp. } \mathcal{K}_q(\mathbf{M}^T, \mathbf{g}) = \text{vect} \{ \mathbf{g}, \mathbf{M}^T \mathbf{g}, \dots, (\mathbf{M}^T)^{q-1} \mathbf{g} \} \right). \quad (1.20)$$

Bai [20] et Freund [80] expliquent alors que les  $2q$  premiers moments  $m_j$  de la fonction de transfert (1.18) qui définissent la matrice de Hankel s'obtiennent simplement en calculant les produits entre les espaces de Krylov à droite et à gauche avec  $\mathbf{M} = (\mathbf{A} - s\mathbf{I})^{-1}$ , ce qui conduit aux expressions

$$m_{2j} = \left( (\mathbf{M}^T)^j \mathbf{g} \right)^T \cdot (\mathbf{M}^j \mathbf{d}) \quad \text{et} \quad m_{2j+1} = \left( (\mathbf{M}^T)^j \mathbf{g} \right)^T \cdot (\mathbf{M}^{j+1} \mathbf{d}). \quad (1.21)$$

Les matrices de projection  $\mathbf{V}$  et  $\mathbf{W}$  intervenant dans la construction du modèle d'ordre réduit défini à l'équation (1.13) sont alors construites de manière à ce qu'elles engendrent les mêmes sous-espaces de Krylov que ceux définis par les relations (1.20). Le point crucial de la méthode consiste à choisir correctement les points d'interpolation  $s_0$ .

Cependant, bien que les sous-espaces droit et gauche de Krylov contiennent les informations souhaitées sur les moments, les vecteurs de base engendrant les sous-espaces ne sont pas adaptés. L'astuce consiste donc à employer d'autres vecteurs de base qui engendrent les mêmes sous-espaces de Krylov. Les procédés de Lanczos [140] et d'Arnoldi [15] sont alors des moyens idéaux d'obtenir les vecteurs recherchés et de reconstruire finalement la fonction de transfert du système réduit, d'autant plus que les vecteurs obtenus sont (bi-)orthogonaux. Lorsque l'algorithme de Lanczos est employé, cette approche est connue sous le nom de méthode Padé Via Lanczos (PVL). Deux exemples d'application sont fournis dans [20], l'un concernant un circuit électrique pour lequel les équations de Maxwell ont été discrétisées et l'autre portant sur un système de frein automobile où la dynamique est calculée par une méthode Éléments Finis.

L'inconvénient principal de la méthode de Padé est qu'elle ne conserve plus les propriétés de stabilité et de passivité du système original [20, 214]. Un moyen de contrer ces deux inconvénients a été proposé par Bai et Freund [21]. Il s'agit d'une variante des méthodes PVL appelée  $PVL\pi$  qui agit postérieurement à l'obtention de la fonction de transfert réduite. La méthode est une procédure *trial-and-error* qui corrige les pôles et zéros instables de la fonction de transfert réduite. Silveira *et al.* [214] ont proposé quant à eux une technique analogue reposant cette fois sur l'utilisation de l'algorithme d'Arnoldi. La conservation des propriétés de stabilité du système s'obtient simplement au prix d'un changement de coordonnées dans l'espace d'état. Le modèle d'ordre réduit ainsi construit est directement stable, sans qu'il soit nécessaire de le modifier a posteriori comme cela est le cas dans les méthodes de type  $PVL\pi$ . Enfin, Antoulas [11] et Gugercin [100] ont préféré combiner les méthodes de Krylov à la méthode de troncature équilibrée afin de préserver les avantages de chacune des deux approches. L'idée repose sur le calcul itératif approché des équations de Lyapunov (1.16).

Le champ d'application des méthodes d'approximation de la fonction de transfert est à l'origine celui de l'automatique et du contrôle de systèmes linéaires. La méthode est devenue très populaire pour la réduction de systèmes en simulation de circuits électroniques notamment. Quelques travaux ont néanmoins été effectués dans le domaine de l'aérodynamique. Karpel [127] a ainsi cherché à modéliser les forces aérodynamiques en réponse à une sollicitation harmonique au moyen d'une fonction de transfert d'ordre minimal étant donné un certain degré de précision. Ueda et Dowell [233] ont quant à eux tenté d'étendre ce type d'approximation à des systèmes non-linéaires pour lesquels l'analogie de la fonction de transfert porte le nom de fonction descriptive. Plus récemment, Willcox [247, 248] a utilisé le procédé d'Arnoldi pour construire le modèle d'ordre réduit et simuler le comportement aéroélastique d'un étage de turbomachine dans lequel l'écoulement est modélisé par les équations d'Euler linéarisées. Par rapport à la méthode de projection sur la base des modes propres qui est aussi mise en œuvre dans [247], les vecteurs d'Arnoldi sont peu coûteux à calculer et le procédé d'Arnoldi a l'avantage de générer une transformation congruente qui conserve la stabilité du système initial. Rule *et al.* [197] ont employé la méthode de troncature équilibrée pour traiter le problème du flottement d'un profil dans un écoulement 2D incompressible. En comparaison avec une réduction sur la base des modes propres, la réduction obtenue est satisfaisante mais le coût de calcul de la transformation reste élevé et devient donc prohibitif pour de grands systèmes.

Bien que les méthodes décrites dans ce paragraphe possèdent des propriétés extrêmement intéressantes en terme de conservation de la stabilité et de la passivité du système initial non réduit, leur domaine d'application est exclusivement réservé aux systèmes dynamiques linéaires.

### 1.3.3 Identification du système réduit

Plutôt que de chercher à définir les matrices de projection permettant de construire le modèle d'ordre réduit, il est possible de se servir d'un ensemble de réponses du système à différentes sollicitations afin d'*identifier* directement les matrices réduites. Le principe consiste donc en quelque sorte à résoudre un problème inverse à partir d'une base de données caractéristique de la réponse du système. Une première manière de procéder pour les systèmes linéaires s'appuie sur la construction de la matrice de Hankel associée au système dynamique. Ce dernier est caractérisé par les équations (1.10) avec pour simplifier  $\mathbf{D} = \mathbf{0}_{N_y, p}$ . Le modèle d'ordre réduit de dimension  $q$  gouvernant l'évolution du vecteur d'état  $\hat{\mathbf{x}} \in \mathbb{R}^q$  est recherché

sous la forme du système dynamique linéaire discret en temps suivant, avec  $\mathbf{u} \in \mathbb{R}^p$  :

$$(\widehat{\Sigma}_{\text{lin}}) : \begin{cases} \widehat{\mathbf{x}}(k+1) &= \mathbf{A}_r \widehat{\mathbf{x}}(k) + \mathbf{B}_r \mathbf{u}(k) \\ \mathbf{y}(k) &= \mathbf{C}_r \widehat{\mathbf{x}}(k) \end{cases} . \quad (1.22)$$

Lorsque le système dynamique est excité par une réponse impulsionnelle telle que  $\mathbf{u}(0) = \mathbf{1}$  et  $\mathbf{u}(k) = \mathbf{0}$  pour tout  $k > 0$ , la réponse est caractérisée par l'ensemble de matrices

$$\mathbf{Y}[1] = \mathbf{C}_r \mathbf{B}_r, \quad \mathbf{Y}[2] = \mathbf{C}_r \mathbf{A}_r \mathbf{B}_r, \quad \dots \quad \mathbf{Y}[k] = \mathbf{C}_r \mathbf{A}_r^{k-1} \mathbf{B}_r \quad (1.23)$$

qui découle de la substitution de l'impulsion dans le système dynamique précédent. Les matrices précédentes  $\mathbf{Y}[k] = \mathbf{C}_r \mathbf{A}_r^{k-1} \mathbf{B}_r$  sont les paramètres de Markov du système dynamique. Le principe de superposition est valable puisque le système dynamique est supposé linéaire. Par conséquent la réponse à n'importe quelle sollicitation s'exprime comme une combinaison linéaire des réponses impulsionnelles du système et les paramètres de Markov sont donc reliés à la matrice généralisée de Hankel

$$\mathbf{H}_{RM}[k-1] = \begin{bmatrix} \mathbf{Y}[k] & \dots & \mathbf{Y}[k+t_{M-1}] \\ \vdots & & \vdots \\ \mathbf{Y}[k+j_{R-1}] & \dots & \mathbf{Y}[k+j_{R-1}+t_{M-1}] \end{bmatrix} \quad (1.24)$$

qui utilise  $R$  fenêtrages de taille  $j_r$  de la réponse impulsionnelle pour  $M$  instants discrets  $t_m$ . Le choix de  $R$  et  $M$  est arbitraire tant que le rang de la matrice est plus grand que la dimension  $q$  du système réduit à identifier. Par exemple la réponse impulsionnelle  $\mathbf{y}(k)$  calculée avec le modèle d'ordre élevé pour  $M$  instants discrets est employée pour former la matrice de Hankel de taille  $p \times M$

$$\mathbf{H}(0) = [\mathbf{y}(1) \quad \mathbf{y}(2) \quad \dots \quad \mathbf{y}(M-1)]. \quad (1.25)$$

Si les équations (1.22) du modèle d'ordre réduit sont en mesure de caractériser correctement la réponse du système dynamique, la matrice de Hankel calculée avec le modèle d'ordre élevé est en théorie reliée aux matrices du système dynamique réduit par la relation

$$\mathbf{H}(0) = \mathbf{C}_r [\mathbf{B}_r \quad \mathbf{A}_r \mathbf{B}_r \quad \dots \quad \mathbf{A}_r^{M-2} \mathbf{B}_r]. \quad (1.26)$$

La technique consiste alors à identifier les matrices du système réduit à partir de la matrice de Hankel construite au moyen des réponses impulsionnelles du modèle d'ordre élevé. Pour cela, une décomposition en valeurs singulières tronquée de la matrice de Hankel (1.25) est pratiquée afin d'éliminer les informations redondantes et de réduire la taille des matrices. Seules les  $q$  premières valeurs singulières les plus importantes sont conservées dans l'approximation de la matrice de Hankel qui se décompose donc en<sup>4</sup>  $\mathbf{H}(0) \approx \mathbf{U}_q \Sigma_q \mathbf{V}_q^T$ . Les matrices  $\mathbf{B}_r$  et  $\mathbf{C}_r$  du système réduit sont alors définies par

$$\mathbf{B}_r = \Sigma_q^{1/2} \mathbf{V}_q^T \mathbf{E}_M \quad \text{et} \quad \mathbf{C}_r = \mathbf{E}_p^T \mathbf{U}_q \Sigma_q^{1/2}, \quad (1.27)$$

avec  $\mathbf{E}_M^T = [\mathbf{I}_M, \mathbf{0}, \dots, \mathbf{0}]$  et  $\mathbf{E}_p^T = [\mathbf{I}_p, \mathbf{0}, \dots, \mathbf{0}]$ . L'identification de la matrice  $\mathbf{A}_r$  nécessite par ailleurs la construction de la matrice de Hankel  $\mathbf{H}(1) = [\mathbf{y}(2) \quad \mathbf{y}(3) \quad \dots \quad \mathbf{y}(M)]$ . De cette manière, la relation (1.26) devient  $\mathbf{H}(1) = \mathbf{C}_r \mathbf{A}_r [\mathbf{B}_r \quad \mathbf{A}_r \mathbf{B}_r \quad \dots \quad \mathbf{A}_r^{M-2} \mathbf{B}_r]$  ce qui permet d'identifier la matrice  $\mathbf{A}_r$  avec la relation

$$\mathbf{A}_r = \Sigma_q^{-1/2} \mathbf{U}_q^T \mathbf{H}(1) \mathbf{V}_q \Sigma_q^{-1/2}. \quad (1.28)$$

Tang *et al.* [225] ont ainsi employé cette méthode pour identifier les coefficients du modèle d'ordre réduit développé pour une configuration d'aile delta 3D oscillante. Silva et Bartels [213] ont quant à eux étudié la réponse aéroélastique en flottement de l'aile AGARD 445.6. Pour cela, la structure est représentée par ses quatre premiers modes propres qui sont ensuite excités séparément par une impulsion dans le code fluide. Les réponses impulsionnelles sont alors utilisées pour identifier les coefficients du modèle d'état à l'aide de l'algorithme ERA (Eigensystem Realization Algorithm) développé par Juang et Pappa [123]. Bien que le modèle d'ordre réduit obtenu soit performant, Silva et Bartels [213] ont indiqué que certaines précautions

<sup>4</sup>Voir l'annexe A.2 pour les détails relatifs à la décomposition en valeurs singulières.

doivent être prises pour imposer les réponses impulsionnelles, notamment en ce qui concerne la résolution temps/fréquence et la dépendance de la solution avec l'amplitude des impulsions. Enfin, Allen *et al.* [4] ont identifié un modèle d'ordre réduit suivant le même procédé et ont construit des lois de contrôle afin de repousser la limite de flottement d'un profil portant bidimensionnel.

La théorie de Volterra fournit une extension possible de la méthode d'identification précédente pour les systèmes dynamiques non-linéaires. L'expression de la réponse du système dynamique en fonction des *noyaux de Volterra* et d'un ensemble d'excitations impulsionnelles supposées ici scalaires est donc

$$\begin{aligned} \mathbf{x}(t) = & \mathbf{h}_0 + \int_{t_0}^{t_0+t} \mathbf{h}_1(t-\tau) u(\tau) d\tau + \int_{t_0}^{t_0+t} \int_{t_0}^{t_0+t} \mathbf{h}_2(t-\tau_1, t-\tau_2) u(\tau_1) u(\tau_2) d\tau_1 d\tau_2 \\ & + \dots + \int_{t_0}^{t_0+t} \dots \int_{t_0}^{t_0+t} \mathbf{h}_N(t-\tau_1, \dots, t-\tau_N) u(\tau_1) \dots u(\tau_N) d\tau_1 \dots d\tau_N. \end{aligned} \quad (1.29)$$

Le premier terme de la réponse correspond à la partie stationnaire  $\mathbf{x}_s$  de la réponse du système dynamique. Le second terme représente la convolution entre le premier noyau de Volterra  $\mathbf{h}_1(\tau)$  et l'excitation impulsionnelle  $u(\tau)$  et caractérise la réponse linéaire du système. Le premier noyau de Volterra est identifié en mesurant la réponse du système dynamique à un Dirac et revient de ce fait à la technique d'identification précédente. Cependant, la théorie de Volterra permet de tenir compte des effets non-linéaires par le biais des noyaux d'ordres supérieurs. Lorsque le système dynamique est seulement quadratique, Silva [212] est parvenu à identifier les trois premiers noyaux de Volterra  $\mathbf{h}_0 = \mathbf{x}_s$ ,  $\mathbf{h}_1(\tau_1) = 2\mathbf{x}_0(\tau_1) - 1/2\mathbf{x}_2(\tau_1)$  et  $\mathbf{h}_2(\tau_1, \tau_2) = 1/2[\mathbf{x}_1(\tau_1, \tau_2) - \mathbf{x}_0(\tau_1) - \mathbf{x}_0(\tau_2)]$  au moyen de réponses impulsionnelles adéquates. Dans les expressions précédentes,  $\mathbf{x}_s$  est la réponse stationnaire du système dynamique,  $\mathbf{x}_0(\tau_1)$  est celle due à un Dirac appliqué à l'instant initial et  $\mathbf{x}_2(\tau_1)$  est celle due à un Dirac d'amplitude deux fois plus grande que celle de  $\mathbf{x}_0(\tau_1)$  imposé aussi à l'instant initial, et enfin  $\mathbf{x}_1(\tau_1, \tau_2)$  est la réponse provoquée par deux impulsions à des instants différents.

Comme l'ont fait remarquer Lucia *et al.* [153], le calcul des noyaux de Volterra n'est pas une fin en soi, même si il est essentiel pour la construction d'un modèle d'ordre réduit qui reste encore à déterminer dans le cas d'un système non-linéaire. Lucia *et al.* [153] ont ainsi calculé les deux premiers noyaux de Volterra de l'écoulement transsonique autour d'un profil RAE en mouvement de plongé. La réponse a été comparée pour trois amplitudes croissantes d'oscillation du profil. Dans ce cas assez simple, un seul noyau suffit à bien capter la réponse du système, mais la différence se creuse lorsque l'amplitude des oscillations, et donc le caractère non-linéaire du système, augmente. Sibetheros et Niedzwecki [210] ont pour leur part analysé le comportement non-linéaire de cylindres placés dans un écoulement au moyen d'un développement en série de Volterra tronqué à l'ordre trois. L'objectif était d'étudier la dépendance linéaire, quadratique ou cubique des fréquences présentes dans les données mais aucun modèle d'ordre réduit n'a été construit.

Enfin des applications intéressantes sont en cours de développement pour combiner ensemble la décomposition orthogonale propre qui sera introduite dans le paragraphe suivant et la théorie de Volterra. Ainsi, Lucia *et al.* [153, 154] ont développé un modèle d'ordre réduit hybride POD/Volterra dans lequel les modes propres de la structure étaient impulsés et la réponse du fluide était projetée sur la base POD pour obtenir l'historique temporel des amplitudes modales. Seul le premier noyau de Volterra a été calculé afin de construire le modèle d'état linéaire à l'aide de l'algorithme ERA. Le modèle d'ordre réduit obtenu a été appliqué à la prédiction des cycles limites d'oscillation d'un panneau élastique dans un écoulement supersonique parfait 2D [154]. La structure du panneau est non-linéaire mais le fluide est traité de façon linéaire, afin de pouvoir identifier le modèle d'ordre réduit. Enfin, pour éviter d'avoir à calculer les modes POD à partir d'une simulation supplémentaire du système complet, une tentative a été menée pour obtenir ces modes à l'aide des réponses impulsionnelles utilisées pour former la matrice de Hankel. Toutefois, les modes POD obtenus ne semblent pas adaptés.

Pour terminer ce paragraphe, il est important de mentionner l'apparition récente de méthodes d'identification adaptées aux systèmes non-linéaires. Sans entrer dans les détails, le principe consiste à résoudre un problème inverse afin d'identifier les opérateurs du modèle d'ordre réduit à partir d'une structure générale des équations [22, 191, 236]. Une approche d'identification modale est alors nécessaire pour diminuer la taille du problème inverse. On reviendra plus tard sur ce type d'approche employé notamment par Perret *et al.* [177] dans le cadre de la POD.

### 1.3.4 Décomposition Orthogonale Propre

À l'instar des méthodes d'identification précédentes, la procédure de décomposition orthogonale propre repose sur un ensemble de réponses du système dynamique à partir duquel une base de faible dimension est extraite. La méthode est connue sous des appellations différentes dans des domaines d'application aussi variés que l'océanographie, la météorologie, la turbulence, l'identification d'image, la compression de données, l'économie, la psychologie, la biologie, ... Son champ d'application est vaste puisque la méthode est intimement liée à la *décomposition de Karhunen-Loève* introduite dans le domaine des probabilités [126, 150], ou encore à l'*analyse en composantes principales* issue des statistiques [113]. Des liens étroits entre ces deux méthodes et la décomposition en valeurs singulières ont récemment été mis en évidence et formalisés par Volkwein [241]. Liang *et al.* [144] ont par ailleurs démontré l'équivalence des trois méthodes lorsque celles-ci sont employées pour des vecteurs discrets aléatoires de dimension finie. Selon les auteurs, la décomposition orthogonale propre peut être considérée soit comme étant la décomposition de Karhunen-Loève, soit comme un terme générique regroupant à la fois la décomposition de Karhunen-Loève, l'analyse en composante principale ou encore la décomposition en valeurs singulières suivant les données auxquelles la procédure est appliquée.

Le principe consiste à rechercher les fonctions de base qui représentent de façon optimale l'ensemble des données auxquelles on attribue généralement le nom de *clichés*. Les clichés sont par exemple les solutions  $x^{(m)}$  ou  $x_h^{(m)}$  des systèmes dynamiques (1.1) et (1.2) respectivement, obtenues à différents instants  $t_m$  ou pour différentes valeurs  $\mu_m$  du paramètre  $\mu$ . En fonction du domaine d'application, les vecteurs de la base portent le nom de fonctions empiriques, de modes principaux, de modes propres orthogonaux, ou encore de modes POD par abus de langage. La procédure de construction d'un modèle d'ordre réduit, détaillée par Holmes *et al.* [112] dans le cadre des systèmes dynamiques, est brièvement reprise ici pour décrire le principe. L'objectif est de déterminer une base  $\Phi$  orthonormale de dimension  $q$  dont les vecteurs  $\varphi^{(j)}$  contenus dans un espace de Hilbert  $H$  muni du produit scalaire  $\langle \cdot, \cdot \rangle$  approchent le mieux possible en moyenne les clichés  $x^{(m)}$  de l'ensemble de données. Il s'agit donc de minimiser en moyenne la différence entre les clichés et leur projection orthogonale sur le sous-espace engendré par les modes POD  $\varphi^{(j)}$ . La base POD est donc solution du problème d'optimisation

$$\Phi = \left\{ \varphi^{(1)}, \dots, \varphi^{(q)} \right\} = \arg \min_{\{\varphi^{(j)}\}_{j=1}^q \in H} E \left[ \left\| x^{(m)} - \sum_{i=1}^q \langle x^{(m)}, \varphi^{(i)} \rangle \varphi^{(i)} \right\|^2 \right], \quad (1.30)$$

avec la contrainte supplémentaire  $\langle \varphi^{(i)}, \varphi^{(j)} \rangle = \delta_{ij}$  qui traduit l'orthonormalité des modes POD entre eux [135]. L'opérateur de moyenne  $E[\cdot]$  est souvent défini comme une moyenne en temps lorsque les clichés sont des solutions du système dynamique à différents instants. On peut donc interpréter les modes POD  $\varphi^{(j)}$  comme les vecteurs qui correspondent aux minima de la fonctionnelle, ou encore comme les vecteurs qui sont en moyenne le plus parallèles possibles à l'ensemble des données.

D'un point de vue pratique, le problème (1.30) se ramène à la résolution du problème aux valeurs propres

$$R \varphi^{(j)} = \lambda_j \varphi^{(j)}, \quad (1.31)$$

où l'opérateur  $R$  est une matrice de corrélation temporelle entre les différents clichés, pour laquelle il a été démontré qu'il existe un nombre fini de valeurs propres [112]. Le spectre des valeurs propres décroît en général rapidement, ce qui signifie que les premiers vecteurs propres sont en mesure de capturer les principales structures communes aux différents clichés. Cette propriété d'optimalité explique pourquoi la méthode est particulièrement populaire lorsqu'il s'agit d'analyser un grand ensemble de données afin d'y repérer des structures particulières. À l'origine employée pour détecter objectivement les structures cohérentes d'un écoulement turbulent, la décomposition orthogonale propre a été depuis employée à de multiples reprises en mécanique pour analyser le comportement des systèmes. Kerschen *et al.* [130] ont ainsi suggéré d'employer la méthode pour caractériser la dynamique du système et détecter en particulier certaines non-linéarités. Tamura *et al.* [224] l'ont quant à eux employé pour extraire les structures principales dans le champ de pression fluctuant exercé sur la paroi d'un bâtiment. Enfin, Cizmas et Palacios [47] se sont intéressés aux structures des modes POD de l'écoulement dans un étage de turbomachine. Quelques applications pour l'analyse de problèmes couplés avec des applications en aéroélasticité ont aussi vu le jour [234, 252].

En plus de ses qualités pour l'analyse de données, la méthode possède encore bien d'autres atouts en ce qui concerne la construction d'un modèle d'ordre réduit. En effet, les  $q$  premiers vecteurs propres associés aux plus grandes valeurs propres forment une base de projection  $\Phi$  idéale. Puisque la base des modes POD constitue la meilleure approximation<sup>5</sup> qui soit des clichés  $x^{(m)}$ , on peut faire l'hypothèse que la variable d'état  $x$  du système dynamique se décompose à tout instant sur la base  $\Phi$  selon la relation

$$x(t) = \sum_{j=1}^q a_j(t) \varphi^{(j)}. \quad (1.32)$$

Cette décomposition est alors introduite dans l'équation (1.1) du système dynamique. La projection de Galerkin sur la base des modes POD conduit alors au système d'équations

$$\dot{a}_i = \left\langle f \left( \sum_{j=1}^q a_j(t) \varphi^{(j)}, \mu \right), \varphi^{(i)} \right\rangle, \quad (1.33)$$

pour tout  $i = 1, \dots, q$ . L'expression précédente conduit à un modèle d'ordre réduit exploitable de façon efficace lorsque le second membre  $f$  du système dynamique est linéaire ou multilinéaire. En particulier lorsque les non-linéarités sont de type polynomial, l'opérateur de sommation peut être extrait et le modèle d'ordre réduit prend la forme<sup>6</sup>

$$\dot{a}_i = \mathcal{K}_i + \sum_{j=1}^q \mathcal{L}_{ij} a_j(t) + \sum_{j,k=1}^q \mathcal{Q}_{ijk} a_j(t) a_k(t) + \dots \quad (1.34)$$

Le système d'équations (1.34) comprend  $q$  équations différentielles ordinaires non-linéaires qui constituent un modèle d'ordre réduit *non-linéaire*. La projection de Galerkin du système dynamique sur la base des modes POD fournit donc une procédure efficace pour construire le modèle d'ordre réduit, au moins lorsque les non-linéarités du système dynamique sont polynomiales. Si l'expression du second membre  $f$  fait apparaître des fractions rationnelles ou des non-linéarités plus complexes, l'extraction des amplitudes modales telle qu'elle a été pratiquée entre les équations (1.33) et (1.34) n'est en général pas possible sans avoir recours à certaines approximations.

La décomposition orthogonale propre est apparue en mécanique des fluides dans un premier temps pour analyser les structures turbulentes dont le contenu stochastique peut être examiné par des outils statistiques. Lumley [156] a ainsi introduit la décomposition orthogonale propre pour identifier les structures cohérentes d'un écoulement turbulent incompressible et s'en est ensuite servi pour former un modèle d'ordre réduit capable de reproduire approximativement la dynamique de l'écoulement. Dans ce cas, les modes POD correspondent aux structures spatiales contenant le maximum d'énergie cinétique. La décomposition orthogonale propre s'est alors imposée comme une généralisation de la décomposition de Fourier dans les directions inhomogènes ou non périodiques de l'écoulement [112]. L'utilisation de la méthode a été par la suite révolutionnée par la découverte de formulations équivalentes au problème aux valeurs propres (1.31). Ainsi, dans la *méthode des clichés* de Sirovich [218], l'opérateur de corrélation temporelle est remplacé par un opérateur de corrélation spatiale de dimension nettement plus faible pour lequel l'extraction des vecteurs propres est aisée. Par ailleurs, la décomposition en valeurs singulières de l'ensemble des clichés produit les modes POD au moyen d'algorithmes peu coûteux.

L'objectif des modèles d'ordre réduit était à l'origine de pouvoir décrire simplement l'écoulement au moyen des structures cohérentes dont l'importance est capitale pour la compréhension et la modélisation des écoulements turbulents. Elles jouent en effet un rôle capital dans les processus de mélange, la génération de bruit ou la création de traîné. Par conséquent, le but ultime est de pouvoir contrôler les structures turbulentes d'un écoulement et la construction d'un modèle d'ordre réduit est un moyen efficace d'y parvenir si celui-ci est capable de reproduire fidèlement l'écoulement. Pour une description détaillée des applications en turbulence, le lecteur pourra consulter en particulier la revue effectuée par Berkooz *et al.* [32]. Parmi les

<sup>5</sup>Dans le sens où l'erreur de projection orthogonale est minimisée.

<sup>6</sup>Les détails de la projection de Galerkin menant à l'expression (1.34) seront explicités dans le paragraphe 2.4.3 p.52 dans le cas général d'opérateurs multilinéaires.

travaux menés, Aubry *et al.* [18] ont développé un modèle d'ordre réduit des équations de Navier-Stokes pour la zone proche de la paroi d'une couche limite turbulente. Les effets des modes négligés sont pris en compte par un modèle de fermeture de type Heisenberg censé préserver la stabilité du système réduit. Dès les premiers travaux, la nécessité de stabiliser le modèle d'ordre réduit est donc apparue. Rempfer [186, 187] s'est aussi intéressé à la construction et à l'analyse des structures turbulentes dans une couche limite de transition. Quelques temps plus tard, il démontrait sur un exemple dans lequel la solution pouvait être exactement représentée par un petit nombre de modes POD, que la réponse du modèle d'ordre réduit construit par une projection de Galerkin pouvait exhiber un comportement physiquement instable et que les structures dans l'espace d'état pouvaient être différentes de celles de l'écoulement réel alors même qu'aucune erreur ne pouvait être attribuée à la troncature de la base POD [188].

Depuis lors, de nombreux travaux pour les écoulements d'un fluide incompressible turbulent se sont orientés vers le développement de modèles d'ordre réduit pour lesquels la stabilité est une priorité. Parmi la littérature abondante sur le sujet, on peut mentionner les travaux de Cazemier *et al.* [41, 42] qui ont tenté de reproduire la circulation d'un fluide dans une cavité entraînée. Cet exemple classique a par la suite été ré-employé pour valider une méthode de calcul des modes POD au moyen de l'algorithme de Lanczos [70] et une méthode de stabilisation [124]. Un grand nombre d'études a par ailleurs été mené pour l'étude de l'écoulement dans le sillage de divers types d'obstacles, tels qu'un cylindre de section cylindrique [171, 215, 216] ou carrée [35, 48, 50, 81, 82, 124], une marche descendante [50, 51] ou encore un profil portant fortement incliné [73]. L'ensemble de ces développements a permis progressivement d'envisager l'utilisation des modèles d'ordre réduit afin de contrôler les écoulements [31, 49, 175, 185, 226, 244, 245].

Des applications plus originales de la méthode POD aux écoulements incompressibles ont aussi vu le jour : Allery *et al.* l'ont ainsi employé pour l'étude de l'effet Coanda [7] et l'étude de la dispersion de particules transportées par le fluide [6]; Liberge [145, 146] a construit un modèle d'ordre réduit pour reproduire certains phénomènes résultant de l'interaction fluide-structure entre un cylindre en mouvement et l'écoulement. Par ailleurs des travaux théoriques portant sur la définition d'estimateurs d'erreur pour les équations de Navier-Stokes d'un fluide incompressible ont été récemment menés afin de parvenir à contrôler la qualité des modèles d'ordre réduit POD-Galerkin [56, 110, 136].

La majorité des modèles d'ordre réduit construits au moyen de la décomposition orthogonale propre concerne les fluides incompressibles puisque dans ce cas les équations de Navier-Stokes peuvent être assimilées à un système d'équations différentielles ordinaires dont le second membre est polynomial et quadratique. Les modes POD calculés fournissent alors les structures d'énergie cinétique maximale et se prêtent bien à la projection de Galerkin. Il en est tout autrement pour les fluides compressibles puisque dans ce cas la définition du produit scalaire ne conduit plus nécessairement à une interprétation des modes en terme d'énergie physique. Rowley *et al.* [192, 194] ont été parmi les premiers à se pencher sur la définition correcte d'un produit scalaire pour les écoulements compressibles. Aucune définition adéquate n'a été obtenue sauf lorsque les équations de Navier-Stokes isentropiques sont considérées. Ce même modèle a par la suite été réutilisé par Gloerfelt [87] pour l'analyse des radiations sonores auto-entretenues produites par une cavité, ainsi que par Moreno *et al.* [169] qui ont comparé la réponse du modèle d'ordre réduit d'un jet rectangulaire supersonique à des résultats expérimentaux. Vigo [239, 240] a pour sa part choisi d'employer les équations de Navier-Stokes formulées non pas avec le jeu de variables conservatives usuel, mais avec des variables primitives modifiées de façon à rendre les équations quadratiques. Toutefois, le produit scalaire employé conduit à des modes POD qui maximisent une énergie purement mathématique. Le même problème de stabilité que pour les modèles d'ordre réduit des écoulements incompressibles apparaît et Iollo *et al.* [116, 117] ont proposé pour remédier à ce problème de changer la définition du produit scalaire afin de tenir compte des dérivées premières en espace du champ aérodynamique. La généralisation au cas des espaces de Sobolev avec la prise en compte de dérivées d'ordre supérieur avait déjà été appliquée par Kirby [132] pour développer un modèle d'ordre réduit des équations de Kuramoto-Sivashinsky. Dans le même esprit, Barone *et al.* [23] ont modifié le produit scalaire pour symétriser le problème et ont cherché à tenir compte explicitement des conditions aux limites dans le modèle d'ordre réduit. Enfin, Bourguet *et al.* [34] ont construit puis corrigé un modèle d'ordre réduit au moyen de méthodes similaires à celles développées par Couplet *et al.* [51]. Finalement, Gloerfelt [87] a effectué une comparaison exhaustive des méthodes de correction pour les modèles d'ordre réduit POD-Galerkin développés pour les écoulements compressibles.

Dans les références citées jusqu'ici, les modèles d'ordre réduit ont pu être développés suivant le principe exposé au début de ce paragraphe puisque les équations non-linéaires du fluide ont été écrites sous



la forme d'un système d'équations polynomiales quadratiques. Ces modèles d'ordre réduit sont donc souvent qualifiés de modèles d'ordre réduit *POD-Galerkin*. D'autres alternatives sont envisageables lorsque les équations du fluide sont linéaires ou ont été linéarisées, ou encore lorsque les équations non-linéaires ne peuvent se ramener à un système polynomial. Lucia *et al.* [153] ont ainsi distingué en dehors des modèles d'ordre réduit *POD-Galerkin* les modèles d'ordre réduit obtenus par *projection directe* pour les systèmes linéaires<sup>7</sup>, et enfin la *projection sur le sous-espace* qui consiste à projeter le second membre  $f$  de l'équation (1.1) sur le sous-espace engendré par les modes *POD*<sup>8</sup>. L'inconvénient principal de cette dernière méthode provient du fait que l'évaluation du second membre requiert la reconstruction du champ non-réduit avant de pouvoir pratiquer la projection à chaque instant. Cette technique a néanmoins permis de reproduire des écoulements modélisés par les équations d'Euler non-linéaires [30, 75, 155, 178], mais le bénéfice en terme de coût de calcul est relativement faible en comparaison avec celui atteint avec l'approche *POD-Galerkin*. Lucia *et al.* [155] ont toutefois signalé que le pas de temps peut être considérablement augmenté pour l'intégration en temps du modèle d'ordre réduit puisque les hautes fréquences sont filtrées par la base tronquée des modes *POD*. Une comparaison des différentes approches a été proposée par Lucia et Beran [152] et une tentative d'extension au cas d'un système aéroélastique a été effectuée par Anttonen *et al.* [12, 13].

Le gain en temps de calcul est aussi substantiel pour les modèles d'ordre réduit construits à partir des systèmes dynamiques linéaires au moyen de la projection directe. Romanowski [189] a été le premier à développer un modèle d'ordre réduit des équations d'Euler linéarisées pour reproduire le comportement aéroélastique d'un profil portant dans un écoulement compressible. La méthode a ensuite été transposée par Kim [131] dans le domaine fréquentiel afin de pouvoir déterminer rapidement les limites de stabilité du système dynamique. Les progrès réalisés ont conduit à des applications pour des profils portants isolés, des voilures, puis des cascades de profils représentatives d'un étage de turbomachine [60, 104, 170]. Récemment, Lieu *et al.* [148] sont parvenus à réaliser le modèle d'ordre réduit de l'écoulement autour d'un avion complet pour tenir compte des effets aéroélastiques et la réponse sur une certaine gamme de fréquences d'excitation et de nombres de Mach a été obtenue au moyen de techniques d'interpolation [10, 147]. Concernant les systèmes aubagés, Epureanu *et al.* [68] ont développé un modèle d'ordre réduit à partir des équations potentielles complétées par une équation intégrale de couche limite au voisinage des parois pour tenir compte des effets visqueux. Enfin, Willcox [247] a comparé la projection sur la base des modes *POD* aux techniques d'identification basées sur l'équilibrage des moments de la fonction de transfert.

Avant conclure cet état de l'art des modèles d'ordre réduit construits à partir des modes *POD*, il faut signaler qu'un certain nombre d'applications ont aussi vu le jour en mécanique des structures. Kerschen *et al.* [130] ont ainsi livré une revue historique des applications en mécanique des structures en ce qui concerne le contrôle actif, la détection de dommages, la réduction de modèle, la validation de capteurs,... La projection sur la base des modes *POD* s'affiche ainsi comme une extension de la projection sur la base des modes propres pour les systèmes non-linéaires. Feeny et Kappagantu [74] et Kerschen et Golinval [129] ont fourni une interprétation physique des modes *POD* dans le cadre de problèmes de vibrations en les reliant notamment aux modes propres du système dynamique. Une conclusion importante de Feeny et Kappagantu [74] est que pour les vibrations libres non amorties d'un système linéaire symétrique avec une matrice de masse unitaire, les modes *POD* convergent vers les modes propres lorsque le nombre de clichés employé pour calculer les modes *POD* augmente. Des connexions existent aussi dans le cas de vibrations forcées. Toutefois, pour les systèmes linéaires la projection sur la base des modes *POD* ne présente pas d'intérêt puisque les modes propres constituent déjà une méthode d'approximation très efficace. En revanche, la méthode s'avère intéressante lorsque des non-linéarités sont présentes [8, 28] ou que le système est soumis à des vibro-impacts [19], même si Sampaio et Soize [200] ont souligné que la projection sur la base des modes propres est tout aussi performante dans le cadre de l'élastodynamique non-linéaire.

Pour être complet, il faut signaler pour terminer qu'il existe des variantes de la décomposition orthogonale propre et que celle-ci est parfois combinée avec d'autres méthodes de réduction. Une variante introduite par Aubry [16] sous le nom de *décomposition bi-orthogonale* permet de déterminer deux familles de fonctions propres qui portent respectivement le nom de *topos* pour les fonctions spatiales et de *chronos* pour les fonctions temporelles. La méthode est toutefois restreinte à l'analyse des écoulements puisque dans ce cas les fonctions temporelles  $a_j(t)$  sont obtenues au même titre que les fonctions spatiales et ne

---

<sup>7</sup>Dans ce cas la projection de Galerkin correspond à une simple projection matricielle analogue à celle effectuée avec la base des modes propres à l'équation (1.8).

<sup>8</sup>Voir les paragraphes 2.4.1 et 2.4.2 pour les détails relatifs à ces deux approches.

peuvent par conséquent pas être utilisées comme inconnues pour formuler un modèle d'ordre réduit. Hémon et Santi [109] ont ainsi appliqué la décomposition bi-orthogonale à l'étude d'un problème d'interaction fluide-structure pour analyser le champ de pression sur la paroi de corps non profilés puis pour réduire la taille des données permettant de représenter un champ de vitesse de vent qui est utilisé ensuite pour des simulations du comportement aéroélastique de tabliers de ponts. La méthode a aussi été employée par Stephen *et al.* [221] qui l'ont comparée une autre méthode d'analyse de signaux multivariés pour étudier un fluide baroclinique. Les modes POD ont parfois été calculés pour un ensemble de clichés composé de la réponse du système dynamique ainsi que sa dérivée en temps [110] et une variante de la POD connue sous le nom de « Smooth Orthogonal Decomposition » a par ailleurs vu le jour. La technique revient simplement à résoudre un problème aux valeurs propres généralisé dans lequel interviennent deux matrices de corrélations, l'une construite pour la réponse du système et l'autre pour sa dérivée en temps [29, 44]. Enfin, la projection sur la base des modes POD a parfois été combinée à des hypothèses de symétrie [195], au développement en séries de Volterra [154], ou encore à la méthode de troncature équilibrée [115, 193, 246].

La projection sur la base des modes POD présente l'avantage de pouvoir être appliquée à des systèmes dynamiques non-linéaires. L'expression du modèle d'ordre réduit résultant de la projection est particulièrement intéressante lorsque le second membre du système dynamique est polynomial. Dans le cas contraire il est difficile de formuler un modèle d'ordre réduit exploitable, à moins de recourir à la technique de projection sur le sous-espace dans laquelle le second membre doit être calculé au moyen du modèle d'ordre élevé ce qui limite fortement le gain en terme de réduction. Par ailleurs, il faut préciser que le modèle d'ordre réduit est susceptible d'être instable et doit donc être corrigé la plupart du temps. Enfin il s'agit à l'origine d'une méthode d'analyse de données si bien que les modes POD dépendent des clichés qui ont servi à construire la base de projection. Cette dépendance peut remettre en cause la capacité du modèle d'ordre réduit à reproduire des réponses correctes sur une certaine gamme de paramètres et la méthode demande donc à être éventuellement adaptée au moyen de techniques inspirées de [10, 147] par exemple.

### 1.3.5 Modes d'oscillation et d'interaction principaux

La principale limitation de la décomposition orthogonale propre est que les modes POD sont construits pour extraire des clichés les structures cohérentes *spatiales*, sans tenir compte à aucun moment de l'évolution temporelle du système dynamique dans le processus de calcul. Les modes POD sont donc certes optimaux au sens d'une approximation moindres carrés pour représenter « statiquement » les clichés à chaque instant, mais ils ne le sont pas pour reproduire le comportement dynamique du modèle d'ordre élevé. Hasselmann [107] a alors proposé de construire un modèle d'ordre réduit optimal à partir d'une forme générique, de telle manière que les modes ainsi que les coefficients du modèle d'ordre réduit fournissent la meilleure approximation de la solution de référence. À la différence de la projection sur la base des modes POD où seuls les modes étaient optimaux pour approcher les clichés, la méthode fournit ici simultanément les modes et les coefficients du système dynamique réduit. En procédant ainsi, le modèle d'ordre réduit est optimal au sens spatio-temporel et non plus seulement spatial. Les modes obtenus portent le nom de *modes d'interaction principaux* dans le cas général, et de *modes d'oscillation principaux* pour les systèmes linéaires<sup>9</sup>. Une revue des applications dans le cas linéaire a été réalisée par Von Storch *et al.* [242]. La méthode a été développée à l'origine pour l'étude de phénomènes climatiques tels que El Niño pour n'en citer qu'un. La nécessité de compléter l'analyse à partir des modes POD s'est imposée dans ce domaine puisque les systèmes dynamiques mis en jeu présentent des échelles de temps et d'espace très diverses dans lesquelles les caractéristiques spatio-temporelles sont fondamentales.

Les variables dynamiques du système (1.2) sont alors supposées s'écrire sous la forme

$$\mathbf{x}(t) = \sum_{j=1}^q a_j(t) \tilde{\varphi}^{(j)} + \boldsymbol{\rho} = \tilde{\Phi} \mathbf{a}(t) + \boldsymbol{\rho}, \quad (1.35)$$

où  $\tilde{\varphi}^{(j)}$  désigne un mode d'interaction principal,  $a_j$  est l'amplitude modale pour laquelle le modèle d'ordre réduit est formulé dans ce qui suit, et  $\boldsymbol{\rho}$  est le résidu de l'erreur d'approximation. La matrice  $\tilde{\Phi}$  de taille  $N_v \times q$  regroupe les  $q$  modes d'interaction principaux qui décrivent le modèle d'ordre réduit.

<sup>9</sup>Les acronymes PIPs et POPs sont généralement employés pour abrégier les expressions anglaises « Principal Interaction Patterns » et « Principal Oscillation Patterns ».

Les coefficients d'expansion  $\mathbf{a}^{(m)}$  qui minimisent le résidu  $\rho$  par rapport à une certaine métrique  $\mathbf{M}$  dans la décomposition (1.35) sont solutions du système linéaire approché  $\mathbf{x}^{(m)} \approx \tilde{\Phi} \mathbf{a}^{(m)}$ , où  $\mathbf{x}^{(m)}$  est un cliché à l'instant  $t_m$ . Dans le cas général, le problème n'a pas de solution exacte et il faut recourir à des méthodes de résolution approchées, telles que celles proposées dans l'annexe A.3. La solution des équations normales de la méthode des moindres carrés avec la métrique  $\mathbf{M}$  conduit à l'expression

$$\mathbf{a}^{(m)} \approx \tilde{\Phi}^+ \mathbf{x}^{(m)} = \left( \tilde{\Phi}^T \mathbf{M} \tilde{\Phi} \right)^{-1} \tilde{\Phi}^T \mathbf{M} \mathbf{x}^{(m)}. \quad (1.36)$$

Dans un second temps, il est nécessaire de choisir la forme générique du modèle d'ordre réduit censée pouvoir décrire le comportement du système non réduit. Ce choix résulte d'une certaine connaissance de la physique du problème modélisé ou d'une suite de raisonnements logiques concernant les propriétés que le modèle d'ordre réduit doit pouvoir reproduire. Un choix particulièrement judicieux consiste à utiliser l'expression formelle de la projection de Galerkin des équations du modèle d'ordre élevé pour des modes  $\tilde{\varphi}^{(j)}$  qui seront déterminés comme les modes d'interaction principaux. D'une manière générale, on note  $\dot{\mathbf{a}} = \mathbf{f}(\mathbf{a}, \mathcal{X})$  le système d'équations gouvernant les amplitudes modales  $\mathbf{a}$  en fonction d'un certain nombre de paramètres  $\mathcal{X}$  qui correspondent aux coefficients des équations différentielles ordinaires du modèle d'ordre réduit. À la différence des amplitudes  $\mathbf{a}^{(m)}$ , les amplitudes  $\mathbf{a}$  sont celles calculées à l'aide du système dynamique réduit. Les modes optimaux regroupés dans la matrice  $\tilde{\Phi}$  et les coefficients optimaux  $\mathcal{X}$  du modèle d'ordre réduit sont alors déterminés simultanément de façon à minimiser la fonctionnelle

$$\begin{aligned} \mathcal{Q}(\tilde{\Phi}, \mathcal{X}) &= E \left[ \left( \dot{\mathbf{a}}(t_m) - \dot{\mathbf{a}}^{(m)} \right)^T \tilde{\mathbf{M}} \left( \dot{\mathbf{a}}(t_m) - \dot{\mathbf{a}}^{(m)} \right) \right] \\ &= E \left[ \left( \mathbf{f}(\mathbf{a}(t_m), \mathcal{X}) - \tilde{\Phi}^+ \dot{\mathbf{x}}^{(m)} \right)^T \tilde{\mathbf{M}} \left( \mathbf{f}(\mathbf{a}(t_m), \mathcal{X}) - \tilde{\Phi}^+ \dot{\mathbf{x}}^{(m)} \right) \right], \end{aligned} \quad (1.37)$$

définie comme l'erreur moyenne entre les dérivées des amplitudes modales  $\dot{\mathbf{a}}$  produites par le modèle d'ordre réduit et celles de référence  $\dot{\mathbf{a}}^{(m)}$  associées aux modes d'interaction principaux par la relation (1.36). L'opérateur  $E[\cdot]$  est un opérateur de moyenne en temps identique à celui employé dans la formulation POD et dont l'objectif est de transcrire l'effet de moyenne d'ensemble statistique sous l'hypothèse d'ergodicité du système. Hasselmann [107] conseillait de choisir pour la métrique  $\mathbf{M}$  l'inverse de la matrice de covariance d'un ensemble de clichés  $\mathbf{x}^{(m)}$  et pour  $\tilde{\mathbf{M}}$  l'inverse de la matrice de covariance de la dérivée temporelle des clichés  $\dot{\mathbf{x}}^{(m)}$ .

La méthode fournit alors les modes  $\tilde{\varphi}^{(j)}$  optimaux et les coefficients  $\mathcal{X}$  optimaux du modèle d'ordre réduit pour représenter à la fois les données de départ issues du modèle non réduit et la réponse fournie par le modèle d'ordre réduit. Toutefois, le problème d'optimisation à résoudre est extrêmement lourd puisqu'il faut déterminer, en plus de tous les coefficients du modèle d'ordre réduit, l'ensemble des degrés de liberté pour les modes d'interaction principaux. Or si le modèle d'ordre réduit est construit pour un domaine fluide comportant un nombre important de cellules de contrôle, la résolution n'est pas possible en pratique. C'est la raison pour laquelle Kwasniok [137] a proposé de chercher des modes d'interaction principaux qui soient des combinaisons linéaires des modes POD. Le principe consiste à calculer dans un premier temps les  $q_\varphi$  premiers modes POD  $\varphi^{(j)}$  associés à la réponse du système. Ensuite, le problème d'optimisation (1.37) est formulé pour une fonctionnelle  $\mathcal{Q}'(\lambda, \mathcal{X})$ , où les coefficients  $\lambda$  sont ceux de la décomposition linéaire des modes d'interaction principaux sur les modes POD tels que :  $\tilde{\varphi}^{(j)} = \sum_{n=1}^{q_\varphi} \lambda_{nj} \varphi^{(n)}$ . En procédant de la sorte, le nombre de paramètres à optimiser pour minimiser la fonctionnelle  $\mathcal{Q}'$  est nettement réduit et la résolution devient possible même pour des systèmes comportant un nombre très important de degrés de liberté.

Deux remarques importantes distinguent les modèles d'ordre réduit basés sur les modes d'interaction principaux de ceux utilisant les modes POD. Les modes d'interaction principaux obtenus ne sont pas nécessairement orthogonaux à moins d'imposer une condition supplémentaire par le biais d'une contrainte dans le problème d'optimisation à résoudre par exemple, alors que par construction, les modes POD sont définis comme des vecteurs propres orthogonaux entre eux. Ensuite, la seconde remarque est que les modes d'interaction principaux peuvent convenir à un système d'équations différent — éventuellement simplifié — de celui du modèle d'ordre élevé puisque les coefficients et les modes s'adaptent à la réponse de référence alors que la projection sur la base des modes POD doit nécessairement être pratiquée pour les mêmes équations. Cette dernière remarque signifie par ailleurs que les modes d'interaction principaux peuvent conduire à des modèles d'ordre réduit non physiques lorsque la forme générique des équations est choisie a priori.

### 1.3.6 Équilibrage harmonique

Dans le cas d'écoulements dominés par un comportement périodique, l'utilisation d'une décomposition en série de Fourier de la solution du système dynamique s'avère très intéressante. Le principe consiste à remplacer la décomposition de Fourier dans les équations du système dynamique puis à équilibrer les termes de fréquences identiques. Le système dynamique est alors remplacé par un ensemble d'équations algébriques non-linéaires stationnaires portant sur les coefficients de Fourier. La résolution de ces équations peut être menée à l'aide des techniques d'accélération de la convergence développées pour les simulations stationnaires. La méthode, connue sous le nom d'*équilibrage harmonique* a d'abord été développée pour analyser le comportement harmonique des équations différentielles ordinaires, puis a plus récemment été appliquée à la résolution de problèmes d'aérodynamique par Hall *et al.* [103].

Le principe consiste donc à décomposer le vecteur d'état du système dynamique (1.1) sur les  $n_{\text{HB}}$  premières harmoniques de la manière suivante :

$$x(t) = \sum_{n=-n_{\text{HB}}}^{n_{\text{HB}}} \hat{x}_n e^{j\omega n t}. \quad (1.38)$$

Le terme pour l'harmonique  $n = 0$  correspond au champ moyen. L'introduction de la décomposition précédente dans l'équation (1.1) du système dynamique conduit alors à la relation

$$\sum_{n=-n_{\text{HB}}}^{n_{\text{HB}}} j\omega n \hat{x}_n e^{j\omega n t} = f \left( \sum_{n=-n_{\text{HB}}}^{n_{\text{HB}}} \hat{x}_n e^{j\omega n t}, \mu \right), \quad (1.39)$$

et en développant le second membre on parvient à l'expression

$$\sum_{n=-n_{\text{HB}}}^{n_{\text{HB}}} j\omega n \hat{x}_n e^{j\omega n t} = \sum_{n=-n_{\text{HB}}}^{n_{\text{HB}}} \hat{f}_n(\hat{x}_n, \mu) e^{j\omega n t}. \quad (1.40)$$

dans laquelle les termes de même fréquence ont été regroupés. Dans ce cas les termes de fréquence identique sont identifiés entre eux, ce qui conduit à  $2n_{\text{HB}} + 1$  relations algébriques portant sur les coefficients de Fourier  $\hat{x}_n$  qui s'écrivent

$$j\omega n \hat{x}_n = \hat{f}_n(\hat{x}_n, \mu), \quad \forall n \in \llbracket -n_{\text{HB}}; n_{\text{HB}} \rrbracket. \quad (1.41)$$

Enfin, en notant  $\mathbf{N} = \text{diag}(-n_{\text{HB}}, \dots, 0, \dots, n_{\text{HB}})$  la matrice des indices des harmoniques,  $\hat{\mathbf{x}} = [\hat{x}_{-n_{\text{HB}}}, \dots, \hat{x}_0, \dots, \hat{x}_{n_{\text{HB}}}]$  le vecteur des coefficients de Fourier et  $\hat{\mathbf{f}} = [\hat{f}_{-n_{\text{HB}}}(\hat{\mathbf{x}}, \mu), \dots, \hat{f}_0(\hat{\mathbf{x}}, \mu), \dots, \hat{f}_{n_{\text{HB}}}(\hat{\mathbf{x}}, \mu)]$  le vecteur des seconds membres, l'expression (1.41) se réécrit sous la forme matricielle

$$j\omega \mathbf{N} \hat{\mathbf{x}} = \hat{\mathbf{f}}. \quad (1.42)$$

Bien que le principe soit simple, la mise en œuvre est difficile lorsqu'il s'agit de traiter les équations d'Euler ou de Navier-Stokes puisque la détermination des termes  $\hat{f}_n$  est fastidieuse et coûteuse en terme de temps de calcul. En outre, l'extension de la méthode telle qu'elle a été introduite ici est difficilement envisageable pour les écoulements turbulents. Hall *et al.* [103] ont donc proposé une technique alternative pour évaluer le second membre. Celle-ci consiste à évaluer le vecteur d'état  $x$  du système dynamique à  $2n_{\text{HB}} + 1$  instants différents régulièrement répartis sur une période d'oscillation du système. Les solutions sont alors regroupées dans un vecteur  $\mathbf{x}^* = [x(t_1), \dots, x(t_{2n_{\text{HB}}+1})]$  auquel on applique une transformation de Fourier discrète. Le vecteur des coefficients de Fourier  $\hat{\mathbf{x}}$  est alors relié au vecteur  $\mathbf{x}^*$  par l'opérateur de la transformée de Fourier discrète noté  $\mathbf{E}$  tel que  $\hat{\mathbf{x}} = \mathbf{E} \mathbf{x}^*$ . Le système (1.42) devient alors  $j\omega \mathbf{N} \mathbf{E} \mathbf{x}^* = \hat{\mathbf{f}}$  et finalement, puisque la matrice de la transformée de Fourier discrète est inversible, l'équation est transposée dans le temps pour obtenir l'expression

$$j\omega (\mathbf{E}^{-1} \mathbf{N} \mathbf{E}) \mathbf{x}^* = \mathbf{f}^*. \quad (1.43)$$

L'opérateur  $\mathbf{E}^{-1} \mathbf{N} \mathbf{E}$  correspond en fait à l'approximation spectrale de la dérivée temporelle et le vecteur  $\mathbf{f}^*$  est défini par  $\mathbf{f}^* = \mathbf{E}^{-1} \hat{\mathbf{f}}$ . L'avantage de cette dernière expression est que le second membre se calcule

facilement avec le code de calcul habituellement employé pour le modèle d'ordre élevé. Hall *et al.* [103] ont finalement introduit dans l'équation (1.43) un terme pseudo-temporel afin de pouvoir résoudre le système au moyen des algorithmes de résolution déjà disponibles. Le système s'écrit alors

$$\frac{\partial \mathbf{x}^*}{\partial \tau} + j\omega(\mathbf{E}^{-1} \mathbf{N} \mathbf{E}) \mathbf{x}^* = \mathbf{f}^*. \quad (1.44)$$

De cette manière, le système est intégré vers un pseudo-état stationnaire périodique qui correspond aux coefficients de Fourier de la décomposition (1.38).

La méthode permet de conserver les non-linéarités et a été employée en particulier pour modéliser les écoulements dans les turbomachines. Ainsi, Hall *et al.* [103] ont appliqué avec succès cette technique pour effectuer des calculs instationnaires avec les équations de Navier-Stokes moyennées pour des cascades de profils en flottement harmonique. Thomas *et al.* [227] ont employé l'équilibrage harmonique pour évaluer le cycle limite d'oscillation d'un profil portant en écoulement visqueux transsonique. McMullen *et al.* [164, 165] ont quant à eux cherché à résoudre directement les équations (1.42) à l'aide d'un code fréquentiel. Liu *et al.* [149] se sont quant à eux intéressés à la réponse aéroélastique d'un profil portant dont le déplacement est gouverné par un ressort de rigidité cubique. La réponse du système est alors comparée lorsque celle-ci est calculée par la méthode dite « classique » dans laquelle le système porte directement sur les coefficients de Fourier gouvernés par l'équation (1.41) et avec la méthode dite de « grande dimension » dans laquelle l'équation (1.44) est résolue avec un terme pseudo-temporel. L'utilisation de la transformée de Fourier discrète pour résoudre le problème en temps allège considérablement la résolution, d'autant plus lorsque le nombre d'harmoniques retenues est grand. Cependant, la seconde version de l'équilibrage harmonique est moins performante que la version classique puisque pour une précision équivalente, le double d'harmoniques doit être retenu. Les auteurs attribuent cet accroissement du nombre d'harmoniques au degré de non-linéarité introduit et précisent que le nombre d'harmoniques supplémentaires est inférieur si la non-linéarité est seulement quadratique. Enfin, la méthode a aussi été employée par Gopinath *et al.* [95] pour modéliser l'écoulement instationnaire (i) dans un compresseur à un seul étage en 3D et (ii) dans un compresseur multi-étage en 2D. Dans ce cas, la difficulté consiste à combiner des fréquences associées aux différents étages qui ne sont pas nécessairement multiples entre elles.

Des gains en temps de calcul significatifs sont obtenus avec cette méthode puisque seuls  $2n_{\text{HB}} + 1$  coefficients de Fourier sont calculés. La principale limitation de la méthode est toutefois le fait que l'écoulement doit nécessairement être harmonique et que les fréquences des oscillations doivent être connues a priori. Toutefois certains développements sont en cours pour éviter cette dernière contrainte [94].

La méthode d'équilibrage harmonique peut aussi servir à construire un modèle d'ordre réduit non-linéaire à partir d'un solveur fluide linéarisé, en pratiquant un développement de Taylor du résidu [228]. Le développement obtenu est alors projeté sur les modes POD pour aboutir à un modèle d'ordre réduit non-linéaire, tronqué ici à l'ordre deux. Le principal avantage est que non seulement le modèle d'ordre réduit est capable de reproduire des phénomènes non-linéaires, mais il permet en outre d'obtenir une réponse paramétrique si les modes POD ont été calculés à partir de clichés obtenus pour différents paramètres.

### 1.3.7 Réduction sur une base construite *a priori*

Enfin des travaux récents ont conduit à la construction de modèles d'ordre réduit pour lesquels la base de projection est fabriquée *a priori* de manière itérative. Le terme *a priori* sous-entend qu'il n'est pas nécessaire de générer au préalable un ensemble de clichés, d'où un gain en temps de calcul pour la construction du modèle d'ordre réduit. La méthode de construction de la base *a priori* introduite par Ryckelynck [198, 199] consiste, comme la plupart des autres méthodes rencontrées jusqu'à présent, à effectuer une décomposition linéaire analogue à celle donnée à l'équation (1.32) pour le vecteur d'état du système dynamique. Une projection formelle du système dynamique sur la base à déterminer est pratiquée et le modèle d'ordre réduit résultant a ainsi une forme identique à celle donnée à l'équation (1.34) pour la décomposition sur la base des modes POD lorsque le second membre du système dynamique est polynomial.

La base de projection recherchée est dans un premier temps initialisée avec un vecteur arbitraire, comme par exemple la condition initiale imposée au système dynamique ou la solution stationnaire. La base *a priori* est donc formée initialement d'un seul vecteur :  $\Phi^{(k=1)} = \{\varphi^{(1)}\}$ . Le système dynamique réduit est alors intégré en temps sur une durée  $T_c$  au cours de laquelle le résidu  $R$  du système dynamique est évalué. Celui-ci est défini à partir de l'équation (1.1) du système dynamique par

$$R(t) = \sum_{j=1}^{n_q} \dot{a}_j(t) \varphi^{(j)} - f \left( \sum_{j=1}^{n_q} a_j(t) \varphi^{(j)}, \mu \right). \quad (1.45)$$

Le nombre  $n_q$  de vecteurs  $\varphi^{(j)}$  contenus dans la base de projection est au départ égal à un et le résidu du système dynamique est évidemment élevé lorsque la base a été initialisée avec un vecteur arbitraire. Si la norme du résidu dépasse un certain seuil  $\varepsilon_R$  pour au moins un instant au cours de l'intervalle de temps  $[t_0; t_0 + T_c]$ , cela signifie que les vecteurs  $\varphi^{(j)}$  de la base a priori ne sont pas appropriés et la base de projection doit donc être améliorée et/ou complétée.

La base est dans un premier temps *améliorée* en procédant à une décomposition orthogonale propre des amplitudes modales  $a_j$  intégrées sur l'intervalle de temps  $[t_0; t_0 + T_c]$ . Seuls les  $p$  premiers modes POD associés aux plus grandes valeurs propres du problème (1.31) sont conservés. En notant alors  $\mathbf{V}^{(j)}$  les  $p$  modes POD de dimension  $n_q$ , les vecteurs de la base a priori sont corrigés pour tout  $j = 1, \dots, p$  en

$$\tilde{\varphi}^{(j)} = \sum_{i=1}^{n_q} \varphi^{(i)} V_i^{(j)}. \quad (1.46)$$

La base a priori améliorée est donc constituée de  $p \leq n_q$  vecteurs tels que  $\tilde{\Phi}^{(k)} = \{\tilde{\varphi}^{(1)}, \dots, \tilde{\varphi}^{(p)}\}$ . Finalement la base doit être enrichie par des vecteurs supplémentaires pour améliorer le contenu de la projection. On choisit pour cela d'alimenter la base avec le résidu  $R(t^+)$  où  $t^+$  est le premier instant au cours de l'intégration en temps du modèle d'ordre réduit où le résidu dépasse le seuil de tolérance  $\varepsilon_R$  imposé. Plus généralement, la base est enrichie avec le sous-espace de Krylov à droite de dimension  $m$  défini à l'équation (1.20). Le vecteur  $\mathbf{d}$  correspond alors à la discrétisation  $\mathbf{R}^+$  du résidu  $R(t^+)$  et la matrice  $\mathbf{M}$  est la matrice jacobienne  $\mathbf{J}$  du système dynamique. La nouvelle base enrichie et améliorée

$$\Phi^{(k+1)} = \left\{ \tilde{\Phi}^{(k)}, \mathcal{K}_m(\mathbf{J}, \mathbf{R}^+) \right\} \quad (1.47)$$

comprend au final les  $p$  vecteurs améliorés auxquels viennent s'ajouter les  $m$  vecteurs supplémentaires apportés par l'enrichissement avec le sous-espace de Krylov. En pratique on choisit  $m = 1$  afin de ne pas avoir à évaluer la matrice jacobienne du système dynamique. Les vecteurs de la base  $\Phi^{(k+1)}$  doivent être normés et éventuellement orthogonalisés. Le processus est poursuivi de façon itérative jusqu'à ce que la base enrichie soit en mesure de reproduire correctement la réponse du modèle d'ordre élevé.

La méthode a été mise en œuvre par Ryckelynck [198, 199] pour des problèmes de thermomécanique. Verdon [238] a quant à lui appliqué la méthode aux équations de Navier-Stokes d'un écoulement incompressible couplées au transport de particules. Les gains en terme de réduction du nombre de degrés de liberté et de temps de calcul sont considérables. Le seul reproche qui peut être formulé à propos de cette méthode est que la construction du modèle d'ordre réduit est intrinsèquement liée au code de calcul employé pour résoudre le modèle d'ordre élevé afin de pouvoir évaluer le résidu  $R(t)$ .

Dans le même esprit, Sirisup *et al.* [217] ont proposé une méthode dite d'*intégration projective* pour accélérer la résolution du système dynamique. L'idée est de calculer sur un petit intervalle de temps la solution du modèle d'ordre élevé afin de pouvoir extraire les modes POD ainsi que les amplitudes modales associées aux clichés prélevés sur la durée simulée. Une intégration projective est alors réalisée : elle consiste simplement à évaluer les amplitudes modales pour un instant ultérieur par des techniques classiques d'intégration en temps des équations différentielles ordinaires. La méthode alterne donc entre la résolution du modèle d'ordre élevé et du modèle d'ordre réduit et permet donc d'aboutir à des simulations stables sur de très longues durées, ce qui n'était pas possible avec des modèles POD-Galerkin classiques [215, 216].

### 1.3.8 Méthodes de réduction alternatives et dépendance avec un paramètre

L'état de l'art effectué dans ce paragraphe n'est pas totalement exhaustif et il existe encore bien d'autres techniques de réduction — peu ou pas encore employées en mécanique des fluides — qui se doivent d'être mentionnées ici puisqu'elles présentent un certain potentiel.

**Méthodes de Galerkin non-linéaires et variétés inertielles** L'obtention de modèles d'ordre réduit non-linéaires s'effectue très souvent au moyen d'une projection « classique » ou « linéaire » de Galerkin, mais récemment des méthodes de Galerkin non-linéaires issues de la théorie des variétés inertielles pour les systèmes dynamiques sont apparues. Le principe consiste simplement à construire le modèle d'ordre réduit en tenant compte de la dynamique rapide de certains modes [158]. Comme le notent Steindl et Troger [220], cette approche peut permettre l'élaboration de modèles d'ordre réduit plus performants puisque l'intégration en temps reste stable à long terme et ce, même avec très peu de modes. L'efficacité de la méthode a en particulier été étudiée pour les équations de Kuramoto-Sivashinsky et de Burgers par Jones *et al.* [121].

**Décompositions alternatives** D'autres types de décompositions plus marginales ont aussi été étudiés. Ainsi, l'*analyse archétypale* développée par Cutler et Breiman [54] apparaît comme une extension intéressante de la décomposition orthogonale propre puisqu'elle fournit des modes optimaux pour des ensembles de données non nécessairement elliptiques alors que les modes POD peuvent justement s'interpréter comme les axes principaux d'une ellipsoïde. Une variante de la méthode décrite par Stone et Cutler [222] pour suivre les structures mobiles telles que les ondes progressives demande à être étudiée plus en détail puisqu'elle est prédisposée aux systèmes où les conditions limites sont cycliques comme pour les turbomachines. La méthode semble robuste puisque Adrover et Giona [2] sont parvenus à reproduire les réponses de systèmes dynamiques dans lesquelles des bifurcations importantes interviennent.

La base de projection peut être construite à l'aide des *tessellations de Voronoï*. Les diagrammes de Voronoï ont été à l'origine développés pour produire une décomposition de l'espace en polygones à partir d'un ensemble de points. De nombreuses applications de la méthode sont fournies dans [62] et la méthode a naturellement été employée pour discrétiser le domaine spatial sur lequel un système dynamique est défini [63]. Plus récemment, un modèle d'ordre réduit pour les équations de Navier-Stokes a été construit par Burkardt *et al.* [37]. Dans cet article, les performances du modèle d'ordre réduit sont comparées à celles obtenues avec un modèle d'ordre réduit dont la base de projection est constituée des modes POD. Les résultats ne permettent cependant pas de favoriser une méthode plus qu'une autre.

Enfin, citons aussi les travaux de Ammar *et al.* [9] qui ont introduit une décomposition des variables d'état afin de réduire le système dynamique non-linéaire à un ensemble d'équations linéaires dans chaque direction de l'espace, ou encore la *décomposition spectrale généralisée* de Nouy [173] pour les problèmes gouvernés par un ensemble d'équations aux dérivées partielles stochastiques.

Pour conclure, Chinesta *et al.* [45] ont fait le point sur différentes méthodes de réduction a priori et a posteriori et ont fait le lien, toujours dans l'optique de réduire les coûts de calcul, avec d'autres techniques envisageables comme l'homogénéisation ou la parallélisation de codes de calcul.

**Réponse paramétrique** Les méthodes décrites dans cette partie visent à déterminer une base de projection optimale pour décrire le système dynamique. La construction de la base s'effectue souvent à partir d'un ensemble fini de  $M$  échantillons formant une base de données parfois appelée *base d'apprentissage* et notée

$$X = \{x^{(1)}, \dots, x^{(M)}\}. \quad (1.48)$$

Le choix des échantillons dans la base de données est crucial pour la construction de la base de projection puisque ce sont les seuls éléments qui apportent des informations caractéristiques du système. Plusieurs manières de choisir les échantillons peuvent être envisagées. Si le processus est stationnaire, l'état du système ne dépend que du paramètre  $\mu$  et par conséquent l'ensemble  $X$  est constitué des réponses du modèle d'ordre élevé  $x(\mu_i)$  pour  $M$  valeurs différentes du paramètre. Pour l'application qui nous intéresse ici, le système est instationnaire et par conséquent les échantillons peuvent être prélevés à des instants  $t_i$  différents et pour des valeurs différentes  $\mu_j$  du paramètre  $\mu$ . Les  $M$  échantillons correspondent donc à l'état  $x(t_i, \mu_j)$ . Enfin, il est possible de considérer le temps comme un paramètre et de ne considérer pour les échantillons que les états  $x(t_i)$ . Le choix de la base d'apprentissage peut faire l'objet d'un *plan d'expérience* qui vise à déterminer un ensemble d'échantillons qui soit le plus représentatif possible du modèle d'ordre élevé sur un intervalle de temps ou une gamme de paramètres donnés. Mentionnons juste que l'ensemble  $X$  peut être obtenu par un tirage aléatoire par une méthode de type Monte-Carlo, ou par des techniques d'échantillonnage stratifié tels que les hypercubes latins. L'échantillonnage peut aussi être pratiqué au moyen d'une décomposition en tessellations de Voronoï de l'espace paramétrique, ou à l'aide d'un

algorithme glouton [36]. Weller *et al.* [244] ont quant à eux proposé de combiner les deux approches.

Dans ce document on se concentre sur la construction d'un modèle d'ordre réduit capable de reproduire la réponse temporelle du modèle d'ordre élevé. Par conséquent, le paramètre  $\mu$  ne représente pour l'instant que le temps, mais il pourrait correspondre dans le futur à un paramètre de l'écoulement comme le nombre de Mach par exemple. La base de projection devra donc être construite de manière à tenir compte de ces nouvelles variations paramétriques. De nombreux travaux ont contribué à la construction de modèles d'ordre réduit paramétriques. Citons en particulier la construction de bases paramétriques à partir des sous-espaces de Lagrange, de Taylor ou de Hermite [118], les extensions proposées par Christensen *et al.* [46] et Jørgensen *et al.* [122] pour compléter les éléments de la base de projection, ou encore les méthodes d'interpolations récemment développées par l'équipe de Farhat pour les écoulements compressibles linéarisés [10, 147, 148]. Le problème de la construction d'une base de projection adéquate pour reproduire la réponse du système sur une gamme de paramètres est clairement expliqué par Rozza *et al.* [196] dans le cas de problèmes linéaires, mais la méthode s'applique de manière plus générale aux systèmes non-linéaires. Les auteurs décrivent en détail le principe de construction d'une base paramétrique et définissent la manière de choisir les clichés à partir d'algorithmes gloutons et d'estimateurs d'erreurs a posteriori. L'objectif est à terme d'aboutir à un modèle d'ordre réduit paramétrique dans lequel les paramètres sont différents du temps. De Vuyst et Audouze [55] ont quant à eux suggéré de combiner des modèles réduits construits par projection sur la base des modes POD avec des méthodes d'interpolation des amplitudes modales pour obtenir la réponse sur une gamme de paramètres. Les méthodes d'interpolation en question sont celles présentées dans le paragraphe 1.2 sous le nom de méthodes de réduction de type surface de réponse.



## 1.4 Synthèse et justification du choix de la méthode POD

Dans le paragraphe précédent, un certain nombre de méthodes physiques de réduction de modèle ont été présentées, mais seules quelques unes d'entre elles sont de bons candidats pour construire un modèle d'ordre réduit non-linéaire de l'écoulement aérodynamique compressible dans les étages d'un compresseur ou d'une turbine de turbomachine. Pour nous aider dans le choix de la méthode de projection adoptée, on résume dans le tableau 1.1 plusieurs caractéristiques des principales méthodes décrites dans ce qui précède.

Syst. dyn.	Type de réduction	BdD	Intr.	Spécificité
Linéaire	Décomposition en modes propres			Calcul des vecteurs propres de l'opérateur élastique ou aérodynamique.
	Balanced Truncation			Calcul des vecteurs propres diagonalisant les matrices grammiennes de contrôlabilité et d'observabilité.
	Approximation de Krylov			Calcul d'une base engendrant les mêmes espaces de Krylov que ceux associés aux moments de la fonction de transfert.
	Identification de la matrice de Hankel			Identification des matrices du système réduit à partir de réponses impulsionnelles.
Non-linéaire	Développement en série de Volterra			Calcul des noyaux de Volterra à partir de réponses impulsionnelles.
	Décomposition orthogonale propre			Calcul des vecteurs propres minimisant l'erreur de projection orthogonale.
	Décomposition en modes d'interaction principaux			Calcul des vecteurs et des coefficients optimaux pour le système dynamique.
	Équilibrage harmonique			Calcul de la réponse pour les coefficients de la décomposition en série de Fourier.
	Réduction a priori			Calcul itératif de la base par amélioration et enrichissement avec le résidu.

**Tableau 1.1 :** Comparaison des principales méthodes de réduction de modèle physiques. La première colonne indique à quel type de système dynamique (linéaire ou non-linéaire) s'applique la méthode de réduction proposée dans la deuxième colonne. Les cases colorées dans la troisième et la quatrième colonne signifient respectivement que la construction du modèle d'ordre réduit nécessite une base de données (BdD) ou est intrusive (Intr.). Enfin la dernière colonne décrit très brièvement le principe de la méthode.

Les méthodes de réduction valables pour les systèmes non-linéaires le sont bien évidemment pour les systèmes linéaires qui en constituent une simplification. La première partie des méthodes dédiées aux systèmes linéaires est écartée d'emblée puisque l'application qui nous intéresse ici met en œuvre des systèmes dynamiques non-linéaires. Parmi les méthodes disponibles pour les systèmes non-linéaires, un certain nombre présente des atouts. Un critère supplémentaire pour orienter notre choix concerne la manière dont le modèle d'ordre réduit est construit. La détermination de la base de projection nécessite en général soit de s'introduire dans le code de calcul avec lequel le modèle d'ordre élevé est résolu, soit d'utiliser les résultats produits par ce code. Dans le premier cas on parle de méthodes *intrusives* puisqu'il est nécessaire au cours de la construction du modèle d'ordre réduit d'effectuer un certain nombre d'opérations avec les routines du code de calcul pour récupérer par exemple les matrices de masse et de rigidité d'un système mécanique, ou l'opérateur aérodynamique discret par exemple, ou encore pour calculer le résidu. La seconde classe de méthode n'est pas intrusive mais nécessite de générer une *base de données* caractérisant le système dynamique au moyen d'un certain nombre de réponses. Dans le cadre de cette thèse, on se fixe l'objectif de construire un modèle d'ordre réduit du système dynamique qui soit indépendant du code de calcul et par conséquent les méthodes intrusives sont éliminées.

Le choix de la base de projection est encore large puisque en plus des modes POD et des modes d'interaction principaux recensés dans le tableau 1.1, les modes produits par l'analyse archétypale, ou encore les générateurs associés aux tessellations de Voronoï conviennent. Le développement en série de Volterra est un cas à part puisque même si il permet théoriquement de représenter des systèmes non-linéaires au moyen d'un faible nombre de noyaux, son emploi en tant que modèle d'ordre réduit n'est pas encore formalisé. Notre choix s'oriente vers la construction d'un modèle d'ordre réduit au moyen de la décomposition ortho-

gonale propre afin de bénéficier des nombreux résultats déjà obtenus pour les écoulements incompressibles et des rares travaux effectués pour les écoulements compressibles.

Écoulement	Étude	L	NL	Ael	Projection
Potentiel	Anttonen <i>et al.</i> [12, 13]				Sous-espace
	Epureanu <i>et al.</i> [68]				Discrète
Euler	Barone <i>et al.</i> [23]				Discrète
	Beran <i>et al.</i> [30]				Sous-espace
	Feng et Soulaïmani [75]				Sous-espace
	Hall <i>et al.</i> [104]				Discrète
	Lieu et Farhat [147]				Discrète
	Lucia <i>et al.</i> [155]				Sous-espace
	Lucia et Beran [152]				POD-Galerkin
	Mortchéléwicz [170]				Discrète
	Pettit et Beran [178]				Sous-espace
	Romanowski [189]				Discrète
Willcox [247]				Discrète	
Navier-Stokes isentropique	Gloerfelt [87]				POD-Galerkin
	Moreno <i>et al.</i> [169]				POD-Galerkin
	Rowley <i>et al.</i> [192, 194]				POD-Galerkin
Navier-Stokes	Bourguet <i>et al.</i> [34]				POD-Galerkin
	Gloerfelt [87]				POD-Galerkin
	Iollo <i>et al.</i> [117]				POD-Galerkin
	Mortchéléwicz [170]				Discrète
	Vigo [239, 240]				POD-Galerkin
	Étude actuelle				POD-Galerkin

**Tableau 1.2 :** Résumé de l'utilisation de modèles d'ordre réduit basés sur la décomposition orthogonale propre pour les écoulements d'un fluide compressible. Les deux premières colonnes indiquent le type d'écoulement pour lequel le modèle d'ordre réduit est construit et la référence de l'étude. Dans les colonnes suivantes, on précise si le modèle d'ordre réduit est linéaire (L) ou non-linéaire (NL) et si il existe une application pour l'aéroélasticité (Ael). Enfin dans la dernière colonne le type de projection employé pour obtenir le modèle d'ordre réduit est fourni.

Le tableau 1.2 résume les principaux travaux effectués pour la réduction de systèmes dynamiques au moyen de la décomposition orthogonale propre pour les écoulements d'un fluide compressible. Pour les applications en aéroélasticité, la grande majorité des travaux est effectuée avec les équations linéarisées d'Euler qui permettent d'employer la projection discrète. Quelques applications avec la projection sur le sous-espace ont été réalisées par Beran *et al.* [30] et Feng et Soulaïmani [75] mais comme cela a été dit plus tôt au paragraphe 1.3.4, la projection sur le sous-espace ne permet pas des gains en temps de calcul aussi intéressants que la projection discrète ou celle de Galerkin. À notre connaissance, aucun modèle d'ordre réduit non-linéaire des équations d'Euler ou de Navier-Stokes d'un fluide compressible n'a été développé pour des applications en aéroélasticité. Dans cette thèse, on se propose donc d'étendre la formulation du modèle d'ordre réduit POD-Galerkin non-linéaire développée par Vigo [239] afin de tenir compte d'effets aéroélastiques dus au déplacement d'une structure dans le fluide. La principale difficulté sera d'écrire formellement les équations du modèle d'ordre réduit obtenues par projection de Galerkin.

L'écriture formelle de cette projection sera ici combinée avec les modes POD de l'écoulement afin de construire le modèle d'ordre réduit, mais le choix d'autres vecteurs de base éventuellement plus appropriés est évidemment envisageable. En particulier, il paraît intéressant d'étudier par la suite la construction d'un modèle d'ordre réduit avec une base constituée des modes d'interaction principaux ou issus de l'analyse archétypale puisque ces méthodes permettent de déterminer des modes optimaux en adéquation avec les coefficients du modèle d'ordre réduit.



# Principe de construction d'un modèle d'ordre réduit POD-Galerkin

## Sommaire

<b>2.1</b>	<b>Introduction</b>	<b>36</b>
<b>2.2</b>	<b>Description de la décomposition orthogonale propre pour un système déterministe</b>	<b>37</b>
2.2.1	Décomposition orthogonale propre d'un ensemble de données	37
2.2.2	Détermination des modes POD par la méthode des clichés	41
2.2.3	Propriétés de la décomposition orthogonale propre	43
2.2.3.1	Orthogonalité des amplitudes modales en moyenne	43
2.2.3.2	Énergie associée aux modes POD	44
2.2.3.3	Troncature de la base POD et estimateur d'erreur	45
2.2.3.4	Héritage de propriétés par les modes POD	46
<b>2.3</b>	<b>Un exemple d'application de la POD en analyse de données</b>	<b>47</b>
<b>2.4</b>	<b>Principe de construction du modèle d'ordre réduit à partir de la base POD</b>	<b>50</b>
2.4.1	Construction d'un modèle d'ordre réduit linéaire par projection discrète	50
2.4.2	Construction d'un modèle d'ordre réduit implicite par projection sur le sous-espace	51
2.4.3	Construction d'un modèle d'ordre réduit polynomial par projection de Galerkin	52
2.4.4	Construction d'un modèle d'ordre réduit par identification des coefficients	53
<b>2.5</b>	<b>Décomposition orthogonale propre des systèmes multivariables</b>	<b>56</b>
2.5.1	Description de l'approche POD scalaire	56
2.5.2	Description de l'approche POD vectorielle	58
<b>2.6</b>	<b>Synthèse du chapitre 2</b>	<b>61</b>

### 2.1 Introduction

Après avoir décrit dans le paragraphe précédent un certain nombre de méthodes de réduction, notre choix s'est porté sur la décomposition orthogonale propre qui fournit une base adéquate pour construire un modèle d'ordre réduit selon les critères que nous nous sommes fixés dans le cadre de cette thèse. Avant d'entrer directement dans le vif du sujet, le principe de la décomposition orthogonale propre est d'abord détaillé afin de présenter le cadre théorique ainsi que les principales propriétés des modes POD qui constitueront la base de projection. Un exemple d'application simple sera fourni en conclusion de cette description afin de faire clairement ressortir le caractère optimal de la méthode ainsi que les liens étroits qui existent avec une approximation de type moindres carrés. Pour terminer, les différents types de projection sur la base des modes POD seront décrits et comparés à l'approche POD-Galerkin qui sera employée par la suite pour aboutir à un modèle d'ordre réduit non-linéaire. Enfin, la distinction entre les approches POD scalaire et vectorielle pour les systèmes multivariés sera rappelée et l'importance de la définition du produit scalaire sera évoquée.

## 2.2 Description de la décomposition orthogonale propre pour un système déterministe

Dans le paragraphe 1.3.4 du chapitre précédent, une brève discussion sur les liens existant entre plusieurs méthodes d'analyse de données étroitement liées à la décomposition orthogonale propre a été amorcée. Les travaux de Liang *et al.* [144] ont notamment permis d'établir des équivalences entre la décomposition de Karhunen-Loève, l'analyse en composantes principales et la décomposition en valeurs singulières sous l'hypothèse que les données auxquelles la décomposition est appliquée sont composées de vecteurs discrets aléatoires en nombre fini. Dans ce cas, les trois méthodes peuvent être regroupées sous l'appellation générique de décomposition orthogonale propre. L'utilisation de ces méthodes est alors justifiée dans le cadre du traitement de données statistiques, ou encore pour l'analyse d'un écoulement turbulent. Dans ce contexte, l'aspect stochastique des données est en effet avéré, si bien que les notions probabilistes invoquées dans le théorème de Karhunen-Loève ont un sens.

Cependant, l'objectif est ici d'employer ces méthodes pour des systèmes mécaniques *déterministes*. La méthode est donc formulée dans un tel contexte comme un problème d'optimisation dont la solution est la base optimale pour approcher en moyenne l'ensemble de données. De cette manière, la formulation s'affranchit de tout argument probabiliste.

### 2.2.1 Décomposition orthogonale propre d'un ensemble de données

On considère un système dynamique dont la solution sur l'intervalle de temps fini  $I_e = [t_0; t_0 + T_e] \subset \mathbb{R}_+$  est connue à  $M$  instants  $t_m$  pour  $m = 1, \dots, M$  tels que pour tout  $m$ ,  $t_m \in I_e$ . Ces solutions constituent les *clichés* auxquels la décomposition orthogonale propre va être appliquée pour en extraire les modes POD. Dans la présentation qui suit, chaque cliché est considéré comme un élément de l'espace de Hilbert  $H$  défini sur un corps  $\mathbb{K} = \mathbb{R}$  et muni de la norme  $\|\cdot\|_H$  induite par le produit scalaire  $\langle \cdot, \cdot \rangle_H$  tel que pour tout élément  $x \in H$  on ait  $\|x\|_H^2 = \langle x, x \rangle_H$ . En général l'espace de Hilbert est soit l'espace de Lebesgue  $H = L^2(\Omega)$  des fonctions de carré intégrable définies sur un domaine spatial  $\Omega \subset \mathbb{R}^d$  avec  $d = 1, 2$  ou  $3$ , soit l'espace euclidien  $H = \mathbb{R}^{N_v}$  des vecteurs discrets de dimension finie  $N_v$ . Dans le premier cas, cela conduit à formuler la décomposition orthogonale propre dans un espace de dimension infinie lorsque les solutions continues du système dynamique (1.1) sont considérées, alors que dans le second cas la méthode est plutôt formulée pour les solutions du système discrétisé (1.2) impliquant un nombre fini  $N_v$  de degrés de liberté. Dans les deux cas on supposera ici que les solutions du système dynamique ne sont connues qu'aux instants  $t_m$  ce qui conduit à adopter une formulation discrète en temps. On notera ensuite  $N = \dim(H)$  en prenant garde au fait que  $N$  est infini lorsque  $H = L^2(\Omega)$ . Enfin, l'ensemble des clichés introduit à l'équation (1.48) est désigné ici par l'ensemble  $X$  qui regroupe  $M$  solutions du système dynamique aux instants  $t_m \in I_e$  tel que

$$X = \{x(t_1), \dots, x(t_M)\} = \left\{ x^{(m)} \in H \right\}_{m=1}^M. \quad (2.1)$$

Le principe de la décomposition orthogonale propre consiste à trouver un sous-espace  $S \subset H$  de dimension finie  $q \ll N$  qui fournit la description « optimale » de l'ensemble  $X$ . Le sous-espace est alors caractérisé par une base de fonctions  $\Phi = \{\varphi^{(j)} \in H, j = 1, \dots, q\}$  qui l'engendrent :  $S = \text{vect}\{\varphi^{(1)}, \dots, \varphi^{(q)}\}$ . La détermination des vecteurs de base  $\varphi^{(j)}$  définit complètement le sous-espace  $S$  et par conséquent la décomposition orthogonale propre fournit une approximation  $\hat{x}^{(m)}$  des clichés sur cette base telle que

$$x^{(m)} \approx \hat{x}^{(m)} = \sum_{j=1}^q a_j^{(m)} \varphi^{(j)} \quad \forall m \in \llbracket 1; M \rrbracket. \quad (2.2)$$

Dans la décomposition précédente, les vecteurs  $\varphi^{(j)}$  correspondent aux modes POD et les coefficients  $a_j^{(m)}$  sont des scalaires caractérisant l'amplitude modale associée au mode  $\varphi^{(j)}$  pour le cliché  $x^{(m)}$ . La décomposition recherchée est « optimale » au sens où la base  $\Phi$  doit être meilleure que n'importe quelle autre base de dimension égale utilisée pour reconstruire les données de l'ensemble  $X$ . Pour assurer le caractère optimal de la base recherchée, on impose que la moyenne  $E[\cdot]$  du carré de l'erreur de projection des éléments de  $X$  sur le sous-espace  $S$  soit minimale, ce qui se traduit par la définition suivante [135].

**Définition 2.1 :** Étant donné un ensemble de clichés  $X = \{x^{(m)} \in H\}_{m=1}^M$  et un opérateur de moyenne  $E[\cdot]$ , la base des modes POD  $\Phi = \{\varphi^{(j)} \in H\}_{j=1}^q$  de dimension  $q$  est définie comme la solution du problème de minimisation

$$\Phi = \arg \min_{\{\varphi^{(j)}\}_{j=1}^q \subset H} E \left[ \left\| x^{(m)} - \sum_{j=1}^q \langle x^{(m)}, \varphi^{(j)} \rangle \varphi^{(j)} \right\|^2 \right] \quad \text{tel que} \quad \langle \varphi^{(i)}, \varphi^{(j)} \rangle = \delta_{ij}. \quad (2.3)$$

L'opérateur de moyenne  $E[\cdot]$  agit sur (une fonction) des éléments de l'ensemble de données  $X$ . Il est défini dans le cas présent<sup>1</sup> par une moyenne arithmétique pondérée puisque les clichés sont supposés n'être connus qu'aux instants  $t_m$ . On a alors pour toute fonction  $f$  dépendant des clichés

$$E \left[ f(x^{(m)}) \right] = \sum_{m=1}^M \alpha_m f(x^{(m)}) \quad \text{avec} \quad \sum_{m=1}^M \alpha_m = 1. \quad (2.4)$$

Dans l'expression précédente, les coefficients  $\alpha_m$  sont des poids de pondération éventuellement introduits pour accentuer l'effet de certains clichés par rapport à d'autres. Le choix le plus simple consiste à prendre  $\alpha_m = 1/M$  pour tout  $m \in \llbracket 1; M \rrbracket$  si aucun argument ne permet d'attribuer plus d'importance à certains clichés. L'opérateur de moyenne défini ci-dessus est linéaire par rapport aux  $f(x^{(m)})$  puisqu'il s'agit d'une simple moyenne arithmétique. En particulier si la fonction  $f$  désigne le produit scalaire, la bilinéarité de  $f$  entraîne la commutation entre l'opérateur de moyenne et le produit scalaire. On a donc trivialement les deux propriétés suivantes :

$$E \left[ f(x^{(m)}) + \mu g(x^{(m)}) \right] = E \left[ f(x^{(m)}) \right] + \mu E \left[ g(x^{(m)}) \right] \quad \forall \mu \in \mathbb{R}, \quad (2.5a)$$

$$E \left[ \langle x^{(m)}, y \rangle \right] = \langle E \left[ x^{(m)} \right], y \rangle \quad \forall y \in H. \quad (2.5b)$$

On introduit par ailleurs l'opérateur  $\mathcal{P}_S$  de projection orthogonale des éléments de  $H$  sur le sous-espace  $S$ . Celui-ci est défini comme une application de  $H$  à valeurs dans  $S$  telle que la distance entre l'argument du projecteur et son image soit minimale, d'où

$$\mathcal{P}_S y = \sum_{j=1}^q \langle y, \varphi^{(j)} \rangle \varphi^{(j)} \quad \forall y \in H. \quad (2.6)$$

La condition (2.3) revient de manière équivalente à trouver les vecteurs  $\varphi^{(j)} \in H$  qui maximisent la moyenne du carré de la norme de la projection sur le sous-espace. En effet, il suffit de remarquer que pour toute projection orthogonale  $\mathcal{P}$  agissant sur un élément  $x \in H$ , on a la relation  $\|x\|^2 = \|x - \mathcal{P}x\|^2 + \|\mathcal{P}x\|^2$  puisque  $\|x\|^2 = \|(x - \mathcal{P}x) + \mathcal{P}x\|^2 = \|x - \mathcal{P}x\|^2 + \|\mathcal{P}x\|^2 + 2\langle x - \mathcal{P}x, \mathcal{P}x \rangle$ . Or  $\langle x - \mathcal{P}x, \mathcal{P}x \rangle = 0$  puisque les vecteurs  $(x - \mathcal{P}x)$  et  $(\mathcal{P}x)$  sont orthogonaux par définition du projecteur orthogonal. En appliquant l'opération de moyenne  $E[\cdot]$  à la relation précédente pour  $x = x^{(m)}$  et en utilisant la propriété de linéarité donnée par la relation (2.5a), on obtient

$$E \left[ \|x^{(m)}\|^2 \right] = E \left[ \|x^{(m)} - \mathcal{P}_S x^{(m)}\|^2 \right] + E \left[ \|\mathcal{P}_S x^{(m)}\|^2 \right]. \quad (2.7)$$

La relation précédente résume tout l'enjeu de la décomposition orthogonale propre dont l'objectif est de trouver une base qui approche au mieux la moyenne de la norme des clichés  $E[\|x^{(m)}\|^2]$ . Celle-ci s'obtient donc soit en minimisant l'erreur de projection  $E[\|x^{(m)} - \mathcal{P}_S x^{(m)}\|^2]$  ce qui conduit à la définition du problème donnée à l'équation (2.3), soit en maximisant la projection  $E[\|\mathcal{P}_S x^{(m)}\|^2]$  ce qui donne lieu à la définition équivalente ci-dessous.

<sup>1</sup>L'application de la décomposition orthogonale propre à un système déterministe repose sur une hypothèse d'ergodicité, qui stipule que la moyenne statistique d'un ensemble de réalisations d'un système aléatoire peut être remplacée de manière équivalente par une moyenne temporelle sur une réalisation. L'opérateur de moyenne est noté  $E[\cdot]$  puisque dans le cas d'un système stochastique il représente l'espérance. La pondération introduite dans la moyenne correspond alors à la probabilité.

**Définition 2.2 :** Étant donné un ensemble de clichés  $X = \{x^{(m)} \in H\}_{m=1}^M$  et un opérateur de moyenne  $E[\cdot]$ , la base des modes POD  $\Phi = \{\varphi^{(j)} \in H\}_{j=1}^q$  de dimension  $q$  est définie comme la solution du problème de maximisation

$$\Phi = \arg \max_{\{\varphi^{(j)}\}_{j=1}^q \subset H} E \left[ \left\| \sum_{j=1}^q \langle x^{(m)}, \varphi^{(j)} \rangle \varphi^{(j)} \right\|^2 \right] \quad \text{tel que} \quad \langle \varphi^{(i)}, \varphi^{(j)} \rangle = \delta_{ij}. \quad (2.8)$$

Le carré de la norme de la projection orthogonale se développe au moyen du théorème de Pythagore généralisé puisque les modes POD recherchés sont soumis à la contrainte d'être orthonormés. Ainsi,

$$E \left[ \left\| \sum_{j=1}^q \langle x^{(m)}, \varphi^{(j)} \rangle \varphi^{(j)} \right\|^2 \right] = E \left[ \sum_{j=1}^q \langle x^{(m)}, \varphi^{(j)} \rangle^2 \right] \quad (2.9)$$

et puisque l'opérateur de moyenne est linéaire, les propriétés (2.5) permettent de transformer la relation précédente en

$$E \left[ \sum_{j=1}^q \langle x^{(m)}, \varphi^{(j)} \rangle^2 \right] = \sum_{j=1}^q E \left[ \langle \langle x^{(m)}, \varphi^{(j)} \rangle x^{(m)}, \varphi^{(j)} \rangle \right] = \sum_{j=1}^q \langle E \left[ \langle x^{(m)}, \varphi^{(j)} \rangle x^{(m)} \right], \varphi^{(j)} \rangle. \quad (2.10)$$

Le premier terme du produit scalaire s'interprète finalement comme l'action d'un opérateur linéaire  $R$  sur la variable  $\varphi^{(j)}$ , défini par

$$\begin{aligned} R : H &\longrightarrow H \\ y &\longmapsto Ry = E \left[ \langle x^{(m)}, y \rangle x^{(m)} \right]. \end{aligned} \quad (2.11)$$

La linéarité de l'opérateur  $R$  est triviale puisqu'il est défini à partir (i) de l'opérateur de moyenne  $E[\cdot]$  pour lequel le caractère linéaire a été signalé plus tôt et (ii) du produit scalaire qui est une application bilinéaire. Par conséquent on a bien  $R(\varphi + \mu\psi) = R\varphi + \mu R\psi$  pour tout  $\varphi, \psi$  dans  $H$  et tout  $\mu \in \mathbb{R}$ . Par ailleurs, il est aussi important de remarquer que l'opérateur est auto-adjoint, *i. e.*  $\langle R\varphi, \psi \rangle = \langle \varphi, R\psi \rangle$  pour tout  $\varphi, \psi$  dans  $H$  (voir l'annexe B.2 pour la démonstration).

Le problème d'optimisation (2.8) revient donc à maximiser la somme  $\sum_{j=1}^q \langle R\varphi^{(j)}, \varphi^{(j)} \rangle$ , ou encore à maximiser séparément chacun des termes de la somme avec toujours la contrainte de chercher des modes orthonormaux. On formule donc un problème variationnel pour chaque mode POD  $\varphi^{(j)}$  au moyen de la fonctionnelle  $\mathcal{J}(\varphi^{(j)})$  définie par la relation

$$\mathcal{J}(\varphi^{(j)}) = \langle R\varphi^{(j)}, \varphi^{(j)} \rangle - \lambda_j \left( \|\varphi^{(j)}\|^2 - 1 \right), \quad (2.12)$$

où  $\lambda_j$  est un multiplicateur de Lagrange introduit pour tenir compte de la contrainte d'orthonormalité imposée<sup>2</sup>. Une condition nécessaire pour évaluer les maxima porte sur l'annulation de la dérivée de la fonctionnelle  $\mathcal{J}$  pour toute variation  $\varphi^{(j)} + \delta\psi$  avec  $\delta \in \mathbb{R}$  et  $\psi \in H : d\mathcal{J}(\varphi^{(j)} + \delta\psi)/d\delta|_{\delta=0} = 0$ . Pour satisfaire la condition précédente, il suffit d'abord de détailler  $\mathcal{J}(\varphi^{(j)} + \delta\psi)$  avec l'expression (2.12) en

$$\begin{aligned} \mathcal{J}(\varphi^{(j)} + \delta\psi) &= \langle R(\varphi^{(j)} + \delta\psi), \varphi^{(j)} + \delta\psi \rangle - \lambda_j \left( \|\varphi^{(j)} + \delta\psi\|^2 - 1 \right) \\ &= \langle R\varphi^{(j)}, \varphi^{(j)} \rangle + \delta \langle R\varphi^{(j)}, \psi \rangle + \delta \langle R\psi, \varphi^{(j)} \rangle + \delta^2 \langle R\psi, \psi \rangle \\ &\quad - \lambda_j \left( \langle \varphi^{(j)}, \varphi^{(j)} \rangle + \delta \langle \varphi^{(j)}, \psi \rangle + \delta \langle \psi, \varphi^{(j)} \rangle + \delta^2 \langle \psi, \psi \rangle - 1 \right). \end{aligned} \quad (2.13)$$

La dérivée de la fonctionnelle par rapport à la variation  $\delta$  conduit alors à

$$\frac{d\mathcal{J}(\varphi^{(j)} + \delta\psi)}{d\delta} = \langle R\varphi^{(j)}, \psi \rangle + \langle R\psi, \varphi^{(j)} \rangle + 2\delta \langle R\psi, \psi \rangle - \lambda_j \left( \langle \varphi^{(j)}, \psi \rangle + \langle \psi, \varphi^{(j)} \rangle + 2\delta \langle \psi, \psi \rangle \right). \quad (2.14)$$

<sup>2</sup>Voir l'ouvrage de Holmes *et al.* [112] pour la formulation du problème d'optimisation dans le contexte de la POD et l'ouvrage de Fletcher [77] par exemple pour les détails généraux relatifs à l'optimisation.



Finalement la dérivée de la fonctionnelle évaluée en  $\delta = 0$  se réduit à

$$\begin{aligned} \left. \frac{d \mathcal{J} [\varphi^{(j)} + \delta \psi]}{d\delta} \right|_{\delta=0} &= \langle R \varphi^{(j)}, \psi \rangle + \langle R \psi, \varphi^{(j)} \rangle - \lambda_j (\langle \varphi^{(j)}, \psi \rangle + \langle \psi, \varphi^{(j)} \rangle) \\ &= 2 (\langle R \varphi^{(j)}, \psi \rangle - \lambda_j \langle \varphi^{(j)}, \psi \rangle) = 2 \langle R \varphi^{(j)} - \lambda_j \varphi^{(j)}, \psi \rangle. \end{aligned} \quad (2.15)$$

en utilisant le caractère auto-adjoint de l'opérateur  $R$  et le fait que le produit scalaire est une application symétrique. Si l'on en revient à présent au problème d'origine, les extrema de la fonctionnelle s'obtiennent lorsque l'expression précédente s'annule pour tout  $\psi \in H$ , ce qui finalement est équivalent à résoudre le problème aux valeurs propres

$$\boxed{R \varphi^{(j)} = \lambda_j \varphi^{(j)}}. \quad (2.16)$$

pour tout  $j \in \llbracket 1; q \rrbracket$ . En conclusion, maximiser (resp. minimiser) la projection (resp. l'erreur de projection) de l'ensemble  $X$  des clichés sur le sous-espace  $S = \text{vect}\{\varphi^{(1)}, \dots, \varphi^{(q)}\}$  revient à résoudre un problème aux valeurs propres pour l'opérateur linéaire  $R$  défini à l'équation (2.11).

Pour terminer, il reste à établir que les solutions du problème (2.16) existent. Pour cela, il suffit de démontrer que l'opérateur  $R$  est compact. En effet, cette dernière propriété, associée au caractère auto-adjoint de  $R$ , permet d'invoquer la théorie spectrale pour les opérateurs de Hilbert-Schmidt, qui assure qu'il existe au plus une infinité dénombrable de valeurs propres positives ou nulles [205]. Les vecteurs propres correspondant à ces valeurs propres sont en outre orthogonaux entre eux. Quelques détails concernant les propriétés de ce type d'opérateurs sont donnés dans l'annexe B.2. Enfin, le rang  $r$  de l'opérateur  $R$  est fini si bien que le problème aux valeurs propres (2.16) produit un nombre fini  $r$  de valeurs propres  $\lambda_j$  et de vecteurs propres  $\varphi^{(j)}$ . Lorsque tous les modes sont retenus, les clichés  $x^{(m)}$  se décomposent exactement sur la base des modes POD en

$$x^{(m)} = \sum_{j=1}^r a_j^{(m)} \varphi^{(j)} \quad \forall m \in \llbracket 1; M \rrbracket. \quad (2.17)$$

La décomposition tronquée (2.2) s'obtient donc en ne conservant que les  $q$  premiers modes POD associés aux valeurs propres les plus élevées, ce choix étant justifié ultérieurement au paragraphe 2.2.3. En pratique  $q$  sera choisi le plus petit possible et nécessairement inférieur au rang  $r$  de l'opérateur  $R$ . Pour éviter que le nombre de valeurs propres non nulles soit identiquement nul, il suffit que le rang de l'opérateur  $R$  soit strictement positif ; cette condition est vérifiée dès qu'un des clichés est non nul, ou lorsque les clichés sont linéairement indépendants, auquel cas on a  $r = M$  valeurs propres non nulles.

Les résultats établis dans ce paragraphe peuvent être résumés sous forme du théorème suivant :

### **Théorème 2.1 : Décomposition Orthogonale Propre**

Étant donné un ensemble de données  $X = \{x^{(m)} \in H\}_{m=1}^M$  et un opérateur de moyenne  $E[\cdot]$ , le sous-espace  $S \subset H$  de dimension finie  $q < r$  qui minimise l'erreur de projection  $E[\|x^{(m)} - \mathcal{P}_S x^{(m)}\|^2]$  est engendré par la base orthonormale  $\Phi = \{\varphi^{(j)}\}_{j=1}^q$ , où les  $\varphi^{(j)}$  sont les modes POD solutions du problème aux valeurs propres :

$$R \varphi^{(j)} = \lambda_j \varphi^{(j)}.$$

L'opérateur  $R \varphi^{(j)} = E[\langle x^{(m)}, \varphi^{(j)} \rangle x^{(m)}]$  est compact, auto-adjoint et de rang fini  $r$ . Il possède donc un nombre fini de valeurs propres non nulles  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$ .

La description de la décomposition orthogonale propre effectuée ici fait clairement ressortir le fait qu'il s'agit d'une méthode basée sur l'analyse de données a posteriori. En effet, la recherche des modes POD est entièrement fondée sur un problème d'optimisation du sous-espace approchant l'ensemble des clichés  $X$  fourni au départ. Il est par ailleurs important de noter que la décomposition orthogonale propre est essentiellement une méthode linéaire, dans le sens où la représentation des données obtenue s'écrit sous forme d'une série de fonctions modales multipliées par les coefficients adéquats, et ceci, malgré le fait que dans la plupart des cas les données traitées proviennent d'un système dynamique non-linéaire.

**Remarque 2.1 : Expressions de l'opérateur  $R$  pour des espaces de Hilbert spécifiques**

Avant de poursuivre, on précise ici l'expression de l'opérateur  $R$  lorsque l'espace de Hilbert considéré est l'espace  $H = \mathbb{R}^{N_v}$  des vecteurs réels de dimension finie  $N_v$ , ou l'espace  $H = L^2(\Omega)$  des fonctions de carré intégrable.

Dans le premier cas, la définition de l'opérateur donnée à l'équation (2.11) se développe avec le produit scalaire canonique de  $\mathbb{R}^{N_v}$  en  $R\varphi^{(j)} = E[\sum_{i=1}^{N_v} x_i^{(m)} \varphi_i^{(j)} x^{(m)}]$ , si bien que le problème aux valeurs propres (2.16) se réduit à un simple problème matriciel puisque l'opérateur  $R$  est une matrice de  $\mathcal{M}(N_v, N_v)$  notée  $\mathbf{R}$  et telle que

$$\mathbf{R} \varphi^{(j)} = \lambda_j \varphi^{(j)} \quad \text{où} \quad \mathbf{R} = E \left[ \mathbf{x}^{(m)} \otimes \mathbf{x}^{(m)} \right]. \quad (2.18)$$

Dans l'expression précédente, le symbole  $\otimes$  désigne le produit tensoriel et les notations en gras correspondent à la représentation des quantités vectorielles dans la base canonique de  $\mathbb{R}^{N_v}$  de sorte que  $\mathbf{x}^{(m)} = [x_1^{(m)}, \dots, x_{N_v}^{(m)}]$  et  $\varphi^{(j)} = [\varphi_1^{(j)}, \dots, \varphi_{N_v}^{(j)}]$ . Le problème se ramène donc simplement à la résolution d'un problème d'algèbre linéaire dans lequel il s'agit d'identifier les vecteurs propres  $\varphi^{(j)}$ .

Lorsque l'espace  $H = L^2(\Omega)$  est considéré, le produit scalaire associé conduit à la définition suivante du problème aux valeurs propres en regroupant les termes sur lesquels l'opérateur de moyenne  $E[\cdot]$  opère :

$$\int_{\Omega} \tilde{R}(\xi, \xi') \varphi^{(j)}(\xi) d\xi = \lambda_j \varphi^{(j)} \quad \text{où} \quad \tilde{R}(\xi, \xi') = E \left[ x^{(m)}(\xi) x^{(m)}(\xi') \right]. \quad (2.19)$$

Dans ce cas, l'opérateur  $\tilde{R}$  correspond au tenseur de corrélation entre deux points  $\xi$  et  $\xi'$  du domaine spatial  $\Omega$ . Le problème à résoudre fait alors intervenir une équation intégrale dite de Fredholm.

## 2.2.2 Détermination des modes POD par la méthode des clichés

Le processus décrit dans ce qui précède porte en général le nom de *méthode directe* puisque les modes POD sont obtenus directement par la résolution du problème aux valeurs propres (2.16). Toutefois, la recherche des valeurs et vecteurs propres par la méthode directe devient rapidement très coûteuse lorsque la dimension  $N$  de l'espace dans lequel les clichés sont contenus devient importante. Le problème se pose par exemple pour des clichés issus de simulations numériques où il n'est pas rare que  $N$  soit de l'ordre de  $\mathcal{O}(10^5)$  et bien souvent plus pour les applications en mécanique des fluides. L'opérateur  $R$  a une taille similaire (voir l'équation (2.18) dans le cas discret par exemple) et par conséquent l'extraction des valeurs et vecteurs propres devient fastidieuse, voire impossible en pratique.

Une variante de la méthode directe proposée par Sirovich [218] sous le nom de *méthode des clichés* permet de réduire considérablement le coût de calcul en se basant sur la résolution d'un problème aux valeurs propres annexe dont la taille est nettement inférieure. La nouvelle formulation du problème découle du fait que les modes POD  $\varphi^{(j)}$  s'expriment en fonction des clichés  $x^{(m)}$  selon la combinaison linéaire

$$\varphi^{(j)} = \sum_{m=1}^M c_m^{(j)} x^{(m)} \quad (2.20)$$

qui est justifiée puisque l'image de l'opérateur  $R$  est incluse dans l'espace engendré par les clichés. La démonstration est donnée dans l'annexe B.2 lorsqu'on établit le rang de  $R$ , et la décomposition des modes POD sur les clichés est détaillée dans la remarque B.1 page 220.

Les  $c_m^{(j)}$  sont des coefficients réels décrivant la décomposition du mode  $\varphi^{(j)}$  sur les clichés  $x^{(m)}$  qui correspondent dans la méthode telle qu'elle a été présentée par Sirovich [218] à des solutions du système évaluées à des instants  $t_m$  régulièrement espacés sur l'intervalle de temps  $I_e$ . On généralise dans ce paragraphe le principe d'obtention des modes POD par la méthode des clichés à partir du problème aux valeurs propres de la méthode directe dans le cas de clichés pondérés par l'opérateur de moyenne  $E[\cdot]$ .

Le problème aux valeurs propres de la méthode directe (2.16) écrit pour un mode  $\varphi^{(j)}$  donné est transformé en introduisant la décomposition des modes POD (2.20) dans la définition (2.11) de l'opérateur  $R$  :

$$\begin{aligned} \sum_{m=1}^M \alpha_m \left\langle x^{(m)}, \sum_{l=1}^M c_l^{(j)} x^{(l)} \right\rangle x^{(m)} &= \lambda_j \sum_{k=1}^M c_k^{(j)} x^{(k)} \\ \sum_{m=1}^M \left( \sum_{l=1}^M \alpha_m \langle x^{(m)}, x^{(l)} \rangle c_l^{(j)} \right) x^{(m)} &= \lambda_j \sum_{k=1}^M c_k^{(j)} x^{(k)}. \end{aligned} \quad (2.21)$$

Dans la somme du second membre, l'indice  $k$  est muet et on peut donc à ce titre choisir d'indexer plutôt la somme avec l'indice  $m$ . En regroupant les membres de droite et de gauche d'un même côté et dans une somme commune, on parvient à l'écriture :  $\sum_{m=1}^M \theta_m x^{(m)} = 0$  avec  $\theta_m = \sum_{l=1}^M \alpha_m \langle x^{(m)}, x^{(l)} \rangle c_l^{(j)} - \lambda_j c_m^{(j)}$ . Sous l'hypothèse que l'ensemble des clichés  $X = \{x^{(m)}\}_{m=1}^M$  forme une famille libre, la somme des  $\theta_m$  est nulle si et seulement si tous les  $\theta_m$  sont nuls pour  $m \in \llbracket 1; M \rrbracket$ . En reprenant l'expression des  $\theta_m$ , le problème se résume à résoudre le système linéaire

$$\sum_{l=1}^M \alpha_m \langle x^{(m)}, x^{(l)} \rangle c_l^{(j)} = \lambda_j c_m^{(j)} \quad \forall m \in \llbracket 1; M \rrbracket. \quad (2.22)$$

L'expression du nouveau problème aux valeurs propres est alors

$$\tilde{\mathbf{R}}^* \mathbf{c}^{(j)} = \lambda_j \mathbf{c}^{(j)}. \quad (2.23)$$

Le terme général du nouvel opérateur est  $\tilde{R}_{ij}^* = \alpha_i \langle x^{(i)}, x^{(j)} \rangle$  et  $\mathbf{c}^{(j)} = [c_1^{(j)}, \dots, c_M^{(j)}]^\top$  est le vecteur propre associé à la valeur propre  $\lambda_j$ . Sous cette forme, la matrice  $\tilde{\mathbf{R}}^*$  n'est pas symétrique, à moins que les poids  $\alpha_i$  associés à l'opérateur de moyenne  $E[\cdot]$  ne soient tous égaux. Le caractère symétrique est pourtant crucial puisqu'il assure l'existence des valeurs propres et, par voie de conséquence, celle de la décomposition orthogonale propre. L'opérateur aux valeurs propres  $\tilde{\mathbf{R}}^*$  est donc redéfini en écrivant  $\tilde{\mathbf{R}}^* = \alpha \mathbf{P}$  où  $\alpha = \text{diag}(\alpha_1, \dots, \alpha_M)$  est la matrice diagonale des poids de pondération de l'opérateur de moyenne et  $\mathbf{P}$  est la matrice des produits scalaires telle que  $P_{ij} = \langle x^{(i)}, x^{(j)} \rangle$ . En supposant que les poids de l'opérateur de moyenne soient tous strictement positifs, la matrice  $\alpha$  est définie positive, ce qui permet de définir  $\tilde{\alpha} = \text{diag}(\sqrt{\alpha_1}, \dots, \sqrt{\alpha_M})$  telle que  $\alpha = \tilde{\alpha} \tilde{\alpha}^\top$ . Le changement de variable  $\mathbf{c}^{(j)} = \tilde{\alpha} \mathbf{d}^{(j)}$  transforme alors le problème aux valeurs propres en

$$\underbrace{\tilde{\alpha} \left( \tilde{\alpha}^\top \mathbf{P} \tilde{\alpha} \right)}_{\mathbf{R}^*} \mathbf{d}^{(j)} = \lambda_j \tilde{\alpha} \mathbf{d}^{(j)}. \quad (2.24)$$

Finalement, puisque la matrice  $\tilde{\alpha}$  est définie positive, son inverse existe et la pré-multiplication de la relation précédente par cet inverse donne

$$\boxed{\mathbf{R}^* \mathbf{d}^{(j)} = \lambda_j \mathbf{d}^{(j)}}. \quad (2.25)$$

L'opérateur  $\mathbf{R}^*$  est bien symétrique puisque son terme général est défini par  $R_{ij}^* = \sqrt{\alpha_i \alpha_j} \langle x^{(i)}, x^{(j)} \rangle$ . La taille du problème aux valeurs propres (2.25) est nettement réduite puisque  $\mathbf{R}^* \in \mathcal{M}(M, M)$  alors que l'opérateur impliqué dans le problème (2.16) est une matrice de  $\mathcal{M}(N, N)$ . La méthode des clichés devient donc beaucoup plus avantageuse en terme de temps de calcul dès lors que le nombre de clichés  $M$  est très inférieur à la dimension  $N$  du système. Le choix entre la méthode directe ou la méthode des clichés s'effectue en fonction des valeurs respectives de  $M$  et  $N$ . Lorsque les clichés proviennent d'un dispositif expérimental, la résolution spatiale des données est généralement faible puisqu'elle est limitée par le nombre de capteurs mais la résolution temporelle est souvent fine ; dans ce cas, la méthode directe paraît plus adaptée puisque la dimension du système représenté par les données des capteurs est très inférieure au nombre de clichés ( $N \ll M$ ). Au contraire, les simulations numériques produisent une résolution spatiale importante en raison de la discrétisation fine du maillage alors que la résolution temporelle est plutôt limitée ( $N \gg M$ ) ; dans ce cas, il est préférable d'opter pour la méthode des clichés afin de réduire la taille du problème aux valeurs propre à résoudre.

Il faut par ailleurs mentionner que la méthode des clichés ne fournit pas directement les modes POD. En effet, la résolution du problème (2.25) produit l'ensemble des vecteurs propres  $\mathbf{d}^{(j)}$  pour  $j = 1, \dots, M$ . Dans un premier temps, les coefficients  $c_m^{(j)}$  de la décomposition (2.17) sont calculés à l'aide des vecteurs propres  $\mathbf{d}^{(j)}$  et du changement de variable  $\mathbf{c}^{(j)} = \tilde{\alpha} \mathbf{d}^{(j)}$  effectué pour symétriser le problème. Ensuite, les modes POD  $\varphi^{(j)}$  de la décomposition (2.17) sont reconstruits par combinaison linéaire avec les clichés au

moyen de la relation (2.20). En procédant de cette manière, la condition d'orthonormalité des modes POD  $\langle \varphi^{(i)}, \varphi^{(j)} \rangle = \delta_{ij}$  imposée avec la méthode directe n'est plus respectée. En effet, la condition d'orthonormalité porte à présent sur les vecteurs propres  $\mathbf{d}^{(j)}$  puisqu'ils sont solutions du problème aux valeurs propres (2.25) et de ce fait le produit scalaire de deux modes POD donne

$$\begin{aligned} \langle \varphi^{(i)}, \varphi^{(j)} \rangle &= \left\langle \sum_{k=1}^M c_k^{(i)} x^{(k)}, \sum_{l=1}^M c_l^{(j)} x^{(l)} \right\rangle \\ &= \sum_{k,l=1}^M c_k^{(i)} \langle x^{(k)}, x^{(l)} \rangle c_l^{(j)} \\ &= \left( \mathbf{c}^{(i)} \right)^\top \mathbf{P} \mathbf{c}^{(j)} = \left( \mathbf{d}^{(i)} \right)^\top \left( \tilde{\alpha}^\top \mathbf{P} \tilde{\alpha} \right) \mathbf{d}^{(j)}. \end{aligned} \quad (2.26)$$

Dans la dernière égalité, le terme  $\tilde{\alpha}^\top \mathbf{P} \tilde{\alpha}$  correspond à l'opérateur  $\mathbf{R}^*$ . Le produit  $(\tilde{\alpha}^\top \mathbf{P} \tilde{\alpha}) \mathbf{d}^{(j)}$  est donc remplacé par  $\mathbf{R}^* \mathbf{d}^{(j)} = \lambda_j \mathbf{d}^{(j)}$  puisque  $\mathbf{d}^{(j)}$  est un vecteur propre de l'opérateur  $\mathbf{R}^*$ . Ainsi le produit scalaire de deux modes POD se réduit à la relation  $\langle \varphi^{(i)}, \varphi^{(j)} \rangle = \lambda_j (\mathbf{d}^{(i)})^\top \mathbf{d}^{(j)}$ . Cette dernière se simplifie au moyen de la condition d'orthonormalité des vecteurs  $\mathbf{d}^{(i)}$  qui sont solutions du problème aux valeurs propres (2.25). Le produit scalaire entre deux modes POD devient alors

$$\langle \varphi^{(i)}, \varphi^{(j)} \rangle = \lambda_j \delta_{ij}. \quad (2.27)$$

Les modes POD obtenus par la méthode des clichés après reconstruction doivent donc subir une étape supplémentaire de normalisation par la racine carrée des valeurs propres  $\sqrt{\lambda_j}$  pour disposer d'une condition d'orthonormalité  $\langle \varphi^{(i)}, \varphi^{(j)} \rangle = \delta_{ij}$  analogue à celle obtenue par la méthode directe.

Pour terminer, il est intéressant de comparer la signification des opérateurs pour les méthodes directe et des clichés. Dans la méthode des clichés, le terme général de l'opérateur est proportionnel à  $\langle x^{(i)}, x^{(j)} \rangle$ . Il correspond donc à la somme du produit des contributions spatiales de deux clichés pris à deux instants différents. Autrement dit, le terme général correspond à la *corrélation temporelle* des clichés  $x^{(i)}$  et  $x^{(j)}$ . Dans le cas de la méthode directe, l'expression de l'opérateur  $R$  a été détaillée à la remarque 2.1 dans le cas de deux espaces de Hilbert spécifiques. Le terme général de l'opérateur correspond dans ce cas à des *corrélations spatiales* puisque chaque terme de l'opérateur est la moyenne du produit des contributions temporelles de deux points de l'espace différents. Il existe donc en quelque sorte une dualité espace/temps entre les deux méthodes de calcul des modes POD. Dans les deux cas cependant, les modes POD représentent les distributions spatiales qui maximisent la projection des clichés sur la base des modes POD. On a montré dans ce paragraphe que les deux méthodes produisent des résultats équivalents, sous l'hypothèse que l'ensemble de clichés  $X$  est linéairement indépendant. Cette équivalence est généralisée dans l'annexe B.3 pour des clichés contenus dans un espace de dimension finie sans avoir recours à cette hypothèse, et le lien avec la décomposition en valeurs singulières  $y$  est en outre explicité.

## 2.2.3 Propriétés de la décomposition orthogonale propre

### 2.2.3.1 Orthogonalité des amplitudes modales en moyenne

Les amplitudes modales  $a_j^{(m)}$  de la décomposition (2.17) sont telles que la moyenne temporelle du produit de deux amplitudes est liée aux valeurs propres  $\lambda_j$  de l'opérateur de corrélation par la relation

$$E \left[ a_j^{(m)} a_i^{(m)} \right] = \lambda_j \delta_{ij}. \quad (2.28)$$

Le résultat précédent s'obtient facilement en écrivant que les amplitudes modales se déduisent du produit scalaire entre les clichés et les modes POD. On a en effet le résultat suivant en prenant le produit scalaire entre la décomposition (2.17) et un mode POD  $\varphi^{(i)}$  puisque les modes POD sont orthogonaux :

$$a_j^{(m)} = \langle x^{(m)}, \varphi^{(j)} \rangle. \quad (2.29)$$

La définition est triviale si on se réfère à la définition (2.6) de la projection orthogonale sur le sous-espace  $S$ . Par conséquent, la moyenne du produit de deux amplitudes modales donnée à l'équation (2.28) se récrit

$$\begin{aligned} E \left[ a_j^{(m)} a_i^{(m)} \right] &= E \left[ \langle x^{(m)}, \varphi^{(j)} \rangle \langle x^{(m)}, \varphi^{(i)} \rangle \right] = E \left[ \langle \langle x^{(m)}, \varphi^{(i)} \rangle x^{(m)}, \varphi^{(j)} \rangle \right] \\ &= \langle E \left[ \langle x^{(m)} x^{(m)}, \varphi^{(i)} \rangle \right], \varphi^{(j)} \rangle = \langle R \varphi^{(i)}, \varphi^{(j)} \rangle. \end{aligned} \quad (2.30)$$

La dernière égalité provient de la définition (2.11) de l'opérateur de corrélation  $R$ . Finalement, le problème aux valeurs propres (2.16) permet de conclure puisque  $\langle R \varphi^{(i)}, \varphi^{(j)} \rangle = \langle \lambda_i \varphi^{(i)}, \varphi^{(j)} \rangle = \lambda_i \langle \varphi^{(i)}, \varphi^{(j)} \rangle = \lambda_i \delta_{ij}$ , ce qui démontre la relation (2.28).

### 2.2.3.2 Énergie associée aux modes POD

Les valeurs propres  $\lambda_j$  extraites du problème (2.16) sont souvent interprétées en terme d'« énergie ». La projection de l'équation (2.16) sur un mode  $\varphi^{(j)}$  donne d'après ce qui précède :

$$\langle R \varphi^{(j)}, \varphi^{(j)} \rangle = \lambda_j \langle \varphi^{(j)}, \varphi^{(j)} \rangle = \lambda_j. \quad (2.31)$$

Or l'opérateur  $R$  a été défini à l'équation (2.11) de sorte que  $\langle R \varphi^{(j)}, \varphi^{(j)} \rangle = E[\langle x^{(m)}, \varphi^{(j)} \rangle^2]$ . Par conséquent la valeur propre  $\lambda_j$  est un indicateur du degré de parallélisme entre l'ensemble des clichés et les modes POD puisque la relation

$$\lambda_j = E \left[ \langle x^{(m)}, \varphi^{(j)} \rangle^2 \right] = E \left[ \left( a_j^{(m)} \right)^2 \right]. \quad (2.32)$$

indique que les valeurs propres  $\lambda_j$  les plus grandes sont celles qui maximisent en moyenne la projection des clichés  $x^{(m)}$  sur les modes  $\varphi^{(j)}$ , ou de manière équivalente, qui minimisent en moyenne l'erreur de projection. Une seconde interprétation proposée par Sirovich [218] consiste à considérer que les valeurs propres sont une expression du temps moyen relatif passé par le système le long de l'axe  $\varphi^{(j)}$  dans l'espace fonctionnel.

Par ailleurs, la somme des valeurs propres représente l'énergie moyenne totale contenue dans le système, cette énergie étant définie mathématiquement par  $E[\|x^{(m)}\|^2]$ . On montre ainsi que

$$E \left[ \|x^{(m)}\|^2 \right] = \sum_j^r \lambda_j. \quad (2.33)$$

En effet, en détaillant l'écriture du premier membre avec la décomposition (2.17) on a

$$\begin{aligned} E \left[ \langle x^{(m)}, x^{(m)} \rangle \right] &= \sum_{m=1}^M \alpha_m \left\langle \sum_{i=1}^r a_i^{(m)} \varphi^{(i)}, \sum_{j=1}^r a_j^{(m)} \varphi^{(j)} \right\rangle \\ &= \sum_{m=1}^M \alpha_m \sum_{i,j=1}^r a_i^{(m)} a_j^{(m)} \langle \varphi^{(i)}, \varphi^{(j)} \rangle. \end{aligned} \quad (2.34)$$

Par orthonormalité des modes POD, le produit scalaire subsistant se réduit à  $\delta_{ij}$  si bien que le terme d'énergie moyenne devient

$$E \left[ \|x^{(m)}\|^2 \right] = \sum_{i=1}^r \sum_{m=1}^M \alpha_m a_i^{(m)} a_i^{(m)} = \sum_{i=1}^r E \left[ a_i^{(m)} a_i^{(m)} \right]. \quad (2.35)$$

La deuxième égalité de la relation (2.28) permet alors de conclure la démonstration de la propriété (2.33).

On dira donc par abus de langage que la valeur propre  $\lambda_j$  correspond à l'« énergie », au sens de la norme induite par l'espace  $H$ , captée par le mode POD  $\varphi^{(j)}$  puisque d'après la relation (2.32), elle donne une mesure de la projection moyenne des clichés sur les modes POD. Les modes POD associés aux valeurs propres les plus grandes sont donc ceux qui contiennent le plus d'énergie et par conséquent ceux qui seront le plus à même de reproduire la réponse du système. Ces observations restent valables pour la méthode des clichés puisque les valeurs et vecteurs propres calculés par la méthode directe ou des clichés sont identiques.

### 2.2.3.3 Troncature de la base POD et estimateur d'erreur

L'objectif étant à terme de construire un modèle d'ordre réduit avec un faible nombre de degrés de liberté, la base des modes POD est en pratique tronquée pour ne conserver que les  $q$  premiers modes POD associés aux valeurs propres les plus grandes. La décomposition n'est plus exacte et il est donc souhaitable de disposer d'un indicateur d'erreur afin de pouvoir estimer l'erreur de troncature. Un moyen simple de caractériser cette quantité est de calculer l'erreur pour les clichés qui ont servi à la construction des modes POD. La décomposition orthogonale propre exacte de l'ensemble de clichés  $X$  est donnée par la relation (2.17) dans laquelle l'ensemble des modes POD est conservé. En revanche, la décomposition orthogonale propre tronquée donnée à l'équation (2.2) fournit une approximation des clichés notée  $\hat{x}^{(m)}$ . L'erreur de troncature notée  $\epsilon_q$  est alors définie par la relation

$$\epsilon_q = E \left[ \left\| x^{(m)} - \hat{x}^{(m)} \right\|^2 \right] = E \left[ \left\| \sum_{j=1}^r a_j^{(m)} \varphi^{(j)} - \sum_{j=1}^q a_j^{(m)} \varphi^{(j)} \right\|^2 \right] = E \left[ \left\| \sum_{j=q+1}^r a_j^{(m)} \varphi^{(j)} \right\|^2 \right]. \quad (2.36)$$

L'expression de l'erreur se simplifie en raison de l'orthonormalité des modes POD  $\varphi^{(j)}$  en

$$\epsilon_q = E \left[ \sum_{j=q+1}^r \left( a_j^{(m)} \right)^2 \right] = \sum_{j=q+1}^r E \left[ a_j^{(m)} a_j^{(m)} \right]. \quad (2.37)$$

Finalement, l'interprétation énergétique des valeurs propres (2.32) établie dans le paragraphe précédent fournit l'expression suivante de l'erreur commise en fonction des valeurs propres négligées :

$$\epsilon_q = \sum_{j=q+1}^r \lambda_j \xrightarrow{j \rightarrow r} 0. \quad (2.38)$$

L'indicateur traduit donc le fait que l'erreur commise correspond aux valeurs propres négligées, *i. e.* à l'énergie associée aux modes POD omis dans la base POD tronquée. Ces modes correspondant aux plus basses valeurs propres contiennent peu d'énergie et peuvent donc dans la majorité des cas être négligés sans altérer notablement la reconstruction des clichés par la relation (2.2). En pratique, une fois que les valeurs et vecteurs propres du système sont calculés, l'indice de troncature de la base POD est déterminé comme le nombre de modes POD  $q$  pour lequel l'énergie captée  $\Lambda_q$  représente presque la totalité de l'énergie  $\Lambda_r$  (plus de 99% par exemple). Le ratio  $\eta_q$  caractérisant le pourcentage d'énergie captée par la décomposition tronquée est alors donné par

$$\eta_q = \frac{\Lambda_q}{\Lambda_r} = \sum_{j=1}^q \lambda_j / \sum_{j=1}^r \lambda_j, \quad (2.39)$$

et le pourcentage d'erreur commis correspondant à l'erreur de troncature  $\epsilon_q$  est donc  $1 - \eta_q$ .

L'utilisation de l'erreur de troncature  $\epsilon_q$  comme seul critère pour déterminer la dimension de la base de projection n'est suffisant que dans le cadre de la compression de données où les coefficients  $a_j^{(m)}$  sont ceux issus de la projection des clichés. Lorsque la base POD est employée pour construire un modèle d'ordre réduit, il devient alors nécessaire de déterminer des estimateurs d'erreur basés sur les amplitudes modales  $a_j(t_m)$  issues de l'intégration en temps, ce qui est autrement plus difficile. Le développement d'estimateurs d'erreur pour les systèmes dynamiques réduits a fait l'objet de plusieurs études. Le lecteur intéressé pourra consulter en particulier les références suivantes [56, 110, 135, 136].

Le choix d'une base composée des  $q$  premiers modes POD associés aux valeurs propres les plus grandes telle que  $1 - \eta_q$  soit inférieur à un très faible pourcentage n'est donc pas toujours adéquat. En effet, Henri et Yvon [110] ont signalé que l'ordre des modes POD n'est pas toujours pertinent et l'erreur de reconstruction du modèle réduit peut être diminuée en choisissant des modes POD associés à des valeurs propres plus petites. Ils ont alors proposé de modifier le critère pour la troncature de la base POD en introduisant des conditions supplémentaires portant sur les bornes des estimateurs d'erreur associés au système dynamique.

### 2.2.3.4 Héritage de propriétés par les modes POD

Enfin un aspect fondamental de la décomposition orthogonale propre concerne la conservation de certaines propriétés vérifiées par les clichés. L'expression des modes POD en fonction des clichés établie à l'équation (2.20) indique que si les clichés satisfont une contrainte *linéaire* homogène notée symboliquement  $L(x^{(m)}) = 0$  où  $L : H \rightarrow H$  est un endomorphisme de  $H$ , alors les modes POD vérifient aussi cette contrainte d'après la relation

$$L(\varphi^{(j)}) = L\left(\sum_{m=1}^M c_m^{(j)} x^{(m)}\right) = \sum_{m=1}^M c_m^{(j)} L(x^{(m)}) = 0. \quad (2.40)$$

Le caractère solénoïdal du champ de vitesse d'un fluide incompressible est donc par exemple préservé par les modes POD :  $\forall m = 1, \dots, M, \operatorname{div} x^{(m)} = 0 \implies \forall j = 1, \dots, q, \operatorname{div} \varphi^{(j)} = 0$ .

Le même raisonnement est employé pour déterminer quels types de conditions aux limites sont vérifiées par les modes POD. Soit  $\partial\Omega_D$  et  $\partial\Omega_N$  deux portions du bord du domaine  $\Omega$  de normale extérieure  $n$  sur lesquelles sont respectivement appliquées des conditions aux limites de type Dirichlet et Neumann, alors les variables du système dynamique vérifient à tout instant  $t \in \mathbb{R}_+$

$$x(\partial\Omega_D, t) = \gamma_D(t) \quad \text{et} \quad \frac{\partial x}{\partial n}(\partial\Omega_N, t) = \gamma_N(t). \quad (2.41)$$

À chaque instant  $t_m$ , les clichés sont contraints par les conditions limites précédentes. Les modes POD décrits par la décomposition (2.20) vérifient quant à eux les relations

$$\varphi^{(j)}(\partial\Omega_D) = \sum_{m=1}^M c_m^{(j)} \gamma_D(t_m) \quad \text{et} \quad \frac{\partial \varphi^{(j)}}{\partial n}(\partial\Omega_N) = \sum_{m=1}^M c_m^{(j)} \gamma_N(t_m) \quad (2.42)$$

qui correspondent en quelque sorte à une combinaison linéaire des conditions aux limites vérifiées par les clichés aux instants  $t_m$ . Les conditions aux limites sont cependant exactement vérifiées lorsqu'elles sont homogènes — *i. e.*  $\gamma_D(t) = 0$  et  $\gamma_N(t) = 0$  à tout instant. Ainsi, les conditions de non-glissement sur l'interface  $\Gamma$  entre la paroi et un fluide visqueux ( $x(\Gamma, t) = 0$ ) sont aussi vérifiées par les modes POD.

## 2.3 Un exemple d'application de la POD en analyse de données

Le principe de la décomposition orthogonale propre est illustré pour commencer avec un problème d'approximation d'un ensemble de données formé de vecteurs du plan. Le lien étroit entre cette méthode et une approximation du type moindres carrés sera ainsi clairement révélé.

Soit  $X = \{P_m\}_{m=1}^M$  un ensemble de points repérés par leurs coordonnées  $(x_m, y_m)$  dans le plan de l'espace  $H = \mathbb{R}^2$ . L'ensemble des points  $\{P_m\}_{m=1}^M$  constitue alors l'ensemble  $X$  des clichés pour lequel il s'agit de déterminer l'approximation fournie par la décomposition orthogonale propre. Le sous-espace de projection  $S$  de dimension  $q = 1 < N = 2$  recherché est engendré par un unique vecteur  $\vec{\varphi} \in \mathbb{R}^2$  tel que  $S = \text{Vect}\{\vec{\varphi}\}$ .

Le sous-espace  $S$  est donc une droite vectorielle sur laquelle les points  $P_m$  sont projetés. Pour mener à bien l'opération de projection, il est nécessaire de définir des vecteurs associés aux points  $P_m$  puisque le projecteur  $\mathcal{P}_S$  relatif au sous-espace  $S$  intervenant dans la formulation de la décomposition orthogonale propre agit sur les vecteurs de  $H = \mathbb{R}^2$  et non pas sur des points. On construit donc les vecteurs  $\vec{OP}_m$  et la projection des points  $P_m$  sur la droite vectorielle  $S$  est alors définie par  $\mathcal{P}_S \vec{OP}_m$ .

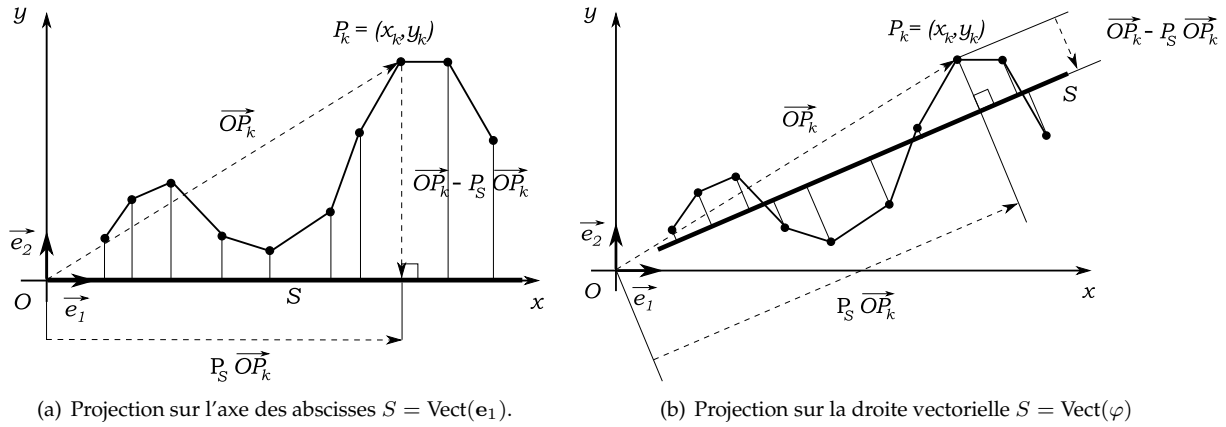


Figure 2.1 : Exemples de projection de l'ensemble de points  $\{P_m\}_{m=1}^M$  sur deux droites vectorielles.

La figure 2.1 donne deux exemples de sous-espaces de projection. On constate immédiatement par exemple qu'en choisissant  $\vec{\varphi} = \vec{e}_1$  l'axe des abscisses n'est pas adapté pour caractériser les données dans le sens où la distance entre les points  $\{P_m\}_{m=1}^M$  et la droite vectorielle formée par l'axe des abscisses n'est pas en moyenne minimale. Intuitivement, on observe que la « meilleure » approximation est obtenue en trouvant la droite qui minimise en moyenne l'écart entre les points initiaux et leurs projections sur un sous-espace à identifier. Cet écart, ou plutôt cette distance correspond à l'erreur de projection  $\epsilon_q$  qu'on cherche à minimiser dans la POD. Le vecteur optimal  $\vec{\varphi}$  est solution du problème de minimisation (2.3). Celui-ci se ramène au problème de maximisation (2.8) dont l'adaptation à l'exemple traité dans cette section donne lieu au problème ci-dessous.

$$\text{Trouver le vecteur } \vec{\varphi} \text{ tel que : } \vec{\varphi} = \arg \max_{\vec{\varphi} \in \mathbb{R}^2, \|\vec{\varphi}\|=1} E \left[ \left\| \langle \vec{OP}_m, \vec{\varphi} \rangle \vec{\varphi} \right\|^2 \right]. \quad (2.43)$$

L'opérateur de moyenne  $E[\cdot]$  correspond ici à la moyenne arithmétique

$$E \left[ f(\vec{OP}_m) \right] = \frac{1}{M} \sum_{m=1}^M f(\vec{OP}_m). \quad (2.44)$$

Le produit scalaire employé est le produit euclidien canonique. En utilisant les mêmes arguments variationnels que dans la méthode générale, le problème revient à résoudre l'équation aux valeurs propres  $R\vec{\varphi} = \lambda \vec{\varphi}$ . Pour cela, il faut tout d'abord détailler l'expression (2.11) de l'opérateur  $R$ , qui dans le cas présent se résume à une matrice de  $\mathcal{M}(2, 2)$ , notée  $\mathbf{R}$  et définie par

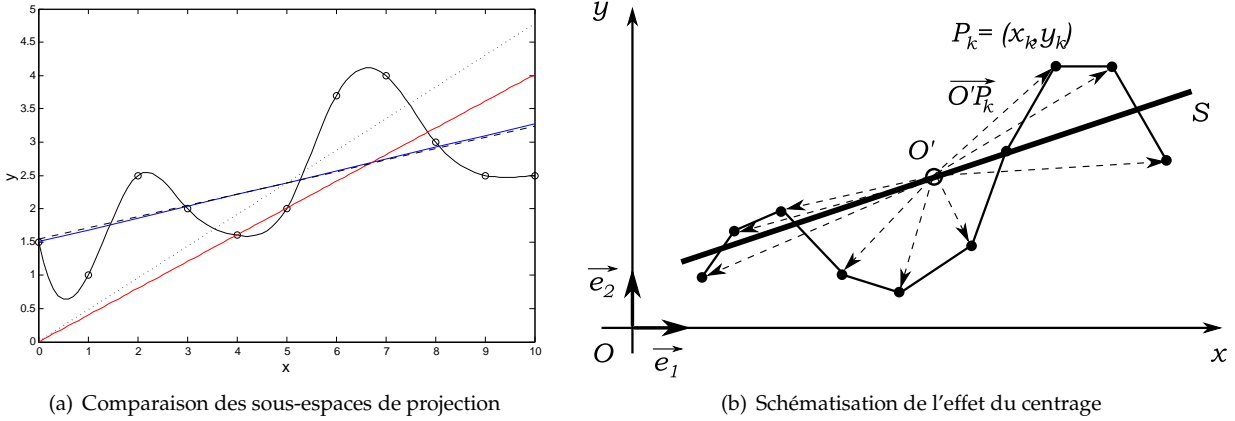


$$\mathbf{R}\vec{\varphi} = E \left[ \langle \overrightarrow{OP_m}, \vec{\varphi} \rangle \overrightarrow{OP_m} \right] = \frac{1}{M} \sum_{m=1}^M \begin{bmatrix} x_m x_m \varphi_1 + x_m y_m \varphi_2 \\ x_m y_m \varphi_1 + y_m y_m \varphi_2 \end{bmatrix} = \frac{1}{M} \sum_{m=1}^M \begin{bmatrix} x_m x_m & x_m y_m \\ y_m x_m & y_m y_m \end{bmatrix} \begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix}, \quad (2.45)$$

en notant  $\overrightarrow{OP_m} = [x_m \ y_m]^T$  et  $\vec{\varphi} = [\varphi_1 \ \varphi_2]^T$  la décomposition des vecteurs sur la base canonique de  $\mathbb{R}^2$ . La dernière égalité est particulièrement intéressante puisqu'elle permet d'identifier l'opérateur  $\mathbf{R}$  à partir de la matrice des clichés  $\mathbf{X} = [\overrightarrow{OP_1} \ \dots \ \overrightarrow{OP_M}]$  au moyen de l'expression

$$\mathbf{R} = \frac{1}{M} \sum_{m=1}^M \begin{bmatrix} x_m x_m & x_m y_m \\ y_m x_m & y_m y_m \end{bmatrix} = \frac{1}{M} \mathbf{X} \mathbf{X}^T. \quad (2.46)$$

Au passage, on peut noter que la forme de l'opérateur (2.46) correspond bien à l'expression du terme général<sup>3</sup> lorsque tous les poids de pondération  $\alpha_m$  sont constants et égaux à  $1/M$ . Connaissant à présent l'expression de  $\mathbf{R}$ , il ne reste plus qu'à en extraire les valeurs et vecteurs propres pour déterminer le vecteur  $\vec{\varphi}$  optimal pour la projection des données. La résolution du problème fournit autant de valeurs et de vecteurs propres orthogonaux que le rang de l'opérateur  $\mathbf{R}$ , soit deux en l'occurrence. Le sous-espace  $S$  recherché est alors caractérisé par le mode POD associé à la valeur propre la plus grande, *i. e.* celle pour laquelle le maximum d'« énergie » est captée.



**Figure 2.2 :** Effet du centrage des clichés sur l'approximation POD d'un ensemble de points de  $\mathbb{R}^2$ .

La méthode décrite ci-dessus est mise en application pour un ensemble de  $M = 11$  points du plan. Le résultat est tracé sur le graphique 2.2(a) où figurent plusieurs courbes. Les cercles correspondent à l'ensemble  $X$  des clichés  $\overrightarrow{OP_m}$  qui peuvent être interprétés comme un échantillonnage du signal continu tracé avec la ligne noire. La décomposition orthogonale propre est appliquée à la matrice  $\mathbf{X}$  qui représente l'ensemble  $X$  des clichés, ce qui revient à calculer les vecteurs propres de l'opérateur (2.46). Le premier vecteur propre correspond au vecteur directeur de la droite vectorielle approchant au mieux l'ensemble des données. Cette droite est représentée en rouge sur le graphique et correspond au sous-espace  $S = \text{Vect}\{\vec{\varphi}\}$  recherché.

L'approximation des données est relativement mauvaise, puisque la droite obtenue passe nécessairement par l'origine du repère. En procédant de la sorte, le mode POD prépondérant correspond grossièrement au vecteur

$$\overrightarrow{OO'} = E \left[ \overrightarrow{OP_m} \right] = \frac{1}{M} \sum_{m=1}^M \begin{bmatrix} x_m \\ y_m \end{bmatrix}, \quad (2.47)$$

où le point  $O'$  représente la position moyenne du nuage de points. La droite vectorielle engendrée par le vecteur  $\overrightarrow{OO'}$  est tracée en pointillés noirs sur la figure 2.2(a). L'accord entre les droites vectorielles  $S = \text{Vect}\{\vec{\varphi}\}$  et  $\text{Vect}\{\overrightarrow{OO'}\}$  n'est pas parfait mais la tendance montre que le premier mode POD est proche du vecteur moyen.

<sup>3</sup>Voir dans l'annexe B l'expression (B.3) du terme général de l'opérateur et sa forme matricielle donnée à l'équation (B.21) en dimension finie.

Ce phénomène est généralement observé lorsque la décomposition orthogonale propre est appliquée à des données de moyenne non nulle. À l'origine cependant, le théorème de Karhunen-Loève est formulé pour un ensemble de variables aléatoires « centrées », *i. e.* à moyenne nulle. Le choix de retrancher ou non la moyenne aux clichés avant de procéder à leur décomposition orthogonale propre influe alors sur les modes POD obtenus. Tamura *et al.* [224] et Gloerfelt [87] se sont ainsi intéressés au calcul des modes POD lorsque des données « brutes » ou des données « centrées » sont employées. Les résultats de Tamura *et al.* [224] révèlent que le spectre des valeurs propres est fortement affecté lorsque les données ne sont pas centrées : en effet, lorsque la moyenne est incluse dans les clichés, le premier mode POD correspond environ à la moyenne et capte près de 98% de l'énergie dans leur exemple, alors que lorsque la décomposition orthogonale propre est effectuée sur les données centrées, le premier mode ne représente plus que 40%. Lorsque la moyenne est soustraite, les modes POD sont plus régulièrement espacés et captent mieux les diverses structures du système en dehors de celle représentée par la moyenne. Gloerfelt [87] a observé le même type de comportement mais a signalé que l'inclusion de la moyenne dans les clichés peut se révéler parfois utile pour améliorer le comportement dynamique du modèle d'ordre réduit — lorsque celui-ci est construit — puisque la première équation du modèle d'ordre réduit gouverne l'évolution du champ moyen.

D'après les remarques précédentes, le nuage de points est projeté sur l'espace affine  $S' = \text{Vect}\{\overrightarrow{OO'} + \overrightarrow{\varphi}\}$ . La partie moyenne  $\overrightarrow{OO'}$  définie à l'équation (2.47) est donc retranchée à l'ensemble des clichés  $\overrightarrow{OP_m}$  avant d'en effectuer leur décomposition orthogonale propre. Le calcul du vecteur optimal  $\overrightarrow{\varphi}$  est alors mené pour les fluctuations  $\overrightarrow{O'P_m}$  autour de cette nouvelle origine comme le montre le schéma de la figure 2.2(b). La nouvelle droite vectorielle  $S' = \text{Vect}\{\overrightarrow{OO'} + \overrightarrow{\varphi}\}$  est tracée en bleu sur la figure 2.2(a). Dans ce cas, la droite de projection est très proche d'une régression linéaire classique tracée en traits discontinus noirs sur la figure 2.2(a). Strang [223], Chap. 3.2 propose par exemple de calculer le coefficient directeur  $a$  et l'ordonnée à l'origine  $b$  de la droite de régression  $y = ax + b$  en résolvant le système linéaire

$$\begin{bmatrix} M & \sum_{m=1}^M x_m \\ \sum_{m=1}^M x_m & \sum_{m=1}^M x_m^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{m=1}^M y_m \\ \sum_{m=1}^M x_m y_m \end{bmatrix}. \quad (2.48)$$

Le système précédent découle de la relation  $ax_m + b = y_m$  écrite pour  $m = 1, \dots, M$ . La détermination des coefficients  $a$  et  $b$  se fait alors en cherchant à minimiser l'erreur  $d_m = y_m - ax_m - b$  pour chaque point  $P_m$ , cette erreur correspondant alors à la distance verticale. La méthode de régression linéaire diffère légèrement de la décomposition orthogonale propre en ce sens que la direction de projection correspond dans le premier cas à l'axe des ordonnées tandis que dans le second cas la projection s'effectue selon la direction orthogonale à la droite vectorielle  $S$ . Les droites obtenues par la décomposition orthogonale propre des données centrées et la méthode de régression linéaire sont néanmoins très proches, ce qui confirme bien l'idée que la POD peut être vue comme une méthode de minimisation de type moindres carrés.

Ceci conclut cet exemple élémentaire d'application de la décomposition orthogonale propre, au moyen duquel l'importance de l'utilisation de clichés centrés a pu être mise en évidence. La structure de l'opérateur  $R$  a par ailleurs été exhibée pour un espace de Hilbert de type  $\mathbb{R}^N$  et a permis de retrouver l'expression générale établie à l'équation (B.21).

## 2.4 Principe de construction du modèle d'ordre réduit à partir de la base POD

Le principe de construction du modèle d'ordre réduit repose sur l'hypothèse que la solution — continue ou discrète — du système dynamique peut être approchée à chaque instant par une combinaison linéaire analogue à la décomposition (2.2). Ainsi, les premiers modes POD fournissent une approximation relativement précise de l'ensemble des clichés  $X$  et par extension on suppose qu'ils sont aussi valables pour représenter le comportement dynamique du système à tout instant  $t$  différent des instants  $t_m$  correspondants aux clichés. On attire ici l'attention sur le fait que les coefficients de la décomposition (2.2) correspondent aux amplitudes modales  $a_j^{(m)}$  définies par la relation (2.29), tandis que le modèle d'ordre réduit fournit un ensemble d'équations gouvernant les amplitudes modales inconnues  $a_j(t)$  à tout instant afin de pouvoir reconstruire la solution  $x(t)$  du système dynamique. La première étape de construction du modèle d'ordre réduit consiste donc à introduire la décomposition des variables dans les équations du système dynamique. Vient ensuite dans un second temps la phase de projection pour réduire le nombre d'équations.

Comme on l'a évoqué dans le paragraphe 1.3.4, plusieurs types de projection sont envisageables pour parvenir au modèle d'ordre réduit en fonction du degré de non-linéarité des équations du système dynamique initial. On distingue ainsi la *projection discrète*, la *projection sur le sous-espace* et la *projection de Galerkin* pour reprendre la terminologie de Lucia *et al.* [153]. Enfin une dernière technique consiste à identifier les coefficients du modèle d'ordre réduit à partir d'une forme pré-supposée.

### 2.4.1 Construction d'un modèle d'ordre réduit linéaire par projection discrète

La projection discrète est l'approche la plus simple pour obtenir le modèle d'ordre réduit d'un système dynamique *linéaire* pour lequel les équations discrétisées de la forme (1.2) sont disponibles. La discrétisation en espace des opérateurs du système dynamique s'écrit sous la forme générique

$$\mathbf{A}_1 \dot{\mathbf{x}}_h(t) + \mathbf{A}_0 \mathbf{x}_h(t) = \mathbf{b}. \quad (2.49)$$

Le vecteur  $\mathbf{x}_h$  contient l'ensemble des degrés de liberté  $x_i$  pour  $i = 1, \dots, N_v$  qui sont associés aux approximations de la variable continue  $x$  dans le domaine de calcul  $\Omega$ . Les matrices  $\mathbf{A}_1$  et  $\mathbf{A}_0$  contiennent les coefficients algébriques résultant de la discrétisation des opérateurs impliqués dans le système d'équations aux dérivées partielles associé à (1.1), et enfin  $\mathbf{b}$  est composé de coefficients algébriques provenant de la discrétisation des valeurs connues de  $\mathbf{x}_h$ , en particulier celles sur les conditions limites.

Si les clichés notés  $\mathbf{x}_h^{(m)}$  proviennent de la résolution numérique des équations précédentes, alors la résolution du problème aux valeurs propres de la méthode directe ou des clichés construit à partir des clichés discrets (voir le tableau B.1 p.224) fournit les modes POD  $\varphi^{(j)}$  sous forme discrète. La matrice tronquée des modes POD  $\Phi = [\varphi^{(1)}, \dots, \varphi^{(q)}]$  est alors employée pour décomposer directement le vecteur des degrés de liberté  $\mathbf{x}_h$  sur la base POD tronquée selon la relation

$$\mathbf{x}_h(t) \approx \Phi \mathbf{a}(t). \quad (2.50)$$

Le modèle d'ordre réduit gouvernant les amplitudes modales  $\mathbf{a} = [a_1, \dots, a_q]^T$  s'obtient alors en remplaçant la décomposition (2.50) dans les équations semi-discrètes (2.49) du modèle d'ordre élevé puis en projetant au moyen d'une simple multiplication matricielle :  $(\Phi^T \mathbf{A}_1 \Phi) \dot{\mathbf{a}}(t) + (\Phi^T \mathbf{A}_0 \Phi) \mathbf{a}(t) = \Phi^T \mathbf{b}$ . Le système réduit est alors explicitement déterminé par un système matriciel qui n'implique que des matrices de  $\mathcal{M}(q, q)$  et des vecteurs de  $\mathbb{R}^q$ , où  $q \ll N_v$ . Celui-ci s'écrit

$$\mathcal{M} \dot{\mathbf{a}}(t) = \mathcal{L} \mathbf{a}(t) + \mathcal{K} \quad \text{avec} \quad \begin{cases} \mathcal{M} = \Phi^T \mathbf{A}_1 \Phi \\ \mathcal{L} = -\Phi^T \mathbf{A}_0 \Phi \\ \mathcal{K} = \Phi^T \mathbf{b} \end{cases} \quad (2.51)$$

et découle de la projection *discrète* qui s'appuie sur l'approximation discrète (2.50) des degrés de liberté du système. La réduction opère donc dans l'espace euclidien réel  $H = \mathbb{R}^{N_v}$ .

Lorsque le système est non-linéaire, une pratique courante consiste à linéariser les équations du système afin de se ramener à une équation de la forme (2.49). C'est une technique souvent employée pour

traiter les problèmes d'aéroélasticité pour lesquels on cherche à déterminer les limites de stabilité en flottement par exemple. Le principe consiste à supposer que les quantités physiques sont caractérisées par de petites perturbations  $\mathbf{x}'$  autour d'un état stationnaire  $\mathbf{x}_s$  éventuellement non-linéaire, de sorte que le vecteur des degrés de liberté est décomposé en  $\mathbf{x}_h(t) = \mathbf{x}_s + \mathbf{x}'(t)$ . La partie stationnaire  $\mathbf{x}_s$  peut être calculée soit au moyen des équations linéarisées, soit à partir des équations non-linéaires. Ces deux approches correspondent respectivement aux modèles *entièrement linéaire* et *dynamiquement linéaire* décrits par Dowell *et al.* [60]. L'utilisation de l'écoulement stationnaire non-linéaire est particulièrement utile lorsqu'il s'agit de calculer un écoulement transsonique dans lequel les non-linéarités restent confinées à l'écoulement stationnaire tandis que les perturbations peuvent être considérées comme linéaires. Connaissant la partie stationnaire  $\mathbf{x}_s$ , la décomposition  $\mathbf{x}_h(t) = \mathbf{x}_s + \mathbf{x}'(t)$  est introduite dans les équations (2.52), puis le terme  $\mathbf{A}_0(\mathbf{x}_h(t))$  est linéarisé. Le système d'équations résultant a une forme similaire à celle donnée à l'équation (2.49) à la différence près que les inconnues sont à présent les fluctuations  $\mathbf{x}'$ . Le principal inconvénient de ce type de modèle d'ordre réduit pour les systèmes linéaires est la nécessité de recalculer la partie stationnaire à chaque fois qu'un paramètre est modifié.

## 2.4.2 Construction d'un modèle d'ordre réduit implicite par projection sur le sous-espace

Dans le cas d'un système non-linéaire, la semi-discrétisation conduit à l'écriture

$$\mathbf{A}_1 \dot{\mathbf{x}}_h(t) + \mathbf{A}_0(\mathbf{x}_h(t)) \mathbf{x}_h(t) = \mathbf{b}. \quad (2.52)$$

L'introduction de la décomposition (2.50) dans les équations semi-discrètes (2.52) et la projection sur la base tronquée des modes POD conduit à  $(\Phi^T \mathbf{A}_1 \Phi) \dot{\mathbf{a}}(t) + [\Phi^T \mathbf{A}_0(\Phi \mathbf{a}(t)) \Phi] \mathbf{a}(t) = \Phi^T \mathbf{b}$ . Le modèle d'ordre réduit ainsi obtenu est défini implicitement en raison du caractère non-linéaire de l'opérateur  $\mathbf{A}_0$  dont la projection produit un opérateur  $\tilde{\mathcal{L}}(\mathbf{a}(t))$  qui nécessite d'être évalué à chaque instant  $t$  puisqu'il dépend encore des amplitudes modales instationnaires  $\mathbf{a}(t)$ . Le modèle d'ordre réduit du système non-linéaire prend donc la forme implicite

$$\mathcal{M} \dot{\mathbf{a}}(t) = \tilde{\mathcal{L}}(\mathbf{a}(t)) \mathbf{a}(t) + \mathcal{K} \quad \text{avec} \quad \begin{cases} \mathcal{M} = \Phi^T \mathbf{A}_1 \Phi \\ \tilde{\mathcal{L}} = -\Phi^T \mathbf{A}_0(\Phi \mathbf{a}(t)) \Phi \\ \mathcal{K} = \Phi^T \mathbf{b} \end{cases} . \quad (2.53)$$

Deux alternatives peuvent être envisagées pour traiter le problème d'un système non-linéaire avec une approche discrète. La première consiste à évaluer effectivement l'opérateur  $\tilde{\mathcal{L}}(\mathbf{a}(t))$ , ce qui nécessite de recourir aux routines du modèle d'ordre élevé. La seconde alternative consiste à pratiquer une linéarisation du système (2.52) autour de l'état stationnaire  $\mathbf{x}_s$  comme on l'a expliqué dans le paragraphe précédent.

Lucia *et al.* [153] qualifient de *projection sur le sous-espace* ou de *projection sur le résidu* la première technique qui consiste à évaluer à chaque instant le terme non-linéaire. En effet, le résidu  $\mathbf{R}_h$  de l'équation semi-discrète (2.52) est défini à chaque instant par

$$\mathbf{R}_h(t) = \mathbf{A}_1 \dot{\mathbf{x}}_h(t) + \mathbf{A}_0(\mathbf{x}_h(t)) \mathbf{x}_h(t) - \mathbf{b} \quad (2.54)$$

et les équations (2.52) sont vérifiées lorsque le résidu est nul. Par analogie, la solution du modèle d'ordre réduit satisfait les équations (2.53) lorsque le résidu projeté  $\Phi^T \mathbf{R}_h(t)$  est nul.

Le principe de la projection sur le sous-espace repose alors sur l'intégration en temps du système (2.53) mais puisque le terme non-linéaire n'est pas explicitement connu pour les amplitudes modales, le champ non réduit  $\mathbf{x}_h$  est reconstruit à chaque instant avec la décomposition (2.50) pour évaluer l'opérateur non-linéaire au moyen des routines du système non réduit. La solution est alors calculée itérativement jusqu'à l'annulation du résidu projeté.

Les non-linéarités du système dynamique initial donc sont préservées puisque l'opérateur  $\tilde{\mathcal{L}}(\mathbf{a}(t))$  correspondant à la projection de  $\mathbf{A}_0(\mathbf{x}_h)$  est calculé tel quel. Le modèle d'ordre réduit est donc capable de reproduire n'importe quel type de non-linéarité. Toutefois, cette approche est intrusive puisque l'évaluation des termes non-linéaires est effectuée au moyen des routines du code de calcul du modèle d'ordre élevé. Par ailleurs, la réduction du temps de calcul n'est pas optimale puisqu'il est nécessaire à chaque pas de temps de reconstruire le vecteur  $\mathbf{x}_h$  à partir de la décomposition (2.50) afin de pouvoir évaluer le terme non-linéaire.

### 2.4.3 Construction d'un modèle d'ordre réduit polynomial par projection de Galerkin

Lorsque les non-linéarités du système dynamique sont « polynomiales », la *projection de Galerkin* des équations continues permet d'aboutir à une forme explicite des opérateurs non-linéaires de sorte que le modèle d'ordre réduit résultant soit optimal en terme de réduction du nombre de degrés de liberté et du temps de calcul.

À la différence de l'approche discrète, on considère la forme continue (1.1) du système dynamique. La variable continue  $x$  est alors décomposée sur la base tronquée des modes POD selon la relation

$$x(t) \approx \sum_{j=1}^q a_j(t) \varphi^{(j)}. \quad (2.55)$$

La méthode de Galerkin est un cas particulier de la méthode des résidus pondérés dont Bathe [26] donne un bref aperçu. Le principe consiste à approcher la solution continue de l'espace vectoriel  $\mathcal{V}$  de dimension infinie par une solution approchée définie dans l'espace  $\mathcal{V}_h$  ou dans l'espace  $\mathcal{V}_q$  (voir la figure 1.2). La méthode des Éléments Finis correspond par exemple à une approximation dans  $\mathcal{V}_h$ , alors que la projection sur la base des modes POD au moyen de la décomposition (2.55) représente une approximation dans  $\mathcal{V}_q$ . Bien que le principe de la projection soit identique dans les deux cas, la dimension des espaces de projection diffère de façon significative. La particularité de la projection de Galerkin provient de l'utilisation des mêmes fonctions  $\varphi^{(j)}$  pour l'approximation des variables continues et comme fonctions de pondération dans la projection. Cette dernière est réalisée au moyen du produit scalaire  $\langle \cdot, \cdot \rangle$  associé à l'espace de Hilbert  $H$  adopté pour décrire les solutions du système dynamique. Ainsi, la projection de Galerkin des équations (1.1) du système dynamique donne la relation

$$\langle \dot{x}(t), \psi \rangle = \langle f(x(t)), \psi \rangle \quad (2.56)$$

qui revient à minimiser l'erreur de projection du résidu  $R(t) = \dot{x}(t) - f(x, \mu)$ , d'où le nom de la méthode des résidus pondérés. Avec la projection de Galerkin, les fonctions de pondération  $\psi$  sont naturellement les modes POD de la base tronquée  $\Phi = \{ \varphi^{(i)} \}_{i=1}^q$  qui sont orthonormaux entre eux par construction. L'introduction de la décomposition (2.55) dans la projection (2.56) simplifie alors légèrement l'expression en

$$\dot{a}_i(t) = \left\langle f \left( \sum_{j=1}^q a_j(t) \varphi^{(j)} \right), \varphi^{(i)} \right\rangle = \widehat{f}_i(\mathbf{a}(t)). \quad (2.57)$$

Le second membre  $\widehat{f}_i(\mathbf{a}(t))$  de l'expression précédente est encore exprimé implicitement en fonction des amplitudes modales. Toutefois, lorsque le second membre  $f$  du système dynamique est linéaire ou multilinéaire, la bilinéarité du produit scalaire peut être mise à profit pour dériver une expression du modèle d'ordre réduit dans laquelle apparaissent explicitement les composantes du vecteur des amplitudes modales  $\mathbf{a}(t)$ . On considère dans le cadre de cette thèse que le second membre  $f$  du système dynamique s'écrit sous la forme d'une somme d'opérateurs multilinéaires, ce qui conduit donc à reformuler le second membre avec l'expression

$$f(x) = \widetilde{f}_0 + \sum_{n \geq 1} \widetilde{f}_n(x, \dots, x), \quad (2.58)$$

où la fonction  $\widetilde{f}_0$  représente un terme constant et les fonctions multilinéaires  $\widetilde{f}_n$  sont définies de  $H^n$  dans  $H$  pour un  $n$ -uplet  $(x, \dots, x)$ . Cette transformation permet alors d'introduire la décomposition (2.50) pour chaque terme du  $n$ -uplet. La transformation (2.58) est toujours possible lorsque les termes de  $f$  s'écrivent sous forme « polynomiale », dans le sens où  $f$  est seulement défini à partir de produits des variables  $x$  sur lesquels agissent éventuellement les opérateurs aux dérivées partielles spatiales. Ainsi, la divergence  $\operatorname{div} x$  sera considérée comme une fonction linéaire  $\widetilde{f}_1(x)$ , tandis que l'opérateur d'advection  $x \cdot \nabla x$  correspond à un terme quadratique  $\widetilde{f}_2(x, x)$ .

Sous l'hypothèse que le second membre du système dynamique s'écrive sous la forme (2.58), l'introduction de la décomposition (2.55) dans le produit scalaire  $\langle f(x), \varphi^{(i)} \rangle$  donne

$$\langle f(x), \varphi^{(i)} \rangle = \langle \widetilde{f}_0, \varphi^{(i)} \rangle + \sum_{n \geq 1} \left\langle \widetilde{f}_n \left( \sum_{j_1=1}^q a_{j_1}(t) \varphi^{(j_1)}, \dots, \sum_{j_n=1}^q a_{j_n}(t) \varphi^{(j_n)} \right), \varphi^{(i)} \right\rangle. \quad (2.59)$$

Finalement, puisque les fonctions  $\tilde{f}_n$  sont multilinéaires et que le produit scalaire est une application bilinéaire, les sommes dans les arguments du  $n$ -uplet des fonctions  $\tilde{f}_n$  peuvent être extraites et le modèle d'ordre réduit prend la forme explicite

$$a_i(t) = \langle \tilde{f}_0, \varphi^{(i)} \rangle + \sum_{n \geq 1} \sum_{j_1, \dots, j_n=1}^q \langle \tilde{f}_n(\varphi^{(j_1)}, \dots, \varphi^{(j_n)}), \varphi^{(i)} \rangle a_{j_1}(t) \dots a_{j_n}(t). \quad (2.60)$$

Le modèle d'ordre réduit est complètement défini par les coefficients  $\langle \tilde{f}_n(\varphi^{(j_1)}, \dots, \varphi^{(j_n)}), \varphi^{(i)} \rangle$  qui sont à présent indépendants des amplitudes modales  $\mathbf{a}(t)$ . Il est donc possible de calculer les coefficients dans un premier temps avant l'intégration en temps du modèle d'ordre réduit contrairement à l'approche de projection sur le sous-espace où le calcul des termes  $\mathcal{L}(\mathbf{a}(t))$  doit être effectué à chaque pas de temps. Le système d'équations (2.60) offre donc un gain substantiel en terme de réduction du nombre de degrés de liberté et du temps de calcul.

L'obtention de la forme (2.60) requiert toutefois que le second membre du système dynamique s'écrive comme une somme d'opérateurs multilinéaires. Dans le cas contraire, il n'est pas possible d'extraire les opérateurs de sommation et les non-linéarités ne peuvent être explicitées. Le recours à la projection sur le sous-espace est donc indispensable dans ces situations, à moins d'effectuer un développement de Taylor des termes non-linéaires, auquel cas la forme du second membre  $f$  peut être interprétée comme ici en tant que somme d'opérateurs multilinéaires.

Avant de poursuivre, la forme du modèle d'ordre réduit est fournie lorsque le second membre présente des non-linéarités au plus quadratiques puisque les systèmes dynamiques rencontrés dans le cadre de cette thèse seront ramenés à cette forme dans la mesure du possible. Dans ce cas  $n \in \llbracket 1; 2 \rrbracket$  et la relation (2.60) se simplifie en

$$a_i(t) = \langle \tilde{f}_0, \varphi^{(i)} \rangle + \sum_{j=1}^q \langle \tilde{f}_1(\varphi^{(j)}), \varphi^{(i)} \rangle a_j(t) + \sum_{j,k=1}^q \langle \tilde{f}_2(\varphi^{(j)}, \varphi^{(k)}), \varphi^{(i)} \rangle a_j(t) a_k(t). \quad (2.61)$$

L'expression linéaire (2.51) est donc généralisée par la relation quadratique

$$a_i(t) = \mathcal{K}_i + \sum_{j=1}^q \mathcal{L}_{ij} a_j(t) + \sum_{j,k=1}^q \mathcal{Q}_{ijk} a_j(t) a_k(t). \quad (2.62)$$

#### 2.4.4 Construction d'un modèle d'ordre réduit par identification des coefficients

Une dernière manière de construire le modèle d'ordre réduit consiste à identifier les coefficients plutôt que de les calculer analytiquement à partir des expressions issues de la projection des opérateurs du système dynamique sur les modes POD. La principale difficulté consiste à déterminer dans un premier temps la structure des équations du modèle d'ordre réduit dont on cherche les coefficients. Le choix de la projection de Galerkin des équations du modèle d'ordre élevé est particulièrement judicieux puisqu'il découle d'un processus physique. Une fois la forme du modèle d'ordre réduit connue, les coefficients sont déterminés comme les solutions d'un problème inverse.

En pratique, le recours à la décomposition orthogonale propre est capital pour réduire le nombre de paramètres à déterminer. Par ailleurs, l'hypothèse d'un modèle d'ordre réduit polynomial est encore une fois fondamentale puisque les coefficients sont alors facilement calculés comme les solutions d'un problème moindres carrés *linéaire*. La méthode a déjà été employée par exemple par Perret *et al.* [177] pour obtenir un modèle d'ordre réduit à partir de données expérimentales. Puisque les clichés dont ils disposaient provenaient de résultats expérimentaux pour lesquels aucun système d'équations n'était associé en particulier, les auteurs avaient donc supposé que le modèle d'ordre réduit s'écrivait de la manière la plus générale qui soit comme un système d'équations différentielles ordinaires polynomial au maximum cubique. Lorsque cela est possible, les coefficients identifiés seront de préférence ceux correspondant à la projection de Galerkin afin de préserver le lien physique avec le modèle d'ordre élevé.

Cette approche sera qualifiée de méthode d'identification des coefficients du modèle réduit, pour reprendre la terminologie employée par Perret *et al.* [177]. Elle a très récemment été reprise et améliorée

[48, 244], notamment pour régulariser le problème inverse à résoudre. La méthode est robuste puisque les erreurs liées à la troncature de la base POD, ainsi que les erreurs numériques ou de modélisation qui dégradent la réponse des modèles d'ordre réduit POD-Galerkin n'interviennent plus ici. Les coefficients sont directement évalués de manière à s'adapter à la forme pré-supposée des équations et par conséquent aucune phase de correction a posteriori n'est nécessaire. Un autre avantage de la méthode est que le temps de calcul alloué à la construction du modèle d'ordre réduit et notamment à l'évaluation des coefficients du modèle d'ordre réduit est considérablement diminué.

La construction d'un modèle d'ordre réduit par identification des coefficients repose toujours sur la décomposition orthogonale propre d'un ensemble  $X$  de clichés  $x^{(m)}$ , à partir duquel les modes POD  $\varphi^{(i)}$  et les amplitudes modales  $a_i^{(m)}$  sont déterminés. Si l'on suppose que la forme du modèle d'ordre réduit est donnée par le système d'équations quadratiques (2.62) par exemple<sup>4</sup>, l'objectif est alors de déterminer les coefficients  $\mathcal{K}_i$ ,  $\mathcal{L}_{ij}$  et  $\mathcal{Q}_{ijk}$  de manière à ce que la dérivée des amplitudes  $\dot{a}_i^{(m)}$  issues de la décomposition (2.2) des clichés coïncide avec le second membre  $\widehat{f}_i(\mathbf{a}^{(m)})$  évalué pour les amplitudes correspondant aux clichés aux mêmes instants  $t_m$ . L'idée sous-jacente est que si les dérivées des amplitudes coïncident, alors l'intégration en temps du modèle d'ordre réduit devrait pouvoir fournir les bonnes amplitudes modales  $a_i(t_m)$ . En supposant que les amplitudes modales  $a_i^{(m)}$  correspondant aux clichés satisfont les équations du modèle d'ordre réduit, la relation suivante peut être écrite à chaque instant  $t_m$  pour  $m = 1, \dots, M$  :

$$\dot{a}_i^{(m)} = \widehat{f}_i(a_j^{(m)}) = \mathcal{K}_i + \sum_{j=1}^q \mathcal{L}_{ij} a_j^{(m)} + \sum_{j,k=1}^q \mathcal{Q}_{ijk} a_j^{(m)} a_k^{(m)}. \quad (2.63)$$

Les amplitudes modales de référence  $a_i^{(m)}$  sont connues à chaque instant  $t_m$  puisqu'elles ont été déterminées en même temps que les modes POD lors de la décomposition orthogonale propre des clichés. Les dérivées en temps des amplitudes modales peuvent ensuite être évaluées de sorte que les termes  $\dot{a}_i^{(m)}$  soient aussi connus. Si l'échantillonnage des clichés est assez faible, ou si le schéma de dérivation en temps est d'ordre assez bas, la dérivée des amplitudes modales peut être évaluée en deux étapes : les amplitudes modales connues aux instants  $t_m$  sont d'abord interpolées au moyen de splines par exemple, puis ces fonctions sont dérivées en temps au moyen d'un schéma aux différences finies centrées.

Dans le problème (2.63), les inconnues sont donc les coefficients du modèle d'ordre réduit  $\mathcal{K}_i$ ,  $\mathcal{L}_{ij}$  et  $\mathcal{Q}_{ijk}$ . L'équation gouvernant chaque amplitude modale est gouvernée par 1 coefficient constant  $\mathcal{K}_i$ ,  $q$  coefficients linéaires  $\mathcal{L}_{ij}$  et  $q^2$  coefficients quadratiques  $\mathcal{Q}_{ijk}$ . En fait, la moitié des termes quadratiques est redondante puisque les contributions  $a_j^{(m)} a_k^{(m)}$  sont identiques à celles produites par  $a_k^{(m)} a_j^{(m)}$ . Par conséquent, on considère seulement les  $q(q+1)/2$  premiers coefficients quadratiques et le vecteur des inconnues de dimension  $N_{c,i} = 1 + q + q(q+1)/2$  est donc

$$\mathcal{X}_i = \left[ \mathcal{K}_i \quad \mathcal{L}_{i1} \quad \dots \quad \mathcal{L}_{ij} \quad \dots \quad \mathcal{L}_{iq} \quad \mathcal{Q}_{i11} \quad \dots \quad \mathcal{Q}_{ijk, k \geq j} \quad \dots \quad \mathcal{Q}_{iqq} \right]^T. \quad (2.64)$$

La relation (2.63) fournissant la dérivée des amplitudes modales à chaque instant d'échantillonnage s'écrit alors sous la forme d'une équation linéaire en fonction du vecteur des inconnues :  $\dot{a}_i^{(m)} = \sum_{l=1}^{N_{c,i}} \mathcal{A}_l(\mathbf{a}^{(m)}) \mathcal{X}_{il}$  où  $\mathcal{X}_{il}$  est la  $l$ -ième composante du vecteur  $\mathcal{X}_i$ . Le regroupement des  $M$  relations précédentes écrites aux instants  $t_m$  conduit alors à la forme matricielle

$$\boxed{\mathcal{A} \mathcal{X}_i = \mathcal{B}_i}. \quad (2.65)$$

La matrice  $\mathcal{A} \in \mathcal{M}(M, N_{c,i})$  contient les corrélations temporelles entre les différentes amplitudes évaluées aux instants  $t_m$  et le vecteur  $\mathcal{B}_i \in \mathbb{R}^M$  contient les dérivées des amplitudes modales, tels que

$$\mathcal{A} = \begin{bmatrix} 1 & a_1^{(1)} & \dots & a_q^{(1)} & a_1^{(1)} a_1^{(1)} & \dots & a_q^{(1)} a_q^{(1)} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & a_1^{(M)} & \dots & a_q^{(M)} & a_1^{(M)} a_1^{(M)} & \dots & a_q^{(M)} a_q^{(M)} \end{bmatrix} \quad \text{et} \quad \mathcal{B}_i = \begin{bmatrix} \dot{a}_i^{(1)} \\ \vdots \\ \dot{a}_i^{(M)} \end{bmatrix}. \quad (2.66)$$

<sup>4</sup>Tant que le second membre est composé d'opérateurs multilinéaires, la méthode décrite dans ce qui suit reste valable et conduit toujours à la résolution d'un problème linéaire, même lorsque le modèle d'ordre réduit est décrit par un système polynomial d'ordre supérieur tel que celui donné à l'équation (2.60).

Les coefficients inconnus sont donc solutions d'un problème linéaire du type (A.20) qui est résolu au moyen des méthodes décrites dans le paragraphe A.3 p.214. Lorsque le système est mal posé et que la matrice  $\mathcal{A}$  est mal conditionnée, le problème est résolu de façon approchée au moyen du pseudo-inverse  $\mathcal{A}^+$  défini par la relation (A.24) et la solution est approchée par

$$\tilde{\mathcal{X}}_i = \mathcal{A}_p^+ \mathcal{B}_i. \quad (2.67)$$

Il est important de rappeler que si la forme générique choisie pour le modèle d'ordre réduit n'est pas polynomiale, alors il n'est pas possible de formuler le problème sous la forme (2.65), et la détermination des coefficients optimaux pour reproduire les amplitudes modales doit se faire par le biais d'algorithmes d'optimisation plus complexes [33, 143, 159]. Toutefois, le problème devient rapidement très lourd à traiter lorsque le nombre de coefficients augmente et la méthode n'est plus viable.

Une seconde remarque concerne l'échantillonnage des clichés : si l'intervalle  $\delta t_c = t_{m+1} - t_m$  entre deux clichés est trop important, alors les dérivées temporelles des amplitudes modales  $\dot{a}_j^{(m)}$  risquent d'être mal évaluées. Par conséquent le second membre  $\mathcal{B}_i$  du problème (2.65) est peu précis et ceci est d'autant plus grave que la matrice  $\mathcal{A}$  est en général mal conditionnée, ce qui rend la résolution du problème extrêmement sensible aux perturbations du second membre. Pour plus de précautions, on a donc choisi ici d'interpoler d'abord les amplitudes modales avec des splines avant de procéder à la dérivation en temps, de manière à ce que les schémas différences finies employés puissent être appliqués pour des pas de temps plus petits que celui associé à l'échantillonnage.

Pour terminer, on précise la terminologie employée dans la suite de ce document. L'approche POD-Galerkin consiste à calculer les coefficients du modèle d'ordre réduit à partir des expressions analytiques établies comme la projection des opérateurs du système non réduit sur les modes POD, alors que la méthode d'identification consiste à optimiser les coefficients pour reproduire la réponse du système non réduit. Il est important de mentionner toutefois que la projection formelle de Galerkin sur les modes POD est une étape indispensable puisque dans le premier cas elle fournit les expressions analytiques tandis que dans le second elle donne la forme générique des équations qui serviront à identifier les coefficients.

Les deux approches sont en quelque sorte complémentaires et ne diffèrent en réalité que dans l'utilisation qui est faite des modes POD et des amplitudes modales issus de la décomposition orthogonale propre pour construire le modèle d'ordre réduit. Ainsi, l'approche POD-Galerkin repose sur l'emploi des modes POD pour calculer les coefficients du modèle d'ordre réduit, alors que les coefficients sont identifiés à partir des (dérivées des) amplitudes modales dans la seconde méthode. Les deux méthodes sont donc d'une certaine manière « duales » puisque la construction du modèle d'ordre réduit avec chacune d'entre elles repose soit sur les modes POD, soit sur les amplitudes modales. Or ces deux quantités contiennent chacune des informations quant au comportement du système : les modes POD représentent les structures spatiales d'énergie maximale dans la réponse du système, alors que les amplitudes modales décrivent plutôt le comportement dynamique de la réponse. L'idéal serait donc de disposer d'une méthode de construction du modèle d'ordre réduit qui tire parti à la fois des informations spatiales et temporelles. Sur ce sujet, il serait intéressant d'approfondir les méthodes de construction de modèles d'ordre réduit basées sur les modes d'interaction principaux qui ont été brièvement décrites au paragraphe 1.3.5.



## 2.5 Décomposition orthogonale propre des systèmes multivariés

Dans la section 2.2, la décomposition orthogonale propre a été décrite dans le contexte d'un système caractérisé par une unique variable  $x$  contenue dans un espace de Hilbert  $H$ . Cependant, les systèmes dynamiques complexes tels que ceux rencontrés en mécanique des fluides sont décrits par un ensemble de  $n_v$  variables indépendantes gouvernées chacune par une équation du type (1.1). On note alors  $\mathbf{q} = [x_1, \dots, x_{n_v}]^\top$  le vecteur des inconnues du système dynamique qui appartient à l'espace  $H^{n_v}$ . Le système dynamique (1.1) est donc remplacé par le système multivariable

$$\dot{\mathbf{q}}(t) = \mathbf{f}(\mathbf{q}) \iff \begin{cases} \dot{x}_1 = f_1(x_1, \dots, x_{n_v}) \\ \vdots \\ \dot{x}_{n_v} = f_{n_v}(x_1, \dots, x_{n_v}) \end{cases}. \quad (2.68)$$

De même que lorsque le système est monovariable, un ensemble  $Q$  de clichés  $\mathbf{q}^{(m)}$  est généré soit par simulation numérique, soit par des dispositifs expérimentaux. La question du traitement des différentes variables du vecteur  $\mathbf{q}$  se pose alors lorsqu'il s'agit d'en effectuer une décomposition orthogonale propre : faut-il, puisqu'il y a à présent plusieurs variables, calculer un jeu de modes POD pour chaque variable, ou, au contraire, extraire des modes « globaux » calculés à partir de l'ensemble des variables ? Rowley [192] compare les deux méthodes et qualifie la décomposition orthogonale propre de *scalaire* dans le premier cas puisque chaque variable  $x_k$  pour  $k = 1, \dots, n_v$  est traitée séparément comme un scalaire possédant sa propre décomposition, et de *vectorielle* dans le second cas en raison de la décomposition qui porte directement sur le vecteur  $\mathbf{q}$  des variables considéré dans son ensemble.

D'après les remarques formulées dans l'exemple de la section 2.3, la décomposition orthogonale propre sera de préférence pratiquée pour les clichés centrés dans le reste de ce document. On définit par conséquent la *partie moyenne* des clichés notée

$$\bar{\mathbf{q}} = E[\mathbf{q}^{(m)}] = \sum_{m=1}^M \alpha_m \mathbf{q}^{(m)} \quad (2.69)$$

et obtenue par application de l'opérateur de moyenne  $E[\cdot]$  aux clichés de l'ensemble  $Q$ . La décomposition orthogonale propre n'est donc plus appliquée à l'ensemble des clichés  $Q = \{\mathbf{q}^{(m)}\}_{m=1}^M$  mais à l'ensemble  $Q_c$  des clichés centrés  $\mathbf{s}^{(m)}$  définis par soustraction de la partie moyenne par

$$\mathbf{s}^{(m)} = \mathbf{q}^{(m)} - \bar{\mathbf{q}}, \quad \forall m \in \llbracket 1; M \rrbracket. \quad (2.70)$$

### 2.5.1 Description de l'approche POD scalaire

L'ensemble  $Q_c$  des clichés centrés est décomposé en  $n_v$  ensembles  $Q_{c,k}$  relatifs au champ de la variable  $x_k^{(m)}$  prélevé aux instants  $t_m$  et centré pour donner les  $s_k^{(m)} = x_k^{(m)} - \bar{x}_k$  qui constituent les  $n_v$  composantes du vecteur  $\mathbf{s}^{(m)}$  défini ci-dessus. La *partie moyenne* pour chaque variable  $x_k$  est obtenue par application de l'opérateur de moyenne  $E[\cdot]$  aux clichés  $x_k^{(m)}$  des ensembles  $Q_{c,k}$ . La partie moyenne est alors donnée par

$$\bar{x}_k = E[x_k^{(m)}] = \sum_{m=1}^M \alpha_m x_k^{(m)}. \quad (2.71)$$

Avec l'approche POD scalaire, chaque variable possède son propre jeu de modes POD qui proviennent de la recherche des valeurs et vecteurs propres des  $n_v$  opérateurs de corrélation  $\mathbf{R}_{s_k}^*$  définis pour les variables  $s_k$  à partir de la relation (2.25) si la méthode des clichés est adoptée par

$$\mathbf{R}_{s_k}^* \mathbf{d}_{s_k}^{(j)} = \lambda_{j,s_k} \mathbf{d}_{s_k}^{(j)} \quad \text{avec} \quad (j, k) \in \llbracket 1; r \rrbracket \times \llbracket 1; n_v \rrbracket. \quad (2.72)$$

Les opérateurs aux valeurs propres sont construits pour les clichés centrés et le terme général s'écrit alors  $R_{i,j,s_k}^* = \sqrt{\alpha_i \alpha_j} \langle x_k^{(i)} - \bar{x}_k, x_k^{(j)} - \bar{x}_k \rangle_H$ . Les vecteurs propres  $\mathbf{d}_{s_k}^{(j)}$  permettent alors de revenir aux coefficients  $\mathbf{c}_{s_k}^{(j)} = \tilde{\alpha} \mathbf{d}_{s_k}^{(j)}$  utilisés dans la décomposition (2.20) adaptée ici en  $\varphi_{x_k}^{(j)} = \sum_{m=1}^M c_{m,s_k}^{(j)} s_k^{(m)}$  pour les clichés

centrés. Les modes POD définis pour chaque variable s'écrivent alors comme la combinaison linéaire des clichés centrés  $s_k^{(m)}$  pondérés par les composantes du vecteur  $\mathbf{c}_{s_k}^{(j)}$ , de sorte que

$$\varphi_{x_k}^{(j)} \approx \sum_{m=1}^M c_{m,x_k}^{(j)} \left( x_k^{(m)} - \bar{x}_k \right). \quad (2.73)$$

La décomposition orthogonale propre tronquée utilisant seulement les  $q$  premiers modes est adaptée à partir de l'équation (2.55) pour chaque variable centrée  $s_k$ , ce qui conduit à la relation

$$x_k(t) = \bar{x}_k + \sum_{j=1}^q a_j^{x_k}(t) \varphi_{x_k}^{(j)}. \quad (2.74)$$

La décomposition pour le vecteur  $\mathbf{q}$  des variables du système dynamique s'écrit de manière synthétique

$$\mathbf{q}(t) \approx \bar{\mathbf{q}} + \Phi \mathbf{a}(t). \quad (2.75)$$

Le vecteur  $\bar{\mathbf{q}} = [\bar{x}_1, \dots, \bar{x}_{n_v}]^\top$  regroupe les parties moyennes des différentes variables et correspond à la définition (2.69). La matrice  $\Phi$  et le vecteur  $\mathbf{a}$  sont définis respectivement par

$$\Phi = \begin{bmatrix} \Phi_{x_1} & & 0 \\ & \ddots & \\ 0 & & \Phi_{x_{n_v}} \end{bmatrix} = \begin{bmatrix} \varphi_{x_1}^{(1)} \dots \varphi_{x_1}^{(q)} & & 0 \\ & \ddots & \\ 0 & & \varphi_{x_{n_v}}^{(1)} \dots \varphi_{x_{n_v}}^{(q)} \end{bmatrix} \quad \text{et} \quad \mathbf{a} = \begin{bmatrix} \mathbf{a}^{x_1} \\ \vdots \\ \mathbf{a}^{x_{n_v}} \end{bmatrix}. \quad (2.76)$$

Il est important de noter que la matrice  $\Phi \in \mathcal{M}(n_v, qn_v)$  n'est pas à proprement parler la matrice de la base POD, mais seulement une matrice diagonale par blocs qui regroupe les matrices des modes POD  $\Phi_{x_k} = [\varphi_{x_k}^{(1)} \dots \varphi_{x_k}^{(q)}]$  relatives à chaque variable. Ainsi, les colonnes des matrices  $\Phi_{x_k}$  correspondent bien à la base des modes POD ce qui n'est pas le cas de celles de  $\Phi$ .

Le modèle d'ordre réduit est construit en remplaçant la décomposition (2.74) de chaque variable dans l'équation continue du système dynamique relative à cette variable. Le système (1.1) prend la forme (2.68) dans le cas d'un système multivariables. On suppose alors que les seconds membres  $f_k(\mathbf{q})$  s'écrivent comme la somme d'opérateurs multilinéaires, à la manière de la décomposition (2.58). En particulier, on suppose ici pour simplifier que le système est au plus quadratique et chacun des opérateurs  $f_n$  est encore décomposé en une somme d'opérateurs portant exclusivement sur chaque  $x_k$ . L'équation aux dérivées partielles régissant la variable  $x_i$  s'écrit alors

$$\dot{x}_i = K_i + \sum_{n=1}^{n_v} L_{in}(x_n) + \sum_{n,p=1}^{n_v} Q_{inp}(x_n, x_p). \quad (2.77)$$

À présent il est possible de remplacer la décomposition scalaire (2.74) de chaque variable dans l'équation précédente. Les propriétés de multilinéarité des opérateurs permettent de développer l'expression obtenue et enfin la projection sur chaque mode POD  $\varphi_{x_i}^{(j)}$  conduit à l'égalité

$$\begin{aligned} \dot{a}_j^{x_i} &= \left\langle K_i + \sum_{n=1}^{n_v} L_{in}(\bar{x}_n) + \sum_{n,p=1}^{n_v} Q_{inp}(\bar{x}_n, \bar{x}_p), \varphi_{x_i}^{(j)} \right\rangle \\ &+ \left\langle \sum_{k=1}^q \left[ \sum_{n=1}^{n_v} L_{in}(\varphi_{x_n}^{(k)}) a_k^{x_n} + \sum_{n,p=1}^{n_v} \left( Q_{inp}(\bar{x}_n, \varphi_{x_p}^{(k)}) a_k^{x_p} + Q_{inp}(\varphi_{x_n}^{(k)}, \bar{x}_p) a_k^{x_n} \right) \right], \varphi_{x_i}^{(j)} \right\rangle \\ &+ \left\langle \sum_{k,l=1}^q \sum_{n,p=1}^{n_v} Q_{inp}(\varphi_{x_n}^{(l)}, \varphi_{x_p}^{(k)}) a_l^{x_n} a_k^{x_p}, \varphi_{x_i}^{(j)} \right\rangle. \end{aligned} \quad (2.78)$$

Finalement, un changement d'indice dans la deuxième contribution du terme linéaire permet de modifier les amplitudes  $a_k^{x_p}$  en  $a_k^{x_n}$ , de sorte que le modèle d'ordre réduit s'écrit

$$\dot{a}_j^{x_i} = \mathcal{K}_j + \sum_{k=1}^q \sum_{n=1}^{n_v} \mathcal{L}_{jk}^{x_n} a_k^{x_n} + \sum_{k,l=1}^q \sum_{n,p=1}^{n_v} \mathcal{Q}_{jkl}^{x_n x_p} a_l^{x_n} a_k^{x_p}. \quad (2.79)$$

Les coefficients du modèle d'ordre réduit sont donnés par les relations suivantes :

$$\begin{aligned}
 \mathcal{K}_j &= \left\langle K_i + \sum_{n=1}^{n_v} L_{in}(\bar{x}_n) + \sum_{n,p=1}^{n_v} Q_{inp}(\bar{x}_n, \bar{x}_p), \varphi_{x_i}^{(j)} \right\rangle, \\
 \mathcal{L}_{jk}^{x_n} &= \left\langle L_{in}(\varphi_{x_n}^{(k)}) + \sum_{p=1}^{n_v} \left( Q_{ipn}(\bar{x}_p, \varphi_{x_n}^{(k)}) + Q_{inp}(\varphi_{x_n}^{(k)}, \bar{x}_p) \right), \varphi_{x_i}^{(j)} \right\rangle, \\
 \mathcal{Q}_{jkl}^{x_n x_p} &= \left\langle Q_{inp}(\varphi_{x_n}^{(l)}, \varphi_{x_p}^{(k)}), \varphi_{x_i}^{(j)} \right\rangle.
 \end{aligned} \tag{2.80}$$

En effectuant la décomposition orthogonale propre de chaque variable, le modèle d'ordre réduit se compose d'un système d'équations différentielles ordinaires couplées. En effet, les  $q$  équations sur les amplitudes  $a_j^{x_i}$  relatives à la variable  $x_i$  ne peuvent être résolues indépendamment des autres modèles d'ordre réduit puisque les équations font intervenir toutes les amplitudes modales, y compris celles correspondant à d'autres variables  $x_n \neq x_i$ . Par conséquent, il faut résoudre simultanément un ensemble de  $q n_v$  équations, où  $q$  est le nombre de modes POD retenus dans la base de projection et  $n_v$  est le nombre de variables indépendantes du système dynamique. Il serait toutefois possible d'optimiser la taille du modèle d'ordre réduit en ne choisissant pas forcément le même nombre  $q$  de modes POD pour chaque variable  $x_i$ . Ceci pourrait éventuellement être efficace si certaines grandeurs physiques peuvent être représentées par un nombre très restreint de modes alors que d'autres en demandent davantage.

L'approche POD scalaire a été employée principalement par Lucia et Beran [152] et Moreno *et al.* [169]. L'avantage de la méthode est de fournir une décomposition pour chaque variable, pour laquelle les modes POD représentent les structures d'énergie maximale. Cependant, le système d'équations différentielles ordinaires obtenu est relativement compliqué et la formulation du modèle d'ordre réduit devient lourde. Un second reproche peut être formulé vis-à-vis de cette approche : en procédant de la sorte, une partie de la physique du problème est négligée puisque les opérateurs de corrélation sont formés à partir de clichés contenant exclusivement une quantité physique donnée, alors que les différentes variables sont corrélées entre elles. Ainsi, dans le cas d'un fluide incompressible les modes POD ne correspondent plus aux structures d'énergie cinétique maximale, mais aux « énergies cinétiques » dans chaque direction. Il est donc préférable de former un opérateur aux valeurs propres à partir des clichés contenant l'ensemble des grandeurs physiques du problème, même si cela induit d'autres difficultés comme on va le voir.

### 2.5.2 Description de l'approche POD vectorielle

À l'opposé de la méthode précédente, on considère l'ensemble  $Q_c$  des clichés centrés sans aucune distinction entre les différentes variables. L'unique problème aux valeurs propres construit pour des clichés centrés  $s^{(m)}$  contenant l'ensemble des variables du système à différents instants  $t_m$  est donc défini par

$$\mathbf{R}^* \mathbf{d}^{(j)} = \lambda_j \mathbf{d}^{(j)} \quad \text{avec } j \in \llbracket 1; r \rrbracket. \tag{2.81}$$

L'opérateur aux valeurs propres est formulé pour les clichés centrés et le terme général s'écrit donc  $R_{ij}^* = \sqrt{\alpha_i \alpha_j} \langle \mathbf{q}^{(i)} - \bar{\mathbf{q}}, \mathbf{q}^{(j)} - \bar{\mathbf{q}} \rangle_{H^{n_v}}$ . À la différence de l'approche POD scalaire, le produit scalaire est cette fois celui associé à  $H^{n_v}$  et non plus à  $H$ . Les vecteurs propres  $\mathbf{d}^{(j)}$  solutions du problème (2.81) fournissent alors les coefficients  $\mathbf{c}^{(j)} = \tilde{\alpha} \mathbf{d}^{(j)}$  utilisés dans la décomposition

$$\varphi^{(j)} = \sum_{m=1}^M c_m^{(j)} (\mathbf{q}^{(m)} - \bar{\mathbf{q}}). \tag{2.82}$$

Il est toujours possible de décomposer le vecteur des modes POD en différentes contributions relatives à chaque variable, en posant :  $\varphi^{(j)} = [\varphi_{x_1}^{(j)} \dots \varphi_{x_{n_v}}^{(j)}]^\top$ . Bien qu'on retrouve le même type de fonctions que pour la POD scalaire, il est primordial de signaler que les  $\varphi_{x_i}^{(j)}$  ne sont plus des modes POD puisqu'ils ne découlent pas directement de la résolution du problème aux valeurs propres (2.81). En particulier, la propriété d'orthonormalité des modes POD est vérifiée par le vecteur des modes POD  $\varphi^{(j)}$  mais pas par les composantes  $\varphi_{x_i}^{(j)}$  relatives à chaque variable.

La décomposition POD est donc appliquée cette fois directement au vecteur des inconnues  $\mathbf{q}$  plutôt qu'à chaque composante  $x_i$ , ce qui conduit à l'approximation

$$\mathbf{q}(t) \approx \bar{\mathbf{q}} + \sum_{j=1}^q a_j(t) \boldsymbol{\varphi}^{(j)}. \quad (2.83)$$

À la différence de l'approche scalaire, il n'existe qu'un seul jeu d'amplitudes modales  $a_j(t)$  pour l'ensemble des variables du système au lieu d'en avoir un par variable. La décomposition orthogonale propre d'une variable  $x_i$  donnée correspond alors à la  $i$ -ième ligne de (2.83) et s'écrit  $x_i(t) = \bar{x}_i + \sum_{j=1}^q a_j(t) \varphi_{x_i}^{(j)}$ . Cependant, à la différence de (2.74), les amplitudes ne dépendent plus de  $x_i$  et les  $\varphi_{x_i}^{(j)}$  ne correspondent plus aux modes POD mais seulement à la partie du vecteur  $\boldsymbol{\varphi}^{(j)}$  relative à la variable  $x_i$ . Ici, l'écriture matricielle synthétique de la décomposition orthogonale propre des variables du système devient

$$\mathbf{q}(t) \approx \bar{\mathbf{q}} + \boldsymbol{\Phi} \mathbf{a}(t). \quad (2.84)$$

L'expression est similaire à (2.75) mais la structure de la matrice  $\boldsymbol{\Phi}$  et du vecteur  $\mathbf{a}$  est plus simple. En effet,  $\boldsymbol{\Phi} \in \mathcal{M}(n_v, q)$  est une matrice pleine et le vecteur des amplitudes modales ne possède plus que  $q$  composantes identiques quelle que soit la variable  $x_i$ , d'où les définitions

$$\boldsymbol{\Phi} = [\boldsymbol{\varphi}^{(1)} \dots \boldsymbol{\varphi}^{(q)}] = \begin{bmatrix} \varphi_{x_1}^{(1)} & \dots & \varphi_{x_1}^{(q)} \\ \vdots & & \vdots \\ \varphi_{x_{n_v}}^{(1)} & \dots & \varphi_{x_{n_v}}^{(q)} \end{bmatrix} \quad \text{et} \quad \mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_q \end{bmatrix}. \quad (2.85)$$

La construction du modèle d'ordre réduit à partir des équations (2.68) est grandement facilitée puisque la décomposition orthogonale propre est écrite directement pour le vecteur des inconnues  $\mathbf{q}$ . En supposant que le second membre  $\mathbf{f}(\mathbf{q})$  s'écrit à la manière de la décomposition (2.58) comme la somme d'opérateurs multilinéaires au plus quadratiques, le système dynamique prend la forme vectorielle

$$\dot{\mathbf{q}} = \mathbf{K} + \mathbf{L}(\mathbf{q}) + \mathbf{Q}(\mathbf{q}, \mathbf{q}). \quad (2.86)$$

Il suffit alors de remplacer la décomposition (2.84) dans l'équation (2.86) et d'utiliser la multilinéarité des opérateurs  $\mathbf{L}$  et  $\mathbf{Q}$  directement par rapport au vecteur  $\mathbf{q}$ . On aboutit alors à l'égalité

$$\begin{aligned} \sum_{k=1}^q \dot{a}_k \boldsymbol{\varphi}^{(k)} &= \mathbf{K} + \mathbf{L}(\bar{\mathbf{q}}) + \mathbf{Q}(\bar{\mathbf{q}}, \bar{\mathbf{q}}) \\ &+ \sum_{k=1}^q \mathbf{L}(\boldsymbol{\varphi}^{(k)}) a_k + \sum_{k=1}^q \mathbf{Q}(\bar{\mathbf{q}}, \boldsymbol{\varphi}^{(k)}) a_k + \sum_{l=1}^q \mathbf{Q}(\boldsymbol{\varphi}^{(l)}, \bar{\mathbf{q}}) a_l \\ &+ \sum_{k,l=1}^q \mathbf{Q}(\boldsymbol{\varphi}^{(k)}, \boldsymbol{\varphi}^{(l)}) a_k a_l. \end{aligned} \quad (2.87)$$

Enfin la projection de l'équation précédente sur les  $q$  modes POD  $\boldsymbol{\varphi}^{(j)}$  conduit aux équations suivantes du modèle d'ordre réduit :

$$\dot{a}_j = \mathcal{K}_j + \sum_{k=1}^q \mathcal{L}_{jk} a_k + \sum_{k,l=1}^q \mathcal{Q}_{jkl} a_k a_l. \quad (2.88)$$

Les coefficients du modèle d'ordre réduit sont alors donnés par les relations

$$\begin{aligned} \mathcal{K}_j &= \left\langle \mathbf{K} + \mathbf{L}(\bar{\mathbf{q}}) + \mathbf{Q}(\bar{\mathbf{q}}, \bar{\mathbf{q}}), \boldsymbol{\varphi}^{(j)} \right\rangle, \\ \mathcal{L}_{jk} &= \left\langle \mathbf{L}(\boldsymbol{\varphi}^{(k)}) + \mathbf{Q}(\bar{\mathbf{q}}, \boldsymbol{\varphi}^{(k)}) + \mathbf{Q}(\boldsymbol{\varphi}^{(k)}, \bar{\mathbf{q}}), \boldsymbol{\varphi}^{(j)} \right\rangle, \\ \mathcal{Q}_{jkl} &= \left\langle \mathbf{Q}(\boldsymbol{\varphi}^{(l)}, \boldsymbol{\varphi}^{(k)}), \boldsymbol{\varphi}^{(j)} \right\rangle. \end{aligned} \quad (2.89)$$

La forme des équations est proche de celle obtenue par l'approche scalaire donnée à l'équation (2.79), mais est largement simplifiée par l'indépendance des amplitudes et des coefficients vis-à-vis de chaque

variable  $x_i$ . Un des avantages de l'approche vectorielle est qu'il est inutile de décomposer les opérateurs  $\mathbf{K}$ ,  $\mathbf{L}$  et  $\mathbf{Q}$  en différentes contributions exclusivement relatives à un  $x_i$  donné puisque la décomposition orthogonale propre porte directement sur le vecteur  $\mathbf{q}$  dont dépendent les opérateurs.

Rowley [192] a comparé au chapitre 5 de sa thèse les approches POD scalaire et vectorielle. Outre le fait que l'approche vectorielle ne nécessite que de résoudre  $q$  équations au lieu de  $q n_v$  dans l'approche scalaire, il a remarqué que le modèle d'ordre réduit obtenu avec l'approche scalaire a plus facilement tendance à diverger et le comportement ne s'améliore pas lorsque le nombre  $q$  de modes POD contenus dans la base de projection est augmenté. La difficulté principale de l'approche vectorielle est de déterminer le produit scalaire  $\langle \cdot, \cdot \rangle$  associé à l'espace  $H^{n_v}$ . Mathématiquement, les définitions basiques du produit scalaire pour l'espace euclidien de dimension finie  $(\mathbb{R}^N)^{n_v} = \mathbb{R}^{N n_v}$  avec  $N_v = N n_v$  et l'espace de Lebesgue des fonctions vectorielles de carré intégrable  $(L^2(\Omega))^{n_v}$  ne posent pas de problème. Ainsi, le produit scalaire euclidien entre deux clichés discrets tels que  $\mathbf{q} = [x_{1,1} \cdots x_{N,1} \cdots x_{1,n_v} \cdots x_{N,n_v}]^T$  s'écrit

$$\forall(\mathbf{q}, \mathbf{r}) \in \mathbb{R}^{N n_v} \times \mathbb{R}^{N n_v}, \quad \langle \mathbf{q}, \mathbf{r} \rangle = \sum_{i=1}^{N n_v} q_i r_i = \sum_{j=1}^N \sum_{k=1}^{n_v} q_{j,k} r_{j,k}. \quad (2.90)$$

De la même manière, le produit scalaire de deux vecteurs de fonctions continues  $\mathbf{q} = [x_1 \cdots x_{n_v}]^T$  s'écrit

$$\forall(\mathbf{q}, \mathbf{r}) \in (L^2(\Omega))^{n_v} \times (L^2(\Omega))^{n_v}, \quad \langle \mathbf{q}, \mathbf{r} \rangle = \int_{\Omega} \sum_{k=1}^{n_v} q_k(\xi) r_k(\xi) d\xi. \quad (2.91)$$

Les définitions précédentes correspondent certes à un produit scalaire, mais la somme des quantités  $q_{j,k} r_{j,k}$  et  $q_k(\xi) r_k(\xi)$  pour différentes valeurs de  $k$  n'est pas consistante du point de vue des unités physiques si les variables  $x_i$  du vecteur  $\mathbf{q}$  ont des unités différentes. Ce problème n'est pas anodin et la détermination d'un produit scalaire correct est primordial puisque la définition même des modes POD, et par voie de conséquence la qualité de la base de projection, dépend de ce choix. Des détails supplémentaires à propos de ce choix seront donnés plus tard lorsque les équations du fluide seront présentées.

## 2.6 Synthèse du chapitre 2

Dans ce chapitre, la décomposition orthogonale propre a été décrite dans le contexte d'un système déterministe en s'appuyant sur le principe d'optimalité que doit vérifier la base des modes POD afin de minimiser l'erreur moyenne de projection. Après avoir rappelé l'écriture des problèmes aux valeurs propres correspondant à la méthode directe et des clichés, les principales propriétés des modes POD et des amplitudes modales ont été rappelées. Un exemple très simple dans lequel un nuage de point de l'espace  $\mathbb{R}^2$  est approché a permis de mettre en œuvre la décomposition orthogonale propre et d'attirer l'attention sur l'importance de travailler avec des clichés centrés, *i. e.* à moyenne nulle. Le lien étroit avec une méthode classique de régression linéaire a par ailleurs été mis en évidence.

Le principe de construction d'un modèle d'ordre réduit au moyen de la base POD a ensuite été présenté. Quatre techniques ont été décrites pour traiter des systèmes linéaires ou non-linéaires. Le cas particulier des systèmes dynamiques non-linéaires pour lesquels les opérateurs peuvent être considérés comme multilinéaires a ouvert la voie à la construction explicite de modèles d'ordre réduit non-linéaires via une projection de Galerkin. En procédant de la sorte, une réduction optimale en terme de nombre de degrés de liberté et de coût de calcul est réalisée. La technique de projection sur le sous-espace permet certes de traiter n'importe quel type de non-linéarité, mais il s'agit d'une méthode intrusive qui ne permet pas réellement de réduction en terme de nombre de degrés de liberté. Finalement, une dernière technique de construction du modèle d'ordre réduit par identification des coefficients du modèle d'ordre réduit a été décrite. Celle-ci utilise encore l'hypothèse que le système dynamique peut s'écrire en fonction d'opérateurs multilinéaires afin de formuler un problème d'approximation moindres carrés linéaires. Les coefficients calculés de cette manière évitent un certain nombre d'erreurs de calcul qui conduisent parfois à la déstabilisation des modèles d'ordre réduit POD-Galerkin et par conséquent elle apparaît comme une méthode robuste. Enfin, l'application de la décomposition orthogonale propre à un système multivariables a été détaillée avec les approches POD scalaires et vectorielles. Les avantages et les inconvénients des deux méthodes ont été rappelés et l'importance de la définition du produit scalaire a été évoquée.

Ces premières remarques concernant les aspects théoriques de la construction d'un modèle d'ordre réduit au moyen de la base des modes POD permettent de mieux appréhender la méthode qui sera à l'avenir mise en œuvre dans le cas de systèmes mécaniques plus complexes. On s'intéressera en particulier dans le chapitre suivant au cas d'un système dynamique linéaire pour lequel différentes formulations de la décomposition orthogonale propre sont comparées pour commencer, avant de chercher à prendre en compte l'action d'une condition limite de Dirichlet non-homogène.



# Évaluation de l'approche POD-Galerkin pour les systèmes dynamiques linéaires non-autonomes

## Sommaire

<b>3.1</b>	<b>Introduction</b>	<b>64</b>
<b>3.2</b>	<b>Modèles d'ordre réduit pour les systèmes dynamiques linéaires autonomes</b>	<b>65</b>
3.2.1	Formulation analytique du modèle d'ordre réduit	65
3.2.1.1	Génération de clichés à partir de la solution analytique du problème	66
3.2.1.2	Calcul des modes POD	66
3.2.1.3	Construction du modèle d'ordre réduit analytique	67
3.2.2	Formulation discrète du modèle d'ordre réduit	68
3.2.2.1	Génération de clichés à partir de la solution Éléments Finis du problème	68
3.2.2.2	Calcul des modes POD	69
3.2.2.3	Construction du modèle d'ordre réduit par une approche discrète	70
3.2.3	Formulation POD-Galerkin du modèle d'ordre réduit	70
3.2.3.1	Approximation numérique du produit scalaire	71
3.2.3.2	Calcul des modes POD	71
3.2.3.3	Construction du modèle d'ordre réduit par une approche POD-Galerkin	72
<b>3.3</b>	<b>Évaluation du modèle d'ordre réduit POD-Galerkin du système linéaire autonome</b>	<b>74</b>
3.3.1	Évaluation du modèle d'ordre réduit POD-Galerkin	74
3.3.2	Choix de la base des clichés et du nombre de modes POD	77
<b>3.4</b>	<b>Modèles d'ordre réduit pour les systèmes dynamiques linéaires non-autonomes</b>	<b>82</b>
3.4.1	Adaptation de la méthode des pénalités au modèle d'ordre réduit	83
3.4.2	Adaptation de la méthode des multiplicateurs de Lagrange au modèle d'ordre réduit	84
3.4.3	Méthode de la fonction de contrôle	85
<b>3.5</b>	<b>Évaluation du modèle d'ordre réduit POD-Galerkin du système linéaire non-autonome</b>	<b>87</b>
3.5.1	Aspect des modes POD non-homogènes et homogénéisés	87
3.5.2	Réponse du modèle d'ordre réduit non-autonome	87
<b>3.6</b>	<b>Synthèse du chapitre 3</b>	<b>93</b>



### 3.1 Introduction

On se propose dans ce chapitre d'évaluer l'approche POD-Galerkin présentée dans la section 2.4 lorsqu'elle est employée pour la construction du modèle d'ordre réduit d'un système dynamique linéaire. L'objectif est dans un premier temps de comparer la qualité de la réponse reproduite à celle fournie par une formulation analytique, ainsi qu'à la réponse du modèle d'ordre réduit construit par la projection discrète qui devrait être particulièrement adaptée puisqu'il s'agit d'un système linéaire. L'exemple d'application choisi ici est volontairement extrêmement simple afin de pouvoir comparer les résultats à la formulation analytique qui constituera notre référence.

Dans un second temps, la formulation POD-Galerkin sera mise à l'épreuve pour reproduire la réponse du système rendu non-autonome par l'introduction d'une condition limite de Dirichlet non-homogène. Ce type de condition limite est choisi intentionnellement puisqu'il correspond à celui rencontré dans les problèmes d'interaction fluide-structure et notamment d'aéroélasticité pour traiter la condition de couplage cinématique à l'interface entre le milieu fluide et solide. Plusieurs techniques sont alors développées et comparées pour tenir compte de cette condition limite non-homogène. Enfin, la capacité du modèle d'ordre réduit à reconstruire des réponses de divers types ou pour différentes valeurs d'un paramètre à partir d'une unique base de modes POD sera brièvement étudiée.

## 3.2 Développement de modèles d'ordre réduit basés sur la POD pour les systèmes dynamiques linéaires autonomes

Plusieurs formulations d'un modèle d'ordre réduit basé sur la décomposition orthogonale propre sont développées dans cette section pour un système linéaire autonome au moyen de la projection discrète et de la projection de Galerkin. Ces deux formulations supposent respectivement de considérer que les clichés dont on dispose sont discrets ou continus. Les performances des modèles d'ordre réduit sont alors comparées à celles d'une formulation analytique de référence qui peut être développée pour l'exemple simple traité ici. Ce dernier consiste en l'étude des déplacements d'une barre encastée-libre unidimensionnelle, dont la géométrie est décrite sur la figure 3.1 : la barre est de longueur  $L$  et le matériau supposé linéaire élastique est modélisé par sa masse volumique  $\rho$  et son module de Young  $E$  tels que  $c = \sqrt{E/\rho}$  corresponde à la célérité du son dans le milieu. Sous l'hypothèse de petits mouvements, l'équation gouvernant les déplacements de la barre se résume à l'équation des ondes unidirectionnelle lorsque les paramètres structuraux et géométriques sont supposés constants [83]. Celle-ci s'écrit

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0. \quad (3.1)$$

Il s'agit donc d'un système dynamique linéaire monovarié ( $n_v = 1$ ) dans lequel l'inconnue  $u$  désigne le déplacement longitudinal. L'état du système noté auparavant  $x$  correspond à présent au déplacement  $u$  et la notation  $x$  sera réservée à la direction spatiale associée à une base du domaine de calcul  $\Omega = [0; L] \subset \mathbb{R}$ .



Figure 3.1 : Schématisation du problème de la barre en petits mouvements de traction-compression.

Pour spécifier une solution unique, l'équation (3.1) est complétée par les conditions aux limites

$$\begin{cases} u(x=0, t) = 0 \\ \left. \frac{\partial u}{\partial x} \right|_{x=L} = 0 \end{cases} \quad \forall t > 0, \quad (3.2)$$

ainsi que les conditions initiales<sup>1</sup>

$$\begin{cases} u(x, t=0) = (U_0/L)x \\ \left. \frac{\partial u}{\partial t} \right|_{t=0} = 0 \end{cases} \quad \forall x \in [0; L]. \quad (3.3)$$

Dans les paragraphes qui suivent, trois formulations du modèle d'ordre réduit sont détaillées : tout d'abord une formulation analytique dans laquelle les clichés et l'équation de base employée pour projeter le modèle d'ordre réduit sont continus, ensuite une formulation discrète reposant sur un formalisme Éléments Finis où les clichés et l'équation de base du modèle d'ordre réduit sont liés à la discrétisation du système et enfin une formulation POD-Galerkin utilisant des clichés discrets et l'équation continue de base.

### 3.2.1 Formulation analytique du modèle d'ordre réduit

Cette première formulation est développée uniquement dans le but de servir de référence ensuite pour évaluer la qualité des réponses produites avec les autres formulations. La construction des modèles d'ordre réduit pour le système dynamique autonome étudié ici n'a en soi qu'un intérêt très limité en dehors de l'évaluation des performances, puisque la solution exacte est connue.

<sup>1</sup>La condition initiale portant sur les déplacements correspond à la solution statique du problème pour une déformée  $U_0$  appliquée à l'extrémité libre : en effet, il faut satisfaire  $\partial^2 u / \partial x^2 = 0$  avec les conditions aux limites  $u(x=0) = 0$  et  $u(x=L) = U_0$ , ce qui, après deux intégrations, donne bien  $U_0 x/L$ .

### 3.2.1.1 Génération de clichés à partir de la solution analytique du problème

En l'absence d'efforts extérieurs, il est possible de montrer que la solution analytique du problème (3.1) avec les conditions limites (3.2) et les conditions initiales (3.3) s'écrit sous la forme d'une série infinie de fonctions trigonométriques donnant lieu à la décomposition

$$u(x, t) = \sum_{n=1}^{\infty} \zeta_n \Theta_n(t) \psi_n(x) = \sum_{n=1}^{\infty} \left[ (-1)^{n-1} \frac{4\sqrt{2L}U_0}{(2n-1)^2\pi^2} \right] \cos(k_n t) \sqrt{\frac{2}{L}} \sin\left(\frac{k_n x}{c}\right). \quad (3.4)$$

L'expression précédente s'obtient à partir de la méthode de séparation des variables rappelée par Lallane *et al.* [138] par exemple. Les nombres d'ondes  $k_n = \pi c(2n-1)/(2L)$  donnent la pulsation des modulations temporelles  $\Theta_n(t)$  et des variations spatiales de chaque mode  $\psi_n(x)$ . Enfin, les constantes  $\zeta_n$  sont déterminées de manière à satisfaire les conditions initiales (3.3).

Il est intéressant de remarquer que le principe d'obtention de la décomposition (3.4) découle d'une séparation des variables analogue à celle qui a été pratiquée pour rechercher les modes propres au paragraphe 1.4. Les modes  $\psi_n$  seront donc appelés les *modes propres* du système dynamique, à ne pas confondre avec les modes POD qui seront construits dans la suite. Les fonctions  $\psi_n$  sont par ailleurs orthonormales entre elles pour le produit scalaire associé à l'espace de Hilbert  $H = L^2([0; L])$ , d'où la propriété

$$\langle \psi_n, \psi_m \rangle_{L^2} = \frac{2}{L} \int_0^L \sin\left(\frac{k_n x}{c}\right) \sin\left(\frac{k_m x}{c}\right) dx = \delta_{nm}. \quad (3.5)$$

La fréquence fondamentale du système, notée  $T_0$ , est donnée par la relation

$$T_0 = 2\frac{\pi}{k_1} = \frac{4L}{c}. \quad (3.6)$$

En pratique seul un nombre limité  $p$  de fonctions de base est utilisé pour caractériser la réponse. La série tronquée représente alors une approximation de la solution analytique.

### 3.2.1.2 Calcul des modes POD

Le problème aux valeurs propres induit par la décomposition orthogonale propre est formulé au moyen des solutions analytiques (3.4) de sorte que les clichés  $\{u^{(m)} = u(x, t_m), m = 1, \dots, M\}$  soient des éléments de l'espace  $L^2([0; L])$ . L'expression analytique (3.4) des clichés s'écrit alors

$$u^{(m)} = \sum_{n=1}^{\infty} \zeta_n^{(m)} \psi_n(x), \quad (3.7)$$

où  $\zeta_n^{(m)} = \zeta_n \cos(k_n t_m)$  est la modulation temporelle à l'instant  $t_m$  du mode propre  $\psi_n$ .

La première étape de la construction du modèle d'ordre réduit consiste à rechercher les modes POD du système au moyen de la méthode directe ou de la méthode des clichés (voir §2.2). Les deux méthodes étant équivalentes, on décrit simplement le calcul des modes POD au moyen de la méthode des clichés. On procède de préférence à la décomposition orthogonale propre des clichés centrés  $s^{(m)} = u^{(m)} - \bar{u}$ . Puisque le système est monovisible, la partie moyenne  $\bar{u}$  est donnée par l'équation (2.71) et le problème aux valeurs propres est celui donné à l'équation (2.72). Comme pour les clichés, la partie moyenne  $\bar{u}$  est aussi décomposée sur la base des modes propres afin d'introduire les modulations moyennes  $\bar{\zeta}_n$  qui apparaissent naturellement dans l'expression

$$\bar{u} = \sum_{m=1}^M \alpha_m \left( \sum_{n=1}^{\infty} \zeta_n^{(m)} \psi_n \right) = \sum_{n=1}^{\infty} \left( \sum_{m=1}^M \alpha_m \zeta_n^{(m)} \right) \psi_n = \sum_{n=1}^{\infty} \bar{\zeta}_n \psi_n. \quad (3.8)$$

Le terme général  $R_{ij}^*$  de l'opérateur de corrélation se simplifie ici au moyen de l'expression analytique (3.7) des clichés et de la propriété d'orthonormalité (3.5) des modes propres  $\psi_n$ . On obtient alors

$$R_{ij}^* = \sqrt{\alpha_i \alpha_j} \langle s^{(i)}, s^{(j)} \rangle = \sqrt{\alpha_i \alpha_j} \sum_{n=1}^{\infty} \left( \zeta_n^{(i)} - \bar{\zeta}_n \right) \left( \zeta_n^{(j)} - \bar{\zeta}_n \right). \quad (3.9)$$

La somme dans l'expression analytique (3.4) est généralement tronquée pour ne conserver que  $p < \infty$  fonctions de base, ce qui permet d'introduire les vecteurs des clichés  $\zeta^{(m)} = [\zeta_1^{(m)}, \dots, \zeta_p^{(m)}]^\top$  et de la partie moyenne  $\bar{\zeta} = [\bar{\zeta}_1, \dots, \bar{\zeta}_p]^\top$  puis la matrice des clichés centrés  $\mathbf{S} \in \mathcal{M}(p, M)$  dont le terme général est  $S_{ij} = \zeta_i^{(j)} - \bar{\zeta}_i$ . L'opérateur de corrélation de la méthode des clichés s'exprime donc au moyen du produit matriciel

$$\mathbf{R}^* = \tilde{\alpha}^\top \mathbf{S}^\top \mathbf{S} \tilde{\alpha} \quad (3.10)$$

dans lequel la matrice  $\tilde{\alpha} = \text{diag}(\sqrt{\alpha_1}, \dots, \sqrt{\alpha_M})$  contient les poids de pondération de l'opérateur de moyenne. La résolution du problème aux valeurs propres (2.72) fournit les vecteurs propres  $\mathbf{d}^{(j)}$  à partir desquels les coefficients  $c^{(j)}$  sont déterminés. Les modes POD sont alors obtenus au moyen de la combinaison linéaire (2.73). Par ailleurs, la décomposition (3.7) des clichés sur les modes propres fournit l'expression suivante des modes POD en fonction des modes propres  $\psi_n$  :

$$\varphi^{(j)} = \sum_{m=1}^M c_m^{(j)} (u^{(m)} - \bar{u}) = \sum_{m=1}^M c_m^{(j)} \sum_{n=1}^{\infty} (\zeta_n^{(m)} - \bar{\zeta}_n) \psi_n = \sum_{n=1}^{\infty} \left( \sum_{m=1}^M c_m^{(j)} (\zeta_n^{(m)} - \bar{\zeta}_n) \right) \psi_n. \quad (3.11)$$

Par conséquent les modes POD se décomposent analytiquement sur la base des modes propres comme en témoigne la relation

$$\varphi^{(j)} = \sum_{n=1}^{\infty} \mu_n^{(j)} \psi_n \quad \text{avec} \quad \mu_n^{(j)} = \sum_{m=1}^M c_m^{(j)} (\zeta_n^{(m)} - \bar{\zeta}_n). \quad (3.12)$$

En définissant  $\boldsymbol{\mu} \in \mathcal{M}(p, q)$  la matrice des coefficients  $\mu_n^{(j)}$  et  $\mathbf{C} \in \mathcal{M}(M, q)$  la matrice des coefficients  $c_m^{(j)}$ , on obtient  $\boldsymbol{\mu} = \mathbf{S}\mathbf{C}$ . Une interprétation des liens entre les modes propres et les modes POD a été proposée par Feeny et Kappagantu [74] et Kerschen et Golinval [129] dans le cas de systèmes linéaires similaires à celui étudié ici.

### 3.2.1.3 Construction du modèle d'ordre réduit analytique

La solution du système dynamique est alors approchée par la décomposition

$$u(x, t) \approx \bar{u}(x) + \sum_{j=1}^q a_j(t) \varphi^{(j)}(x) \quad (3.13)$$

qui correspond à l'équation (2.74). L'introduction de cette décomposition dans l'équation (3.1) et la projection de Galerkin sur le mode POD  $\varphi^{(i)}$  avec le produit scalaire associé à l'espace  $H = L^2([0; L])$  conduit à l'égalité

$$\left\langle \frac{\partial^2}{\partial t^2} \left( \bar{u}(x) + \sum_{j=1}^q a_j(t) \varphi^{(j)}(x) \right), \varphi^{(i)}(x) \right\rangle - \left\langle c^2 \frac{\partial^2}{\partial x^2} \left( \bar{u}(x) + \sum_{j=1}^q a_j(t) \varphi^{(j)}(x) \right), \varphi^{(i)}(x) \right\rangle = 0. \quad (3.14)$$

La partie moyenne  $\bar{u}$  étant indépendante du temps, le terme  $\partial \bar{u} / \partial t$  disparaît. Par ailleurs, puisque les modes POD sont orthonormaux entre eux, le premier produit scalaire se réduit à  $\ddot{a}_i$ . Le second produit scalaire devient quant à lui

$$\left\langle c^2 \frac{\partial^2}{\partial x^2} \left( \bar{u}(x) + \sum_{j=1}^q a_j(t) \varphi^{(j)}(x) \right), \varphi^{(i)}(x) \right\rangle = c^2 \left\langle \frac{d^2 \bar{u}}{dx^2}, \varphi^{(i)} \right\rangle + c^2 \sum_{j=1}^q \left\langle \frac{d^2 \varphi^{(j)}}{dx^2}, \varphi^{(i)} \right\rangle a_j(t). \quad (3.15)$$

Pour terminer, une expression analytique des dérivées secondes est établie pour simplifier les produits scalaires précédents. En remplaçant la partie moyenne  $\bar{u}$  et le mode POD  $\varphi^{(j)}$  par leurs décompositions respectives (3.8) et (3.12) sur la base des modes propres, on aboutit aux expressions

$$\frac{d^2 \bar{u}}{dx^2} = - \sum_{n=1}^{\infty} \left( \frac{k_n}{c} \right)^2 \bar{\zeta}_n \psi_n \quad \text{et} \quad \frac{d^2 \varphi^{(j)}}{dx^2} = - \sum_{n=1}^{\infty} \left( \frac{k_n}{c} \right)^2 \mu_n^{(j)} \psi_n. \quad (3.16)$$

Ainsi, les produits scalaires entre les dérivées précédentes et le mode POD  $\varphi^{(i)}$  se simplifient en introduisant la décomposition (3.12). L'orthonormalité des modes propres est alors mise à profit pour simplifier le produit scalaire subsistant :  $\langle \psi_n, \varphi^{(i)} \rangle = \mu_n^{(i)}$ . Finalement, la projection (3.14) devient

$$\ddot{a}_i = -c^2 \sum_{n=1}^{\infty} \left( \frac{k_n}{c} \right)^2 \bar{\zeta}_n \mu_n^{(i)} - c^2 \sum_{j=1}^q \left( \sum_{n=1}^{\infty} \left( \frac{k_n}{c} \right)^2 \mu_n^{(j)} \mu_n^{(i)} \right) a_j, \quad (3.17)$$

et lorsque la série trigonométrique est tronquée, le modèle d'ordre réduit linéaire s'écrit sous la forme matricielle

$$\ddot{\mathbf{a}} = \mathcal{L}\mathbf{a} + \mathcal{K} \quad (3.18)$$

en regroupant les amplitudes modales dans le vecteur  $\mathbf{a}$ . Les opérateurs  $\mathcal{K} \in \mathbb{R}^q$  et  $\mathcal{L} \in \mathcal{M}(q, q)$  sont alors définis par

$$\mathcal{K} = -c^2 \boldsymbol{\mu}^\top \text{diag}\left((k_n/c)^2\right) \bar{\boldsymbol{\zeta}} \quad \text{et} \quad \mathcal{L} = -c^2 \boldsymbol{\mu}^\top \text{diag}\left((k_n/c)^2\right) \boldsymbol{\mu}. \quad (3.19)$$

Finalement, pour intégrer en temps le modèle d'ordre réduit il faut préciser la condition initiale  $\mathbf{a}_0 = \mathbf{a}(t=0)$ . Celle-ci peut être calculée comme la projection de la condition initiale (3.3), ce qui conduit à  $\mathbf{a}_0 = \boldsymbol{\mu}^\top (\boldsymbol{\zeta}^{(0)} - \bar{\boldsymbol{\zeta}})$  et  $\dot{\mathbf{a}}_0 = \dot{\mathbf{a}}(t=0) = \mathbf{0}$ . L'équation matricielle (3.18) complétée par les conditions initiales  $\mathbf{a}_0$  et  $\dot{\mathbf{a}}_0$  forme donc un problème de Cauchy pour lequel il existe une solution unique.

Cette première formulation du modèle d'ordre réduit repose sur la solution analytique des clichés qui est exprimée comme une série de modes propres. La décomposition orthogonale propre n'est toutefois pas complètement équivalente à une décomposition sur la base des modes propres puisque la décomposition des variables (3.13) est affine. En ce sens, elle ne coïncide pas avec la décomposition (3.4) mais les deux méthodes ne diffèrent en réalité que par les coefficients qui pondèrent les modes. En effet, la relation (3.12) indique bien que les modes POD sont une combinaison linéaire des modes propres **dans ce cas particulier**. Le lien entre les deux types de décompositions est explicité un peu plus en détail dans l'annexe C.2.

### 3.2.2 Formulation discrète du modèle d'ordre réduit

Cette seconde formulation repose directement sur les solutions discrètes du système dynamique caractérisé par un système d'équations différentielles ordinaires issu de la discrétisation des équations continues (3.1). Le principe de la méthode est celui exposé au paragraphe 2.4.1 : la décomposition orthogonale propre est formulée pour les degrés de liberté discrets du système dynamique, de sorte que les modes POD sont obtenus sous une forme discrète qui peut servir directement pour la projection des équations discrètes du système.

#### 3.2.2.1 Génération de clichés à partir de la solution Éléments Finis du problème

Dans un premier temps, on donne les équations du système dynamique discrétisé qui forment le modèle d'ordre élevé à l'aide duquel une base de clichés discrets est générée. Le système continu (3.1) est ici discrétisé au moyen de la méthode des Éléments Finis afin d'obtenir un système matriciel. Les détails des étapes menant à la forme du système discrétisé pourront être trouvés dans [83] entre autres.

D'une manière générale, la discrétisation des équations du système dynamique est associée à une partition du domaine de calcul en un nombre fini  $N$  d'éléments telle que  $\Omega = \cup_{e=1}^N \Omega_e$ . Dans l'exemple traité ici, on considère que le domaine de calcul  $\Omega = [0; L]$  est partitionné en  $N$  éléments qui correspondent ici à des segments identiques de longueur constante  $h = L/N$ . Le principe de la discrétisation consiste alors à décomposer la solution continue  $u$  sur les  $N$  segments au moyen d'une base de fonctions de forme  $\chi_j(x)$  et à rechercher les valeurs des  $N_v = N + 1$  coordonnées généralisées  $u_j = u(x_j)$  aux extrémités  $x_j$  de ces segments. La solution analytique est alors approchée par

$$u(x, t) \approx \sum_{j=1}^{N_v} u_j(t) \chi_j(x). \quad (3.20)$$

À la différence de la décomposition sur la base des modes POD, le nombre  $N_v$  de fonctions de base est en général grand puisque les fonctions choisies pour l'approximation sont universelles : contrairement aux

modos POD, les fonctions considérées ici sont linéaires par morceaux et s'adaptent à n'importe quel type de réponse du système dynamique, alors que les modos POD sont construits pour une réponse particulière.

Les équations discrètes associées à l'équation continue (3.1) sont obtenues en écrivant une formulation faible dans laquelle les déplacements virtuels correspondent aux fonctions de forme  $\chi_j$ . En ramenant ensuite les fonctions de forme à un domaine de référence, on obtient le système d'équations discrètes

$$\mathbf{M} \ddot{\mathbf{u}} + \mathbf{K} \mathbf{u} = \mathbf{0} \quad (3.21)$$

qui porte sur les coordonnées généralisées  $u_j$  regroupées dans le vecteur  $\mathbf{u} = [u_1 \cdots u_{N_v}]^T$ . Les matrices  $\mathbf{M}$  et  $\mathbf{K}$  sont respectivement les matrices de masse et de rigidité qui appartiennent à  $\mathcal{M}(N_v, N_v)$ . Elles sont respectivement définies par

$$\mathbf{M} = \frac{h}{2} \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 2 & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & 2 & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix} \quad \text{et} \quad \mathbf{K} = \frac{c^2}{h} \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ -1 & 2 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 2 & -1 \\ 0 & \cdots & 0 & -1 & 1 \end{bmatrix}. \quad (3.22)$$

Bien que ce ne soit pas l'objet ici, le système (3.21) peut être réduit au moyen d'une décomposition sur la base des modos propres suivant le principe exposé dans le paragraphe 1.3.1 puisque le problème est linéaire. La comparaison entre la projection sur la base des modos propres et des modos POD a déjà fait l'objet de diverses publications et ne sera donc pas effectuée ici. Le lecteur intéressé pourra consulter entre autres les références suivantes [200, 247].

### 3.2.2.2 Calcul des modos POD

Le système (3.21) est résolu sur l'intervalle de temps  $I_e = [t_0; t_0 + T_e]$  afin de générer un ensemble de clichés discrets  $\mathbf{u}^{(m)}$  représentatifs du système. Les clichés correspondent cette fois à des vecteurs de  $\mathbb{R}^{N_v}$  de sorte que l'ensemble des  $M$  clichés est décrit par la matrice des clichés

$$\mathbf{U} = [\mathbf{u}^{(1)} \cdots \mathbf{u}^{(M)}] = \begin{bmatrix} u_1^{(1)} & \cdots & u_1^{(M)} \\ \vdots & & \vdots \\ u_{N_v}^{(1)} & \cdots & u_{N_v}^{(M)} \end{bmatrix} \in \mathcal{M}(N_v, M). \quad (3.23)$$

Chaque composante  $u_i^{(j)}$  de la matrice représente la  $i$ -ème coordonnée généralisée du  $j$ -ième cliché, *i. e.* celle associée à  $u(x_i, t_j)$ . La partie moyenne des clichés  $\bar{\mathbf{u}} = E[\mathbf{u}^{(m)}]$  est un vecteur dans lequel chaque composante est donnée par  $\bar{u}_i = \sum_{m=1}^M u_i^{(m)}$ . La matrice des clichés centrés  $\mathbf{S}$  est alors définie en retranchant la matrice de la partie stationnaire  $\bar{\mathbf{U}} = \bar{\mathbf{u}} \mathbf{I}_{1,M}$  à la matrice des clichés  $\mathbf{U}$  :

$$\mathbf{S} = \mathbf{U} - \bar{\mathbf{U}}. \quad (3.24)$$

Le terme général de la matrice des clichés centrés est donc donné par  $s_{ij} = s_i^{(j)} = u_i^{(j)} - \bar{u}_i$ . Les modos POD sont alors calculés comme les solutions du problème aux valeurs propres (2.72) si l'on adopte la méthode des clichés. Le produit scalaire considéré ici est celui de l'espace euclidien  $\mathbb{R}^{N_v}$  à la différence de la formulation analytique où le produit scalaire était celui associé à l'espace des fonctions de carré intégrable  $L^2([0; L])$ . Le terme général est donc donné par l'expression

$$R_{ij}^* = \sqrt{\alpha_i \alpha_j} \left\langle \mathbf{s}^{(i)}, \mathbf{s}^{(j)} \right\rangle_{\mathbb{R}^{N_v}} = \sqrt{\alpha_i \alpha_j} \sum_{l=1}^N s_l^{(i)} s_l^{(j)}. \quad (3.25)$$

Le produit scalaire de  $\mathbb{R}^{N_v}$  s'interprète immédiatement en terme de produit matriciel entre la matrice des clichés centrés et sa transposée. Les coefficients  $\sqrt{\alpha_i}$  sont introduits en pré- et post-multipliant par la matrice  $\tilde{\alpha}$  pour rendre l'opérateur symétrique. L'opérateur aux valeurs propres de la formulation discrète

est finalement défini par le même produit matriciel que celui donné à l'équation (3.10) pour la formulation analytique. La seule différence réside dans la définition de la matrice des clichés qui contient ici les coordonnées généralisées  $s_j^{(i)}$  résultant de l'approximation Éléments Finis de la solution du système discret.

Les modes POD se décomposent toujours sur l'ensemble des clichés au moyen de la relation (2.73), mais il s'agit cette fois de clichés discrets. Ainsi, la décomposition s'écrit  $\varphi^{(j)} = \sum_{m=1}^M c_m^{(j)} \mathbf{s}^{(m)}$  et par conséquent les modes POD sont décrits au moyen de leurs coordonnées généralisées telles que  $\varphi^{(j)} = [\varphi_1^{(j)}, \dots, \varphi_{N_v}^{(j)}]^\top$ . La décomposition sur les clichés est alors équivalente au produit matriciel

$$\Phi = \mathbf{S}\mathbf{C}. \quad (3.26)$$

La base de projection est donc représentée par la matrice  $\Phi = [\varphi^{(1)}, \dots, \varphi^{(q)}] \in \mathcal{M}(N_v, q)$  qui contient les coordonnées généralisées des modes POD correspondant au formalisme Éléments Finis employé pour discrétiser l'équation continue du système dynamique.

### 3.2.2.3 Construction du modèle d'ordre réduit par une approche discrète

La décomposition orthogonale propre fournit une approximation des coordonnées généralisées  $\mathbf{u}$ . Par conséquent, la décomposition discrète (2.50) adaptée ici pour des clichés centrés devient

$$\mathbf{u}(t) \approx \bar{\mathbf{u}} + \Phi \mathbf{a}(t). \quad (3.27)$$

La construction du modèle d'ordre réduit détaillée au paragraphe 2.4.1 résulte du remplacement de la décomposition (3.27) dans l'équation discrétisée (3.21) puis de la pré-multiplication par  $\Phi^\top$  qui correspond à l'étape de projection. Le modèle d'ordre réduit s'écrit donc

$$\mathbf{M}_q \ddot{\mathbf{a}} + \mathbf{K}_q \mathbf{a} + \mathbf{F}_q = \mathbf{0}, \quad (3.28)$$

où  $\mathbf{M}_q = \Phi^\top \mathbf{M} \Phi$  est la matrice de masse réduite,  $\mathbf{K}_q = \Phi^\top \mathbf{K} \Phi$  est la matrice de raideur réduite et  $\mathbf{F}_q = \Phi^\top \mathbf{K} \bar{\mathbf{u}}$  est le vecteur correspondant à la projection d'un effort statique induit par la partie moyenne  $\bar{\mathbf{u}}$ . La matrice  $\mathbf{M}_q$  est symétrique, définie positive et par conséquent inversible. De ce fait, le modèle d'ordre réduit s'écrit exactement sous la même forme (3.18) que celle obtenue avec la solution analytique. Les opérateurs linéaires et constants sont cependant donnés respectivement par

$$\mathcal{K} = -\mathbf{M}_q^{-1} \mathbf{F}_q \quad \text{et} \quad \mathcal{L} = -\mathbf{M}_q^{-1} \mathbf{K}_q. \quad (3.29)$$

Enfin les conditions initiales traduites en terme d'amplitudes modales se calculent aussi par projection de la relation  $\mathbf{u}(t=0) = \bar{\mathbf{u}} + \Phi \mathbf{a}(t=0)$  sur la base des modes POD, ce qui conduit à :  $\mathbf{a}_0 = \Phi^\top (\mathbf{u}(t=0) - \bar{\mathbf{u}})$ . Les conditions initiales sur la dérivée des amplitudes restent ici encore nulles, en accord avec (3.3) :  $\dot{\mathbf{a}}_0 = \mathbf{0}$ .

### 3.2.3 Formulation POD-Galerkin du modèle d'ordre réduit

La dernière technique employée pour construire le modèle d'ordre réduit découle de la projection de Galerkin décrite au paragraphe 2.4.3. La projection des équations continues s'effectue au moyen du produit scalaire associé à l'espace  $H = L^2([0; L])$  puisqu'on cherche une décomposition des variables continues similaire à (2.74) lorsque les clichés centrés sont considérés. Toutefois, lorsque les clichés sont calculés comme les solutions numériques d'un système discrétisé tel que (3.21), seuls les vecteurs des coordonnées généralisées des clichés  $\mathbf{u}^{(m)}$  sont connus. Par conséquent les produits scalaires qui apparaissent dans les expressions (2.57) ou (2.60) du modèle d'ordre réduit doivent être explicités au moyen des quantités discrètes dont on dispose<sup>2</sup>.

La manière la plus cohérente d'évaluer les produits scalaires consiste à interpréter les composantes discrètes des clichés comme les coordonnées généralisées intervenant dans l'approximation (3.20) dans le cas où les clichés ont été générés par un calcul Éléments Finis. Cependant, on emploie ici une technique d'approximation plus générale qui ne fait pas directement appel au formalisme de discrétisation employé pour générer les clichés<sup>3</sup>.

<sup>2</sup>D'après le principe de construction, l'approche a été qualifiée de méthode « hybride » ou « mixte » dans [180, 181] puisque la projection des équations s'effectue dans un cadre continu mais les quantités discrètes sont ensuite nécessairement manipulées pour construire le modèle d'ordre réduit.

<sup>3</sup>En particulier la méthode peut être employée si les données proviennent d'un dispositif expérimental. Dans ce cas, les composantes des vecteurs des clichés doivent être associées à un maillage de l'espace à définir.

### 3.2.3.1 Approximation numérique du produit scalaire

Le produit scalaire associé à l'espace  $H = L^2([0; L])$  défini à l'équation (2.91) implique une intégrale sur le domaine  $\Omega$  qu'il est nécessaire d'approcher au moyen des coordonnées généralisées. L'intégrale sur tout le domaine est donc scindée en une somme d'intégrales portant sur chaque élément de la discrétisation si bien que le produit scalaire de deux fonctions intégrables  $f$  et  $g$  définies sur le domaine spatial  $\Omega$  s'écrit :

$$\langle f, g \rangle_{L^2(\Omega)} = \int_{\Omega} f(x)g(x) dx = \sum_{e=1}^N \int_{\Omega_e} f(x)g(x) dx \approx \sum_{e=1}^N \widehat{f}_e \widehat{g}_e \delta\Omega_e. \quad (3.30)$$

Chaque intégrale de la somme est alors approchée par le produit des valeurs moyennes  $\widehat{f}_e$  et  $\widehat{g}_e$  des fonctions  $f$  et  $g$  sur le sous-domaine  $\Omega_e$ , le tout étant multiplié par la mesure de Lebesgue<sup>4</sup>  $\delta\Omega_e = \text{mes}(\Omega_e)$  du sous-domaine  $\Omega_e$ . Lorsque la valeur moyenne sur l'élément est calculée comme la moyenne arithmétique des valeurs aux extrémités des éléments, la méthode porte le nom de méthode des trapèzes puisque sur chaque intervalle, l'intégrale est approchée par l'aire du trapèze délimitant le secteur sous la courbe. L'expression (3.30) se ramène alors au produit matriciel

$$\langle f, g \rangle_{L^2(\Omega)} \approx \widehat{\mathbf{f}}^T \mathbf{\Delta} \widehat{\mathbf{g}} \quad (3.31)$$

en introduisant le vecteur des valeurs moyennes  $\widehat{\mathbf{f}} = [\widehat{f}_1, \dots, \widehat{f}_N]^T$  (resp.  $\widehat{\mathbf{g}}$ ) de la fonction  $f$  (resp.  $g$ ) sur chaque élément et la matrice diagonale  $\mathbf{\Delta} = \text{diag}(\delta\Omega_1, \dots, \delta\Omega_N)$  contenant les mesures de Lebesgue de tous les éléments  $\Omega_e$ .

Le calcul des valeurs moyennes dépend de la discrétisation employée pour générer les clichés, et en particulier de la localisation des coordonnées généralisées qui peuvent être calculées sur les *nœuds* des éléments du maillage ou au contraire au *centre* des éléments. Si les valeurs des fonctions  $f$  et  $g$  sont connues au centre des éléments  $\Omega_e$ , alors on peut considérer qu'il s'agit des valeurs moyennes. Dans le cas contraire, les valeurs moyennes sont calculées à partir des valeurs connues sur les extrémités des éléments  $\Omega_e$ . Puisque l'exemple traité est ici unidimensionnel, les valeurs moyennes sont simplement calculées comme la demi-somme  $\widehat{f}_i = (f_i + f_{i+1})/2$  des valeurs aux nœuds correspondant aux extrémités de chaque élément.

L'approche POD-Galerkin, lorsqu'elle est combinée avec l'utilisation de clichés discrets, nécessite donc de définir la façon dont le produit scalaire est approché ainsi que la manière dont les valeurs moyennes sont obtenues.

### 3.2.3.2 Calcul des modes POD

Les clichés  $\mathbf{u}^{(m)}$  disponibles proviennent de l'intégration en temps du système (3.21) et sont donc discrets. Toutefois, puisque la décomposition orthogonale propre est formulée dans l'espace  $H = L^2([0; L])$ , le terme général de l'opérateur de corrélation (2.72) formulé pour les clichés centrés  $s^{(m)} = u^{(m)} - \bar{u}$  doit être discrétisé au moyen de l'approximation (3.31). Ceci conduit à la relation

$$R_{ij}^* = \sqrt{\alpha_i \alpha_j} \langle s^{(i)}, s^{(j)} \rangle_{L^2} \approx \sqrt{\alpha_i \alpha_j} (\widehat{\mathbf{s}}^{(i)})^T \mathbf{\Delta} \widehat{\mathbf{s}}^{(j)}, \quad (3.32)$$

qui est généralisée de sorte que l'opérateur de corrélation s'écrive sous la forme matricielle

$$\mathbf{R}^* = \widetilde{\boldsymbol{\alpha}}^T \widehat{\mathbf{S}}^T \mathbf{\Delta} \widehat{\mathbf{S}} \widetilde{\boldsymbol{\alpha}}. \quad (3.33)$$

La forme de l'opérateur est légèrement différente de celle obtenue avec les formulations analytique et discrète. En effet, la matrice des clichés contient cette fois les valeurs moyennes des clichés au centre des éléments de discrétisation du domaine spatial, et la matrice des mesures de Lebesgue vient pondérer l'opérateur de corrélation. Il est important de mentionner que la symétrie de l'opérateur est préservée puisque la matrice  $\mathbf{\Delta}$  est diagonale.

<sup>4</sup>La mesure de Lebesgue dépend de la dimension  $d$  du domaine spatial  $\Omega \subset \mathbb{R}^d$  dans lequel évolue le système. Cette mesure correspond à une longueur lorsque  $d = 1$ , à une surface lorsque  $d = 2$  et enfin à un volume lorsque  $d = 3$ .



### 3.2.3.3 Construction du modèle d'ordre réduit par une approche POD-Galerkin

Le modèle d'ordre réduit est construit de la même manière que pour la formulation analytique puisque dans les deux cas la projection est effectuée à l'aide du produit scalaire associé à  $H = L^2([0; L])$ . L'introduction de la décomposition (3.13) dans l'équation (3.1) et la projection de Galerkin sur le mode POD  $\varphi^{(i)}$  avec le produit scalaire associé à l'espace  $H = L^2([0; L])$  conduit à l'équation (3.14). Le développement des produits scalaires donne alors

$$\ddot{a}_i = c^2 \left\langle \frac{d^2 \bar{u}}{dx^2}, \varphi^{(i)} \right\rangle + c^2 \sum_{j=1}^q \left\langle \frac{d^2 \varphi^{(j)}}{dx^2}, \varphi^{(i)} \right\rangle a_j. \quad (3.34)$$

Les produits scalaires sont ensuite approchés au moyen de la relation (3.31) fournie au début de ce paragraphe. Il suffit pour cela de calculer les dérivées secondes<sup>5</sup> des modes POD et de la partie stationnaire puis d'évaluer les valeurs moyennes sur les éléments de discrétisation pour former les vecteurs  $\widehat{d^2 \Phi} = [d^2 \widehat{\varphi}^{(1)}/dx^2, \dots, d^2 \widehat{\varphi}^{(q)}/dx^2]$  et  $\widehat{d^2 \bar{u}} = d^2 \bar{u}/dx^2$  respectivement. La matrice des valeurs moyennes des modes POD  $\widehat{\Phi} = [\widehat{\varphi}^{(1)}, \dots, \widehat{\varphi}^{(q)}]$  est aussi formée afin de pouvoir évaluer les opérateurs du modèle d'ordre réduit. Celui-ci s'écrit donc finalement sous la forme matricielle  $\ddot{\mathbf{a}} = \widetilde{\mathcal{K}} + \widetilde{\mathcal{L}} \mathbf{a}$ , où les opérateurs sont définis par

$$\widetilde{\mathcal{K}} = -c^2 \widehat{\Phi} \Delta \widehat{d^2 \bar{u}} \quad \text{et} \quad \widetilde{\mathcal{L}} = -c^2 \widehat{\Phi} \Delta \left( \widehat{d^2 \Phi} \right)^T. \quad (3.35)$$

L'inconvénient de cette formulation est de devoir évaluer les dérivées secondes de la partie stationnaire et des modes POD, ce qui nécessite d'employer des schémas de dérivation d'ordre élevé. Le problème peut être aisément contourné en tirant profit de la définition du produit scalaire sur l'espace  $L^2([0; L])$ . En effet, les produits scalaires de l'expression (3.34) se transforment par le biais d'une intégration par partie en

$$\begin{aligned} \left\langle \frac{d^2 \bar{u}}{dx^2}, \varphi^{(j)} \right\rangle_{L^2} &= \int_0^L \frac{d^2 \bar{u}}{dx^2} \varphi^{(j)} dx = \underbrace{\left[ \frac{d \bar{u}}{dx} \varphi^{(j)} \right]_0^L}_{=0} - \int_0^L \frac{d \bar{u}}{dx} \frac{d \varphi^{(j)}}{dx} dx \\ \left\langle \frac{d^2 \varphi^{(j)}}{dx^2}, \varphi^{(i)} \right\rangle_{L^2} &= \int_0^L \frac{d^2 \varphi^{(j)}}{dx^2} \varphi^{(i)} dx = \underbrace{\left[ \frac{d \varphi^{(j)}}{dx} \varphi^{(i)} \right]_0^L}_{=0} - \int_0^L \frac{d \varphi^{(j)}}{dx} \frac{d \varphi^{(i)}}{dx} dx. \end{aligned} \quad (3.36)$$

Les termes entre crochets s'annulent puisque les modes POD héritent des conditions aux limites homogènes (3.2) vérifiées par les clichés (voir §2.2.3.4) et par conséquent  $\varphi^{(j)}(x=0) = 0$  et  $d\varphi^{(i)}(x)/dx|_{x=L} = 0$ . De même la partie moyenne se simplifie puisque  $\bar{u}$  est par définition une combinaison linéaire des clichés si bien que les conditions aux limites homogènes sont aussi satisfaites par  $\bar{u}$ .

Les intégrales subsistant dans les expressions (3.36) correspondent encore à des produits scalaires de l'espace  $L^2([0; L])$ . Après intégrations par parties, le modèle d'ordre réduit s'écrit donc

$$\ddot{a}_i = -c^2 \left\langle \frac{d \bar{u}}{dx}, \frac{d \varphi^{(i)}}{dx} \right\rangle - c^2 \sum_{j=1}^q \left\langle \frac{d \varphi^{(j)}}{dx}, \frac{d \varphi^{(i)}}{dx} \right\rangle a_j. \quad (3.37)$$

Les produits scalaires sont approchés par la relation (3.31) et on introduit à cet effet la matrice des valeurs moyennes des dérivées premières des modes POD  $\widehat{d \Phi} = [d \widehat{\varphi}^{(1)}/dx, \dots, d \widehat{\varphi}^{(q)}/dx]$  et le vecteur des valeurs moyennes de la dérivée première de la partie moyenne  $\widehat{d \bar{u}} = d \bar{u}/dx$ . Le système d'équations (3.37) s'écrit donc sous la même forme matricielle (3.18) que celle des formulations précédentes, mais les opérateurs sont maintenant définis par

$$\mathcal{K} = -c^2 \left( \widehat{d \Phi} \right)^T \Delta \widehat{d \bar{u}} \quad \text{et} \quad \mathcal{L} = -c^2 \left( \widehat{d \Phi} \right)^T \Delta \left( \widehat{d \Phi} \right). \quad (3.38)$$

Enfin la condition initiale pour les amplitudes modales s'obtient par projection de la condition initiale (3.3) :  $\mathbf{a}_0 = \widehat{\Phi}^T \Delta (\widehat{\mathbf{u}}(t=0) - \widehat{\bar{\mathbf{u}}})$ . La condition initiale cinématique  $\dot{\mathbf{a}}_0$  reste quant à elle toujours nulle.

<sup>5</sup>Les dérivées sont calculées ici au moyen d'un schéma aux différences centrées.

**Remarque 3.1 : Intégration en temps du modèle d'ordre réduit**

D'après la remarque effectuée dans la note 1 p.9, les systèmes dynamiques réduits d'ordre deux en temps développés dans ce paragraphe peuvent être ramenés à des systèmes équivalents dans lesquels seules des dérivées premières en temps interviennent. Ainsi, le modèle d'ordre réduit (3.18) est transformé en posant  $\mathbf{y} = [\mathbf{a} \ \dot{\mathbf{a}}]^T$ . Le système se réorganise alors en

$$\underbrace{\begin{bmatrix} \mathbf{0}_{q,q} & \mathbf{I}_q \\ \mathbf{I}_q & \mathbf{0}_{q,q} \end{bmatrix}}_{\mathbf{V}} \begin{bmatrix} \dot{\mathbf{a}} \\ \ddot{\mathbf{a}} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathcal{L} & \mathbf{0}_{q,q} \\ \mathbf{0}_{q,q} & \mathbf{I}_q \end{bmatrix}}_{\mathbf{W}} \begin{bmatrix} \mathbf{a} \\ \dot{\mathbf{a}} \end{bmatrix} + \underbrace{\begin{bmatrix} \mathcal{K} \\ \mathbf{0}_{q,1} \end{bmatrix}}_{\mathbf{g}}. \quad (3.39)$$

La matrice  $\mathbf{V}$  étant inversible, le système différentiel d'ordre 2 en temps portant initialement sur le vecteur d'inconnue  $\mathbf{a} \in \mathbb{R}^q$  se ramène à un système différentiel d'ordre 1 en temps portant sur le vecteur  $\mathbf{y} \in \mathbb{R}^{2q}$  :

$$\dot{\mathbf{y}} = \mathbf{V}^{-1} \mathbf{W} \mathbf{y} + \mathbf{V}^{-1} \mathbf{g}. \quad (3.40)$$

Le système (3.40) complété par les conditions initiales  $\mathbf{y}_0 = [\mathbf{a}_0 \ \dot{\mathbf{a}}_0]^T$  définit un problème de Cauchy bien posé puisque le modèle d'ordre réduit nécessite seulement la donnée de conditions initiales. En effet, les conditions aux limites vérifiées par le modèle d'ordre élevé sont implicitement prises en compte dans la manière de calculer les modes POD et interviennent par conséquent au travers des valeurs des coefficients du modèle d'ordre réduit<sup>6</sup>. L'intégration en temps du modèle d'ordre réduit peut alors être menée au moyen des méthodes classiques de résolution proposées dans diverses librairies informatiques. L'intégration en temps est ici réalisée au moyen de la fonction `ode` de Scilab qui n'est en réalité qu'une interface vers les librairies ODEPACK [111]. Le solveur `lsoda` du package ODEPACK est utilisé par défaut. Il choisit automatiquement entre un schéma d'Adams de type prédicteur-correcteur et un schéma de type « Backward Differentiation Formula » (méthode de Gear) adapté aux systèmes raides. La résolution débute avec le schéma d'Adams puis, en fonction de la raideur du problème, le schéma est ensuite choisi dynamiquement au cours de la résolution .

---

<sup>6</sup>En dehors des conditions aux limites homogènes, les modes POD ne satisfont pas en général les conditions aux limites mais une combinaison linéaire des conditions aux limites aux instants correspondant aux clichés (voir §2.2.3.4).

### 3.3 Évaluation du modèle d'ordre réduit POD-Galerkin du système linéaire autonome

Dans cette section, la formulation POD-Galerkin est comparée aux formulations analytique et discrète développées dans le paragraphe précédent dans le cas d'une réponse autonome du système à une condition initiale. Dans un premier temps, les erreurs de reconstruction produites par les différentes formulations sont comparées, puis on s'intéresse dans un second temps au choix de l'ensemble des clichés servant ensuite au calcul des modes POD : l'influence de la distribution des clichés, leur quantité ainsi que le nombre de modes POD retenus dans la base de projection seront en particulier étudiés.

#### 3.3.1 Évaluation du modèle d'ordre réduit POD-Galerkin

La qualité des réponses produites par les différentes formulations du modèle d'ordre réduit est évaluée au moyen de plusieurs types d'erreurs. La première d'entre elles est l'erreur relative spatiale de reconstruction, définie au sens de la norme  $L^2(\Omega)$  pour des fonctions de carré intégrable par

$$\varepsilon_{L^2,u} = 100 \frac{\|u(x,t) - u_r(x,t)\|_{L^2}}{\|u_r(x,t)\|_{L^2}} \quad \text{avec} \quad \forall f \in L^2([0;L]) \quad \|f(x)\|_{L^2} = \sqrt{\int_0^L f(x)^2 dx}. \quad (3.41)$$

L'erreur précédente mesure à tout instant  $t$  l'écart relatif entre la solution  $u(x,t)$  reconstruite par le modèle d'ordre réduit et la solution de référence  $u_r(x,t)$  qui correspond aux clichés aux instants  $t = t_m$ . En pratique, les solutions ne sont connues que sous leur forme discrète et par conséquent l'erreur  $\varepsilon_{L^2,u}$  est évaluée en approchant les intégrales par la règle du trapèze donnée à l'équation (3.30). La solution analytique (3.4) sera en général employée comme solution de référence afin de disposer d'une solution à tout instant et en tout point du domaine.

En complément de cette erreur spatiale, on introduit l'erreur moyenne en temps qui est uniquement évaluée à l'extrémité  $x = L$  de la barre puisque le déplacement  $y$  est maximal. Elle est donnée par

$$\varepsilon_{\text{rms}}(x = L) = \sqrt{\frac{\sum_{j=1}^T |u(x = L, t_j) - u_r(x = L, t_j)|^2}{T}}. \quad (3.42)$$

Enfin, la différence entre la matrice de la solution de référence  $\mathbf{U}_r \in \mathcal{M}(N_v, T)$  et la matrice de même dimension  $\mathbf{U}$  des solutions produites par le modèle d'ordre réduit en  $T$  instants définit l'erreur spatio-temporelle globale

$$\varepsilon_{\infty} = \frac{\|\mathbf{U} - \mathbf{U}_r\|_{\infty}}{\|\mathbf{U}_r\|_{\infty}}. \quad (3.43)$$

L'utilisation de la norme infinie permet d'évaluer l'erreur maximale commise parmi les  $N_v$  coordonnées généralisées et les  $T$  instants d'intégration en temps.

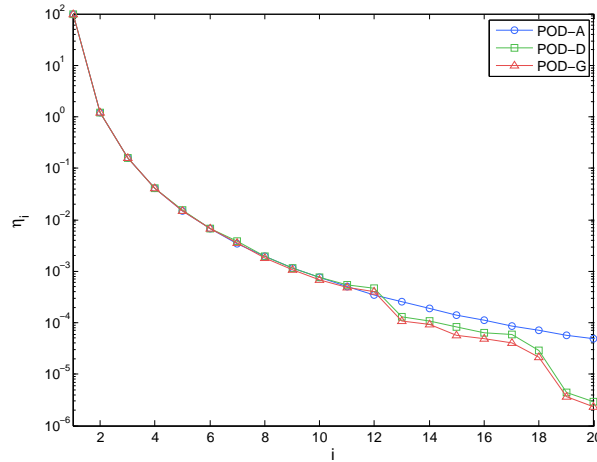
Les différentes formulations du modèle d'ordre réduit décrites aux paragraphes §3.2.1, §3.2.2 et §3.2.3 sont désignées respectivement par la suite au moyen des abréviations suivantes : POD-A pour la formulation analytique, POD-D pour la formulation discrète et POD-G pour la formulation POD-Galerkin.

Les valeurs propres associées aux modes POD de chaque formulation sont comparées sur le graphique 3.2 au moyen du pourcentage d'énergie  $\tilde{\eta}_i$  capté par chaque valeur propre. Par analogie avec le pourcentage  $\eta_q$  d'énergie captée par les  $q$  premières valeurs propres dont la définition a été donnée à l'équation (2.39), le pourcentage  $\tilde{\eta}_i$  est

$$\tilde{\eta}_i = \lambda_i / \sum_{j=1}^r \lambda_j. \quad (3.44)$$

L'allure de la distribution des valeurs propres obtenues avec la formulation analytique reste beaucoup plus régulière qu'avec les deux autres formulations lorsque l'indice  $i$  augmente. Ceci est dû au fait que l'ensemble des clichés utilisé avec les formulations discrète et de Galerkin provient de la résolution des équations discrètes (3.21) alors que la formulation analytique repose sur l'expression exacte des clichés (3.4)

qui n'est pas entachée d'erreurs d'approximation. Les spectres obtenus avec les différentes formulations sont pratiquement confondus jusqu'à la douzième valeur propre au-delà de laquelle une discontinuité dans le spectre est observée lorsque les clichés proviennent de la simulation numérique.



**Figure 3.2 :** Spectre du pourcentage d'énergie capté par les premières valeurs propres associées aux modes POD calculés au moyen des formulations POD analytique, discrète et Galerkin.

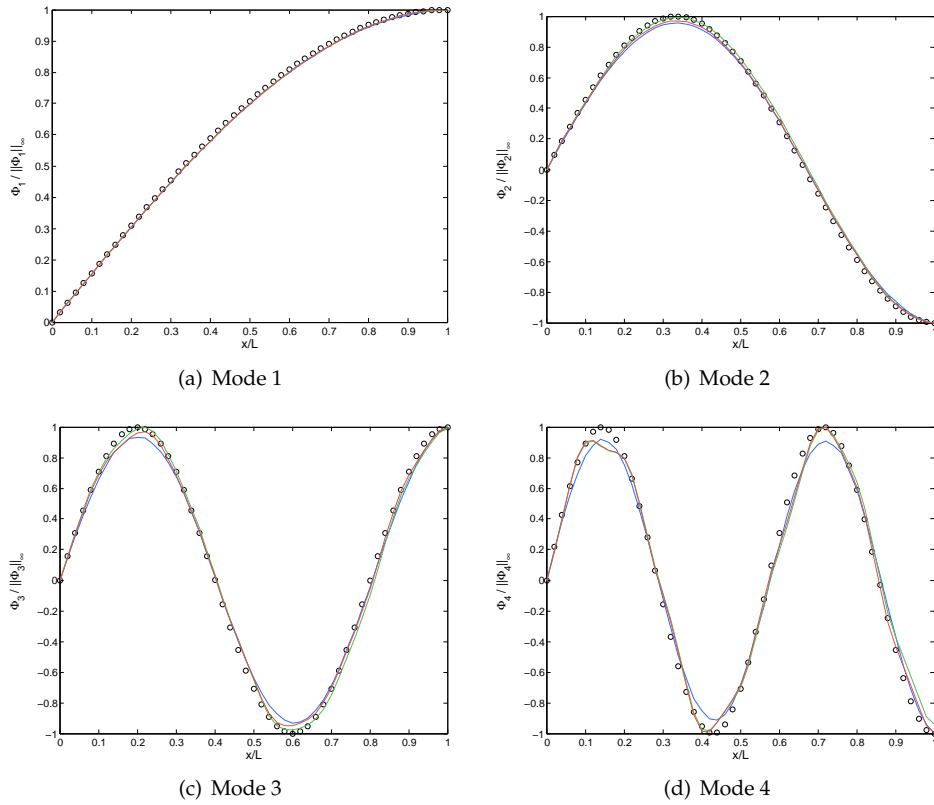
Les valeurs propres évoluent de manière très similaire avec les formulations discrète et de Galerkin. La distinction se traduit uniquement par une translation verticale du spectre qui s'explique par le fait que les opérateurs de corrélation ne diffèrent en réalité que par la métrique qui est employée pour calculer les produits scalaires (voir les expressions (3.25) et (3.32) du terme général des opérateurs de corrélation des deux formulations).

Le spectre des valeurs propres décroît rapidement : la première valeur propre capte  $\tilde{\eta}_1 = 98,6\%$  de l'énergie du système et le pourcentage d'énergie capté par les cinq premières valeurs propres est  $\eta_5 = 99,98\%$ , ce qui signifie que le modèle d'ordre réduit doit pouvoir être construit avec un nombre très restreint de valeurs propres.

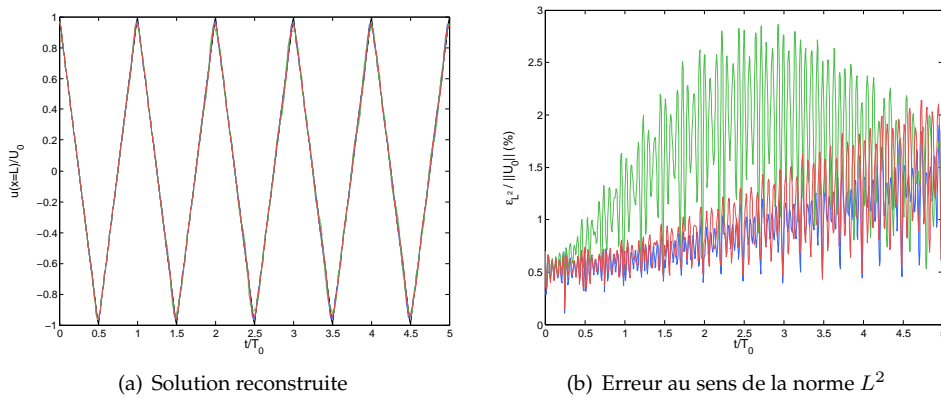
La figure 3.3 donne l'allure des quatre premiers modes POD de la barre calculés au moyen des différentes formulations. Quelle que soit la méthode, la ressemblance entre les modes POD et les modes propres  $\psi_n$  de la barre est flagrante. L'écart entre les modes POD issus de la formulation analytique et ceux produits par les autres formulations se creuse au fur et à mesure pour les modes d'ordre élevé.

Les réponses produites par les modèles d'ordre réduit construits avec les différentes formulations sont tracées sur le graphique de la figure 3.4(a). Globalement la réponse est excellente lorsque le modèle d'ordre réduit est intégré en temps sur une durée d'environ cinq périodes d'oscillation. Toutefois l'étude des erreurs spatiales  $\varepsilon_{L^2,u}$  tracées sur la figure 3.4(b) révèle des comportements sensiblement différents. Ainsi, les réponses reconstruites avec les formulations analytique et de Galerkin suivent une évolution assez similaire produisant une erreur  $\varepsilon_{L^2,u}$  caractérisée par une croissance globalement monotone dans le temps. Au contraire, la formulation discrète produit une erreur qui oscille globalement sur la période de temps considérée. En simulant sur des durées plus longues, on pourrait vérifier que les formulations analytique et de Galerkin produisent invariablement une erreur instantanée globalement croissante, alors que l'erreur associée à la formulation discrète augmente par cycles, la forme étant grossièrement sinusoidale et croissante.

Enfin le tableau 3.1 fournit les valeurs des erreurs  $\varepsilon_{\text{rms}}$  et  $\varepsilon_{\infty}$  pour chaque formulation lorsque la réponse est simulée sur cinq ou dix périodes. La formulation analytique se révèle bien être la plus performante pour reconstituer la réponse du système puisque sur les deux durées simulées, les erreurs rms et infinies sont minimales. La formulation de Galerkin produit des erreurs légèrement plus importantes que la formulation analytique comme le laissait entrevoir le graphique de l'erreur instantanée 3.4(b). Enfin, la formulation discrète apparaît comme la plus « mauvaise » puisque non seulement l'erreur instantanée mais aussi les erreurs rms et infinies montrent qu'en moyenne la réponse reconstruite est de moins bonne qualité.



**Figure 3.3 :** Comparaison des quatre premiers modes POD de la barre calculés avec la formulation analytique (—), la formulation discrète (—) et la formulation de Galerkin (—). L’allure des modes propres  $\{\psi_n\}_{n=1}^4$  définis à l’équation (3.4) est donnée à titre de comparaison (○).



**Figure 3.4 :** Comparaison des solutions produites par le modèle d’ordre réduit construit avec  $q = 5$  modes POD avec la formulation analytique (—), la formulation discrète (—) et la formulation de Galerkin (—). La courbe (—) sur la figure (a) correspond à la solution de référence calculée avec  $p = 50$  modes propres. Le graphique de droite représente l’erreur au sens de la norme  $L^2$  entre la solution reconstruite avec chaque formulation et la solution de référence.

	5 périodes			10 périodes		
	POD-A	POD-D	POD-G	POD-A	POD-D	POD-G
$\varepsilon_{\text{rms}}$	1,59	2,74	1,71	2,38	3,22	2,55
$\varepsilon_{\infty}$	1,89	3,74	2,39	2,76	4,44	3,58

**Tableau 3.1 :** Comparaison des erreurs produites par les réponses calculées avec les formulations analytique, discrète et de Galerkin du modèle d'ordre réduit. La réponse est évaluée sur cinq et dix périodes d'oscillation du système.

La comparaison des trois formulations du modèle d'ordre réduit présentées au paragraphe 3.2 a permis d'évaluer la qualité de la réponse produite par chacune des formulations. La formulation analytique est sans surprise la plus performante puisque les modes POD sont construits à partir de clichés définis exactement comme les solutions du système, ce qui assure une bonne régularité du spectre. La réponse produite par le modèle d'ordre réduit construit avec cette formulation n'engendre que de très faibles erreurs. Toutefois, cette formulation n'a que peu d'intérêt en dehors de celui de servir de point de comparaison puisqu'en pratique les solutions analytiques ne seront jamais disponibles.

Ainsi, il est intéressant de comparer les performances des formulations discrète et de Galerkin par rapport à l'approche analytique qui constitue le modèle d'ordre réduit de référence. Les modes POD calculés avec ces deux formulations sont pratiquement confondus puisque les clichés ayant servis à les générer sont identiques. La seule différence entre les deux méthodes de calcul provient de la métrique employée pour calculer l'opérateur de corrélation, dont l'influence semble ici peu prononcée. Cependant, la réponse du modèle d'ordre réduit basé sur la formulation discrète produit des erreurs généralement plus importantes que celles correspondant à la formulation de Galerkin. Cette dernière apparaît donc comme un bon compromis entre les approches analytique et discrète pour l'exemple traité ici. En effet, elle offre la possibilité d'employer les clichés discrets issus des simulations numériques tout en étant basée sur une formulation continue du problème qui semble diminuer au final l'erreur de reconstruction.

### 3.3.2 Choix de la base des clichés et du nombre de modes POD

Dans ce paragraphe, l'influence de quelques paramètres entrant dans la construction du modèle d'ordre réduit à partir de la base des modes POD est étudiée. On s'intéresse en particulier au nombre  $M$  de clichés employé pour construire l'opérateur de corrélation, ainsi qu'au nombre  $q$  de modes POD conservés dans la base de projection. Un troisième paramètre important noté  $n_p$  correspond au nombre de périodes d'oscillation du système sur lequel les clichés sont prélevés. L'influence de ce paramètre est analysée dans l'annexe C.1 pour ne pas surcharger la présentation. Les résultats obtenus indiquent, entre autres, que l'échantillonnage doit être pratiqué de préférence avec une valeur  $n_p$  telle que la durée d'échantillonnage ne soit pas un multiple de période fondamentale du système lorsque l'échantillonnage est uniforme.

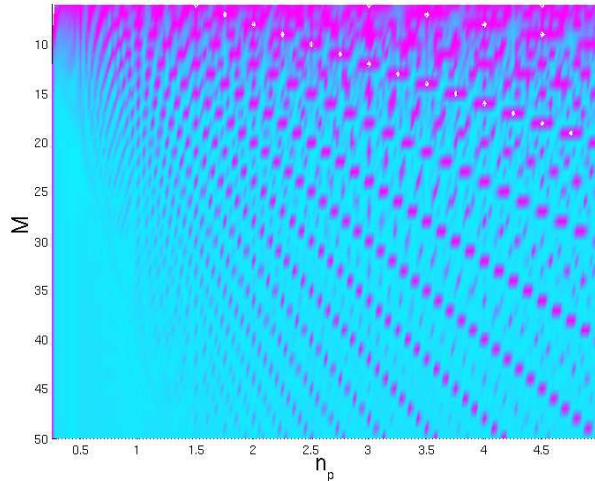
Les résultats présentés ont pour la plupart été obtenus avec le modèle d'ordre réduit construit au moyen de la formulation POD-Galerkin puisqu'il s'agit de la méthode qui sera conservée pour la suite. Néanmoins, une comparaison des différentes formulations est fournie en conclusion de ce paragraphe.

Dans un premier temps, on étudie la convergence de l'erreur de reconstruction lorsque le nombre  $M$  de clichés employés pour former l'opérateur de corrélation augmente. Le nombre de modes POD conservés dans la base de projection est constant et égal à  $q = 5$  puisque d'après les résultats du paragraphe précédent le pourcentage d'énergie captée est déjà  $\eta_q = 99,98\%$ . Les modes POD sont alors calculés pour différentes valeurs de  $M$  variant dans l'intervalle  $[[6; 50]]$ . La borne inférieure est choisie de sorte que la taille du problème aux valeurs propres à résoudre soit supérieure au nombre  $q$  de modes POD choisi pour le modèle d'ordre réduit<sup>7</sup>. La borne supérieure est quant à elle déterminée par le nombre  $N_v$  de degrés de liberté du modèle Éléments Finis correspondant au modèle d'ordre élevé.

La convergence est aussi analysée en fonction du nombre  $n_p$  de périodes sur lequel les clichés sont prélevés. Le résultat tracé sur le graphique de la figure 3.5 laisse apparaître une certaine convergence lorsque le nombre  $M$  de clichés augmente pour  $n_p$  fixé. Toutefois des pics d'erreurs localisés pour certaines combinaisons critiques de  $M$  et  $n_p$  sont observés. Des détails quant à l'origine des pics d'erreurs sont fournis dans

<sup>7</sup>Si les clichés ne sont pas linéairement indépendants, le rang de l'opérateur de corrélation chute et le nombre de modes POD dans la base de projection correspond alors au minimum entre le nombre  $q$  de mode POD souhaité et le rang  $r$  de la matrice de corrélation.

l'annexe C.1. On retrouve en particulier le fait que les principaux pics d'erreur se concentrent le long des lignes où le rapport  $M/n_p$  est entier. Les erreurs les plus faibles se rencontrent dans la zone correspondant aux valeurs élevées de  $M$  et aux valeurs de  $n_p$  comprises approximativement dans l'intervalle  $[0, 5; 1, 5]$ . D'une manière générale, la réponse est de bonne qualité dès qu'une vingtaine de clichés est utilisée, sauf dans le cas où la valeur de  $M$  choisie conduit à une redondance des clichés. Ainsi, lorsque le nombre de clichés augmente pour  $n_p \approx 1,0$ , l'erreur de reconstruction converge assez rapidement vers une valeur asymptotique qui atteint  $\varepsilon_\infty = 2,39\%$ .



**Figure 3.5 :** Convergence de l'erreur infinie de reconstruction  $\varepsilon_\infty$  du modèle d'ordre réduit en fonction du nombre  $M$  de clichés et du nombre  $n_p$  de périodes d'échantillonnage. Les tons bleus correspondent à une erreur faible tandis que les tons roses indiquent les erreurs maximales.

La convergence s'observe aussi au niveau de l'allure des modes POD qui est représentée sur les graphiques de la figure 3.6 pour différentes valeurs de  $M$  : plus le nombre de clichés utilisés pour former l'opérateur de corrélation est grand (de  $M = 6$  en bleu clair à  $M = 50$  en rose), plus la forme des modes POD se rapproche de celle des modes propres  $\psi_n$ . Le premier mode POD coïncide parfaitement avec le premier mode propre du système, et ce, même pour de très faibles valeurs du nombre de clichés. Cependant, le nombre de clichés nécessaire pour capter correctement les modes d'ordre plus élevé augmente rapidement. Ainsi avec  $M < 20$  clichés, le troisième et le quatrième mode sont fortement écrêtés.

Ce phénomène de convergence des modes POD vers les modes propres correspond bien à la situation décrite dans [74, 129] à condition que le système étudié soit linéaire, autonome et que la matrice de masse du système soit proportionnelle à l'identité. Ces conditions sont bien réunies ici<sup>8</sup> et la convergence vers les modes propres est effectivement constatée. Quelques éléments d'explication de la convergence des modes POD vers les modes propres pour le système considéré ici sont fournis dans l'annexe C.2.

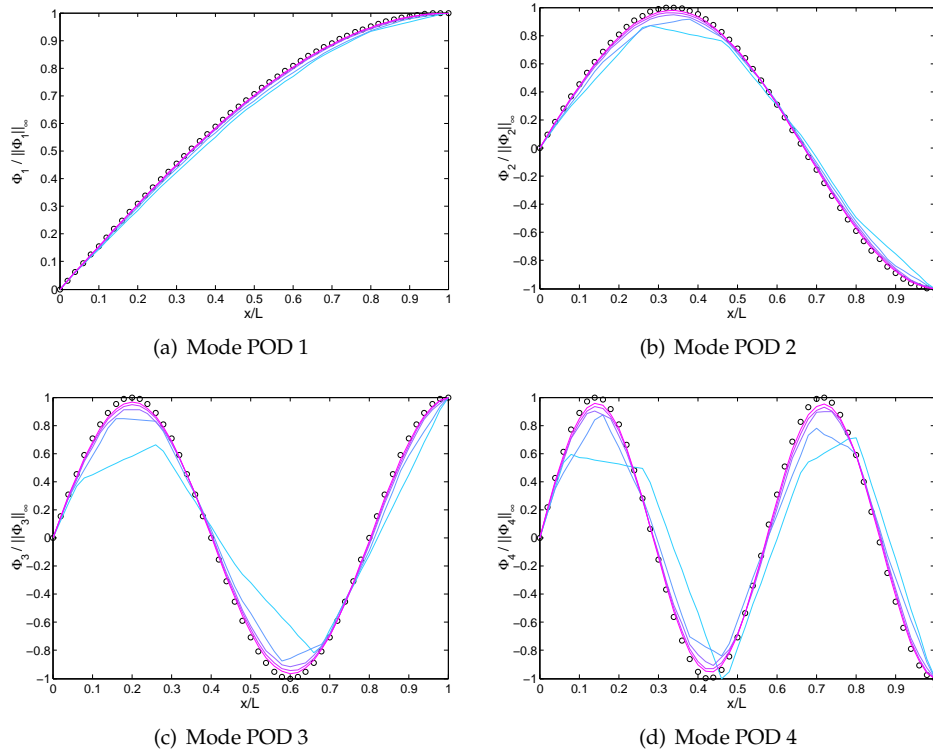
Le dernier paramètre fondamental dans la construction du modèle d'ordre réduit est le nombre  $q$  modes POD conservés dans la base de projection. La réponse du modèle d'ordre réduit est alors évaluée lorsque le nombre  $q$  de modes est compris dans l'intervalle  $[[2; 15]]$ . Il est inutile de considérer plus de modes POD puisque le temps de calcul augmente rapidement sans que la réponse ne soit améliorée de manière significative. La valeur de  $n_p$  est choisie en accord avec les résultats précédents où il a été mis en évidence que pour  $n_p \approx 1$ , la convergence était optimale. L'allure de la distribution de l'erreur infinie  $\varepsilon_\infty$  obtenue avec les paramètres précédents est alors représentée sur le graphique de la figure 3.7 en fonction du nombre de modes POD ainsi que du nombre de clichés.

Le nombre de modes POD a été volontairement tronqué au-dessus d'une certaine limite matérialisée par la ligne noire puisque celui-ci est limité par le rang de la matrice de corrélation qui est lui-même assujéti au nombre de clichés ayant servi pour construire l'opérateur de corrélation par la méthode des clichés :  $q \leq r \leq M$ . Le nombre maximal de modes POD disponible<sup>9</sup> est alors  $q = M - 1$  et lorsque le modèle d'ordre

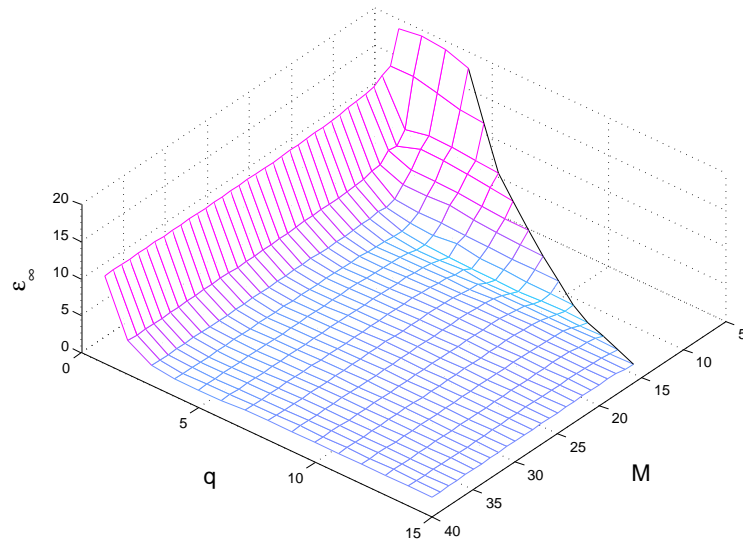
<sup>8</sup>En éliminant dans la matrice de masse (3.22) les termes correspondants aux conditions aux limites.

<sup>9</sup>Le rang chute d'une unité si l'opérateur de corrélation est formulé pour les clichés centrés où la partie moyenne a été soustraite.

### 3.3. Évaluation du modèle d'ordre réduit POD-Galerkin du système linéaire autonome



**Figure 3.6 :** Convergence des quatre premiers modes POD de la barre calculés pour  $n_p \approx 1$  lorsque le nombre de clichés augmente. — :  $M = 6$ , — :  $M = 10$ , — :  $M = 20$ , — :  $M = 30$ , — :  $M = 50$ . L'allure des modes propres de la barre est donnée à titre de comparaison par les cercles ( $\circ$ ).



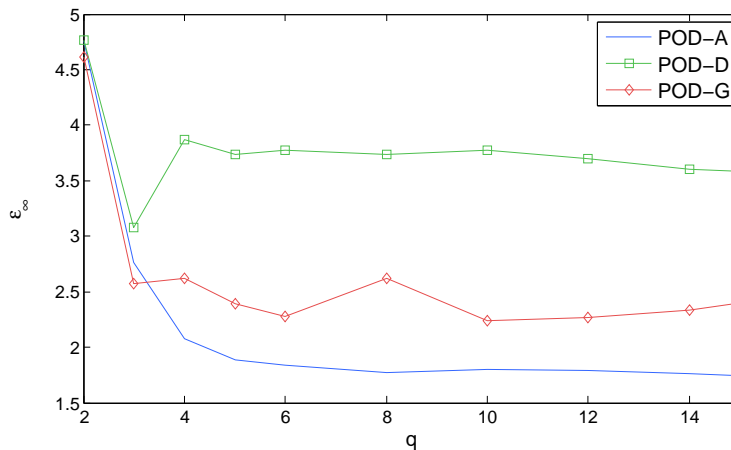
**Figure 3.7 :** Convergence de l'erreur infinie de reconstruction  $\varepsilon_\infty$  du modèle d'ordre réduit en fonction du nombre  $M$  de clichés et du nombre  $q$  de modes POD conservés dans la base de projection POD.



réduit doit être construit pour  $M < q$ , le nombre de modes dans la base de projection n'est en réalité que de  $M - 1$ .

Le graphique de la figure 3.7 met en évidence la convergence de l'erreur vers une valeur asymptotique lorsque  $M$  et  $q$  augmentent. En considérant un nombre  $q$  de modes POD fixé, la courbe de l'erreur présente un léger minimum local qui se situe vers  $M = 13$ . De la même façon, pour un nombre  $M$  de clichés donné, l'erreur décroît fortement, passe par un minimum local situé vers  $q = 4$  — pour  $M \geq 15$ , sinon la position du minimum est plus variable — puis se stabilise vers une valeur asymptotique. Le minimum global atteint pour  $(q, M) \in \llbracket 1; 15 \rrbracket \times \llbracket 2; 50 \rrbracket$  est situé pour  $M = 13$  et  $q = 9$ , l'erreur atteignant alors  $\varepsilon_\infty = 1,23\%$ .

Le tracé de l'erreur en fonction du nombre de modes POD a permis de vérifier que le modèle d'ordre réduit converge globalement vers une erreur asymptotique résiduelle qui ne décroît plus au delà d'un certain d'un nombre de modes POD conservés dans la base de projection. Cependant, la convergence n'est pas strictement monotone et l'existence de minima locaux a pu être mise en évidence. Par ailleurs, l'exemple traité ici montre que pour un  $M$  donné le minimum de l'erreur est atteint rapidement pour de petites valeurs de  $q$  — typiquement  $q = 4$  pour  $M \geq 15$  — et qu'il n'est donc pas nécessaire de conserver beaucoup de modes dans la base de projection. Ces résultats corroborent les remarques de Rowley [192] qui affirmait qu'augmenter le nombre de modes POD au-delà d'une certaine limite n'améliore pas nécessairement le résultat. Park et Kim [175] ont par ailleurs expliqué ce phénomène par les erreurs d'arrondis commises sur les modes POD associés à des valeurs propres de plus en plus faibles.



**Figure 3.8 :** Comparaison de la convergence de l'erreur infinie de reconstruction  $\varepsilon_\infty$  des différentes formulations du modèle d'ordre réduit en fonction du nombre  $q$  de modes POD conservés dans la base de projection.

Pour terminer, on compare sur la figure 3.8 la convergence de l'erreur infinie de reconstruction produite avec les différentes formulations du modèle d'ordre réduit. Seul le nombre de modes POD  $q$  retenus dans le modèle d'ordre réduit est modifié ; le nombre de clichés utilisés pour former les modes POD est constant et égal à  $M = 25$ . On vérifie alors que la formulation analytique présente bien le meilleur comportement en terme de convergence puisque l'erreur de reconstruction affiche une décroissance quasiment monotone et converge rapidement vers une valeur asymptotique proche de  $\varepsilon_\infty = 1,7\%$ . Les autres formulations convergent de façon plus irrégulière. Ainsi, l'erreur de la formulation discrète se stabilise sur l'intervalle  $q = \llbracket 4; 10 \rrbracket$  à une valeur assez élevée  $\varepsilon_\infty \approx 3,7\%$  sans diminuer significativement lorsque  $q$  augmente. Enfin, l'erreur produite par la formulation POD-Galerkin évolue aussi de manière irrégulière puisque quelques pics localisés apparaissent — notamment pour  $q = 8$ . Dans l'ensemble, l'erreur de reconstruction varie autour de  $\varepsilon_\infty = 2,5\%$  ce qui en fait un bon compromis entre les formulations analytique et discrète malgré les irrégularités observées.

Dans ce paragraphe, les formulations analytiques, discrète et de Galerkin ont été dans un premier temps comparées entre elles. L'objectif était de vérifier que les erreurs produites par les formulations discrète et de Galerkin présentaient des niveaux équivalents à celui de la formulation analytique. Les résultats obtenus ont par ailleurs montré que la formulation de Galerkin était plus performante que la formulation discrète, ce qui peut s'expliquer par le fait que la formulation est basée sur les équations continues du système. Dans un second temps, l'influence de trois paramètres de construction du modèle d'ordre réduit a été étudiée. Le nombre  $n_p$  de période semble devoir être choisi de façon à éviter les redondances dans les clichés qui apparaissent lorsque la réponse du système est périodique. Par ailleurs, le nombre de modes POD et de clichés ne doit pas nécessairement être élevé. Le choix du nombre de cliché est dicté en partie par le pourcentage d'énergie capté  $\eta_q$  mais une analyse de la convergence de l'erreur de reconstruction s'avère plus efficace pour déterminer les valeurs adéquates. Les résultats obtenus avec la formulation POD-Galerkin sont encourageants et valident son utilisation pour les développements ultérieurs.

### 3.4 Développement de modèles d'ordre réduit basés sur la POD pour les systèmes dynamiques linéaires non-autonomes

Dans le paragraphe 3.2, plusieurs formulations du modèle d'ordre réduit ont été développées pour un système autonome dans lequel la sollicitation provenait simplement de la condition initiale appliquée sur l'extrémité libre de la barre. La comparaison des différentes formulations a permis de contrôler que la projection de Galerkin était adaptée et on se concentre donc à présent sur l'étude avec cette formulation de la réponse du même système rendu non-autonome par l'introduction d'une sollicitation instationnaire. L'objectif est de parvenir à construire un modèle d'ordre réduit dans lequel les effets de la condition limite imposée interviennent explicitement afin de pouvoir à l'avenir modifier les paramètres de la condition limite directement dans le modèle d'ordre réduit.

La sollicitation imposée sur l'extrémité libre de la barre porte soit sur les déplacements, soit sur la dérivée des déplacements par rapport à la normale externe au domaine. On parlera alors respectivement de condition limite de Dirichlet et de Neumann non-homogènes. La condition limite en  $x = L$  donnée à l'équation (3.2) est donc modifiée par des conditions aux limites du type (2.41).

Les conditions aux limites de type Neumann sont facilement prises en compte dans le modèle d'ordre réduit avec la formulation POD-Galerkin. En effet, le produit scalaire de  $H = L^2([0; L])$  fait intervenir une intégrale à l'aide de laquelle une intégration par parties peut être pratiquée pour introduire explicitement la condition limite dans les équations du modèle d'ordre réduit. La projection des équations (3.1) sur les modes POD conduit à l'expression (3.34) du modèle d'ordre réduit avec la formulation POD-Galerkin. Cependant, seuls les termes correspondant à l'encastrement en  $x = 0$  s'annulent dans les produits scalaires (3.36) tandis que les termes du bord libre en  $x = L$  subsistent. Après simplification, les équations du modèle d'ordre réduit satisfont alors l'égalité

$$\begin{aligned} \ddot{a}_i = & -c^2 \sum_{j=1}^q \left\langle \frac{d\varphi^{(j)}}{dx}, \frac{d\varphi^{(i)}}{dx} \right\rangle a_j - c^2 \left\langle \frac{d\bar{u}}{dx}, \frac{d\varphi^{(i)}}{dx} \right\rangle \\ & + c^2 \left( \sum_{j=1}^q \frac{d\varphi^{(j)}}{dx} \Big|_{x=L} a_j + \frac{d\bar{u}}{dx} \Big|_{x=L} \right) \varphi^{(i)}(x=L). \end{aligned} \quad (3.45)$$

Or la condition limite de Neumann donnée sous forme générale à l'équation (2.41) s'écrit

$$\frac{\partial u}{\partial x} \Big|_{x=L} = \gamma_N(t) \iff \frac{d\bar{u}}{dx} \Big|_{x=L} + \sum_{j=1}^q \frac{d\varphi^{(j)}}{dx} \Big|_{x=L} a_j(t) = \gamma_N(t) \quad (3.46)$$

pour le déplacement  $u$  sur le bord  $\partial\Omega_N$  qui correspond ici au point  $x = L$ . Ainsi, le terme supplémentaire qui subsiste dans l'équation (3.45) s'interprète directement comme la condition de Neumann (3.46) décomposée sur la base des modes POD. La projection des équations donne alors

$$\ddot{a}_i = -c^2 \sum_{j=1}^q \left\langle \frac{d\varphi^{(j)}}{dx}, \frac{d\varphi^{(i)}}{dx} \right\rangle a_j - c^2 \left\langle \frac{d\bar{u}}{dx}, \frac{d\varphi^{(i)}}{dx} \right\rangle + c^2 \gamma_N(t) \varphi^{(i)}(x=L). \quad (3.47)$$

Le modèle d'ordre réduit s'écrit donc comme un système d'équations différentielles ordinaires linéaires non-autonomes dans lequel apparaît un terme source instationnaire supplémentaire qui décrit l'action de la condition limite imposée sur l'extrémité libre. Le système matriciel s'écrit sous la forme synthétique

$$\ddot{\mathbf{a}} = \mathcal{L}\mathbf{a} + \mathcal{K} + \mathcal{S}(t), \quad (3.48)$$

dans laquelle les coefficients des opérateurs sont définis par les relations suivantes :

$$\mathcal{L}_{ij} = -c^2 \sum_{j=1}^q \left\langle \frac{d\varphi^{(j)}}{dx}, \frac{d\varphi^{(i)}}{dx} \right\rangle, \quad \mathcal{K}_i = -c^2 \left\langle \frac{d\bar{u}}{dx}, \frac{d\varphi^{(i)}}{dx} \right\rangle \quad \text{et} \quad \mathcal{S}_i(t) = -c^2 \gamma_N(t) \varphi^{(i)}(x=L). \quad (3.49)$$

Les produits scalaires sont alors évalués selon le principe exposé au paragraphe 3.2.3.1 et les opérateurs constant et linéaire sont donc définis par les relations (3.38) lorsque les produits scalaires sont approchés.

La prise en compte de conditions aux limites de type Dirichlet est plus délicate à mettre en œuvre puisque les déplacements ne peuvent être directement introduits dans les équations du modèle d'ordre réduit au moyen d'intégrations par parties<sup>10</sup>. On peut alors s'inspirer de méthodes issues des techniques Éléments Finis pour tenir compte d'une telle condition limite et recourir ainsi à la méthode des pénalités ou à la méthode des multiplicateurs de Lagrange qui sont par exemple décrites dans [26, 253]. Une troisième solution, inspirée des méthodes de relèvement statique, présente quelques similarités avec les méthodes de synthèse modale [53, 230]. On décrit dans les paragraphes suivants ces trois techniques.

La condition limite de Dirichlet écrite de manière très générale à l'équation (2.41) prend ici la forme suivante :  $u(x = L, t) = \gamma_D(t)$ . Il est important de mentionner que les modes POD ne satisfont pas exactement les conditions aux limites puisque celles-ci ne sont pas homogènes. La relation vérifiée dans ce cas est donnée au paragraphe 2.2.3.4.

### 3.4.1 Adaptation de la méthode des pénalités au modèle d'ordre réduit

Dans le cadre d'une discrétisation Éléments Finis, les facteurs de pénalités peuvent être introduits directement dans le système matriciel sur les lignes correspondant aux degrés de liberté contraints par la condition limite. Toutefois, le modèle d'ordre réduit est formulé pour les amplitudes modales où la notion de déplacement est absente. Une première manière de procéder est d'introduire dans l'équation continue un opérateur de pondération local défini à partir d'une distribution de Dirac qui permet de tenir compte de la condition limite de Dirichlet. Cependant la construction du modèle d'ordre réduit devient difficile à justifier rigoureusement puisqu'il est nécessaire de recourir aux notions de la théorie des distributions. Par conséquent, on préfère ici une seconde approche décrite par Zienkiewicz et Morgan [253], qui consiste à remplacer artificiellement la condition limite de Dirichlet  $u(x = L, t) = \gamma_D(t)$  par une condition limite de Neumann  $\gamma_N(t)$  approchée. La condition est alors relaxée au moyen de la relation

$$\left. \frac{\partial u}{\partial x} \right|_{x=L} = \gamma_N(t) = -\frac{u_L(t) - \gamma_D(t)}{\epsilon_p}. \quad (3.50)$$

Dans l'expression précédente,  $u_L$  représente l'inconnue en déplacement sur l'extrémité libre en  $x = L$  qui doit coïncider avec le déplacement imposé  $\gamma_D(t)$  à tout instant. Lorsque le facteur  $\epsilon_p > 0$  est très faible, la condition de Neumann précédente se réduit à la limite à la condition de Dirichlet souhaitée. On définit alors un facteur de pénalité  $f_p = c^2/\epsilon_p$  dont la valeur doit être élevée pour que la condition limite soit approximativement vérifiée. Puisque la condition limite de Dirichlet a été ramenée à une condition limite de Neumann, il est possible de l'introduire dans les équations du modèle d'ordre réduit au moyen d'une intégration par parties comme on l'a décrit au début de cette section. L'équation (3.47) s'écrit alors en fonction de la condition limite de Dirichlet relaxée au moyen de l'expression (3.50) :

$$\ddot{a}_i = -c^2 \sum_{j=1}^q \left\langle \frac{d\varphi^{(j)}}{dx}, \frac{d\varphi^{(j)}}{dx} \right\rangle a_j - c^2 \left\langle \frac{d\bar{u}}{dx}, \frac{d\varphi^{(i)}}{dx} \right\rangle + f_p (u_L(t) - \gamma_D(t)) \varphi^{(i)}(x = L). \quad (3.51)$$

Finalement, le déplacement  $u_L$  est décomposé sur la base des modes POD puisqu'il s'agit d'un élément de l'ensemble des clichés :  $u_L(t) = \bar{u}_L + \sum_{j=1}^q a_j(t) \varphi_L^{(j)}$  avec  $\varphi_L^{(j)} = \varphi^{(j)}(x = L)$ . L'introduction de cette décomposition dans l'équation (3.51), conduit à l'expression

$$\ddot{a}_i = -c^2 \sum_{j=1}^q \left\langle \frac{d\varphi^{(j)}}{dx}, \frac{d\varphi^{(j)}}{dx} \right\rangle a_j - c^2 \left\langle \frac{d\bar{u}}{dx}, \frac{d\varphi^{(i)}}{dx} \right\rangle + f_p \bar{u}_L \varphi_L^{(i)} + f_p \sum_{j=1}^q \varphi_L^{(i)} \varphi_L^{(j)} a_j - f_p \gamma_D(t) \varphi_L^{(i)}. \quad (3.52)$$

Le modèle d'ordre réduit s'écrit donc sous la même forme matricielle (3.48) que celle obtenue pour la prise en compte de la condition limite de Neumann, mais on précise ici par l'indice MP que les opérateurs découlent de la méthode des pénalités pour une condition limite de Dirichlet. En introduisant le vecteur  $\varphi_L = [\varphi_L^{(1)}, \dots, \varphi_L^{(q)}]^T \in \mathbb{R}^q$ , les opérateurs sont alors définis par les relations

<sup>10</sup>Il est possible d'introduire les conditions aux limites de Dirichlet dans le modèle d'ordre réduit en pratiquant deux intégrations par parties successives, mais dans ce cas il devient nécessaire d'évaluer des dérivées secondes. Les essais menés avec cette technique ont toutefois conduit à un modèle d'ordre réduit dont la réponse diverge à chaque fois.

$$\begin{aligned}
 \mathcal{L}_{\text{MP}} &= \mathcal{L} + \mathcal{L}_p \quad \text{avec} \quad \mathcal{L}_p = f_p \varphi_L \otimes \varphi_L, \\
 \mathcal{K}_{\text{MP}} &= \mathcal{K} + \mathcal{K}_p \quad \text{avec} \quad \mathcal{K}_p = f_p \bar{u}_L \varphi_L, \\
 \mathcal{S}_{\text{MP}}(t) &= -f_p \gamma_D(t) \varphi_L.
 \end{aligned} \tag{3.53}$$

Les opérateurs  $\mathcal{L}$  et  $\mathcal{K}$  sont ceux définis à l'équation (3.49) et sont cette fois complétés par les opérateurs  $\mathcal{L}_p$  et  $\mathcal{K}_p$  qui représentent l'action du terme de pénalité. Enfin le terme source  $\mathcal{S}_{\text{MP}}$  est directement lié à la condition limite de Dirichlet qui peut donc être modifiée dans le modèle d'ordre réduit.

La méthode a été employée par Tang *et al.* [226] qui l'ont comparée à la méthode des fonctions de contrôle présentée dans le paragraphe 3.4.3. Sirisup et Karniadakis [215] ont aussi employé un facteur de pénalité pour imposer une condition limite non-homogène et ont remarqué que la détermination du facteur de pénalité adéquat n'est pas évidente. La meilleure valeur à prendre pour le modèle d'ordre réduit était dans leur cas relativement faible, ce qui laisse penser que la condition limite n'est qu'approximativement vérifiée.

### 3.4.2 Adaptation de la méthode des multiplicateurs de Lagrange au modèle d'ordre réduit

Le principe de la méthode consiste cette fois à adjoindre à l'équation continue du système (3.1) une inconnue supplémentaire notée  $R_\gamma(t)$  qui représente la force de réaction produite sur l'extrémité libre par le déplacement imposé  $u(x = L, t) = \gamma_D(t)$ . Le problème à résoudre est donc transformé en

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = R_\gamma(x, t) \\ u(x = L, t) = \gamma_D(t) \end{cases}. \tag{3.54}$$

La première équation est projetée sur chacun des modes POD suivant la méthode de Galerkin après avoir utilisé la décomposition maintenant habituelle  $u(x, t) = \bar{u}(x) + \sum_{j=1}^q a_j(t) \varphi^{(j)}(x)$ . De manière analogue, la seconde équation est aussi décomposée sur la base des modes POD afin d'obtenir le système

$$\begin{cases} \ddot{a}_i = c^2 \sum_{j=1}^q \left\langle \frac{d^2 \varphi^{(j)}}{dx^2}, \varphi^{(i)} \right\rangle a_j + c^2 \left\langle \frac{d^2 \bar{u}}{dx^2}, \varphi^{(i)} \right\rangle + \left\langle R_\gamma(x, t), \varphi^{(i)} \right\rangle \\ \bar{u}(x = L) + \sum_{j=1}^q a_j(t) \varphi^{(j)}(x = L) = \gamma_D(t) \end{cases}. \tag{3.55}$$

La première ligne du système correspond encore à la forme (3.48) du modèle d'ordre réduit dans laquelle les opérateurs constant et linéaire — notés ici respectivement  $\mathcal{K}_{\text{MML}}$  et  $\mathcal{L}_{\text{MML}}$  pour indiquer qu'il s'agit de la méthode des multiplicateurs de Lagrange — correspondent aux opérateurs  $\mathcal{K}$  et  $\mathcal{L}$  définis à l'équation (3.35). Le terme source  $\mathcal{S}_{\text{MML}}$  est quant à lui défini par la projection de la réaction sur les modes POD.

La réaction  $R_\gamma$  est produite par le déplacement imposé en  $x = L$ , et de ce fait, il est possible de considérer qu'il s'agit d'une force localisée en ce point. Ceci se traduit formellement à l'aide d'une distribution de Dirac par  $R_\gamma(x, t) = r_L(t) \delta(x - L)$ . Par conséquent les composantes du terme source sont transformées en

$$\mathcal{S}_{\text{MML},i}(t) = \left\langle R_\gamma(x, t), \varphi^{(i)} \right\rangle = \int_0^L r_L(t) \delta(x - L) \varphi^{(i)}(x) dx = r_L(t) \varphi_L^{(i)}. \tag{3.56}$$

En utilisant le vecteur  $\varphi_L$  introduit avec la méthode des pénalités dans le paragraphe précédent, le terme source s'écrit alors

$$\mathcal{S}_{\text{MML}}(t) = r_L(t) \varphi_L. \tag{3.57}$$

Enfin la seconde ligne du système (3.55) se réécrit  $\bar{u}_L + \varphi_L^\top \mathbf{a} = \gamma_D(t)$ , de sorte que le système (3.55) se met sous la forme matricielle

$$\begin{bmatrix} \mathbf{I}_q & \mathbf{0}_{q,1} \\ \mathbf{0}_{1,q} & 0 \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{a}} \\ \dot{r}_L \end{bmatrix} = \begin{bmatrix} \mathcal{L}_{\text{MML}} & \varphi_L \\ \varphi_L^\top & 0 \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ r_L \end{bmatrix} + \begin{bmatrix} \mathcal{K}_{\text{MML}} \\ \bar{u}_L - \gamma_D(t) \end{bmatrix}. \tag{3.58}$$

Le problème est donc ramené à un système matriciel de la forme  $\tilde{\mathcal{T}}\ddot{\tilde{\mathbf{a}}} = \tilde{\mathcal{L}}_{\text{MML}}\tilde{\mathbf{a}} + \tilde{\mathcal{K}}_{\text{MML}}$  qui porte sur le vecteur d'inconnues  $\tilde{\mathbf{a}} = [\mathbf{a} \ r_L]^\top$ . Contrairement aux modèles d'ordre réduit construits jusqu'ici, l'intégration en temps doit être menée avec un schéma de Newmark [26] par exemple puisque la matrice  $\tilde{\mathcal{T}}$  n'est pas inversible. Par conséquent il n'est pas possible de mettre le système (3.58) sous la forme générique (3.40) résolue habituellement avec les routines du code ODEPACK.

### 3.4.3 Méthode de la fonction de contrôle

Au lieu de considérer le déplacement imposé comme une contrainte supplémentaire que doit vérifier le modèle d'ordre réduit, l'expression (3.13) de la décomposition orthogonale propre de la variable est modifiée par l'ajout d'un terme supplémentaire correspondant à la sollicitation  $\gamma_D(t)$  imposée au système. Le principe consiste alors à décomposer la solution du système dynamique en deux parties telles que la première contribution  $u_h$  corresponde à la solution satisfaisant des conditions aux limites homogènes tandis que la seconde contribution  $u_\gamma$  représente l'effet de la sollicitation qui vérifie les conditions aux limites non-homogènes. La solution du système dynamique est donc recherchée sous la forme  $u = u_h + u_\gamma$  avec en particulier  $u_h(x = L) = 0$  et  $u_\gamma(x = L) = \gamma_D(t)$ .

La décomposition orthogonale propre est alors appliquée à la partie homogène  $u_h$ . Les modes POD sont calculés à partir de l'opérateur de corrélation construit pour des clichés centrés rendus artificiellement homogènes, notés ici  $s_h^{(m)} = u_h^{(m)} - \bar{u}_h$ . Les modes POD  $\varphi_h^{(j)}$  ainsi obtenus sont homogènes puisqu'ils héritent des conditions aux limites homogènes satisfaites par les clichés centrés  $s_h^{(m)}$  selon la relation (2.40). La partie homogène  $u_h$  se décompose alors en  $u_h(x, t) = \bar{u}_h(x) + \sum_{j=1}^q a_j(t) \varphi_h^{(j)}(x)$  qui est l'homologue de la décomposition (3.13). Pour faciliter le développement du modèle d'ordre réduit, il est judicieux d'effectuer une décomposition espace-temps de la contribution non-homogène telle que  $u_\gamma(x, t) = \gamma_D(t) u_c(x)$  où  $u_c$  est la fonction éponyme de la méthode. La nouvelle décomposition de la variable  $u$  prend donc la forme

$$u(x, t) = u_h(x, t) + u_\gamma(x, t) \quad \text{avec} \quad \begin{cases} u_h(x, t) = \bar{u}_h(x) + \sum_{j=1}^q a_j(t) \varphi_h^{(j)}(x) \\ u_\gamma(x, t) = \gamma_D(t) u_c(x) \end{cases} . \quad (3.59)$$

La décomposition (3.59) nécessite de définir l'expression de la fonction de contrôle  $u_c$ . Le choix est relativement vaste puisque la seule contrainte est que le déplacement à l'extrémité libre coïncide avec la condition limite imposée :  $u_\gamma(x = L, t) = \gamma_D(t)$ . Combinée avec la décomposition  $u_\gamma(x, t) = \gamma_D(t) u_c(x)$ , la contrainte se résume à  $u_c(x = L) = 1$  et la forme de  $u_c$  est a priori quelconque sur l'intervalle  $]0; L[$ . Le choix le plus simple consiste à prendre une distribution de Dirac, nulle partout sauf à l'extrémité libre. Toutefois ce choix n'est certainement pas optimal puisqu'il n'y a aucun lien avec l'équation résolue et, qui plus est, il s'agit d'une distribution qu'il n'est pas toujours simple de manipuler rigoureusement. Un autre choix plus élégant pour la fonction de contrôle consiste à prendre la solution du problème statique en réponse à un déplacement unitaire imposé. Les expressions des fonctions de contrôle données ci-dessous lorsqu'on considère une distribution de Dirac et la solution statique établie dans la note 1 p.65 sont respectivement

$$u_c(x) = \delta(x - L) \quad \text{et} \quad u_c(x) = \frac{x}{L}. \quad (3.60)$$

La nouvelle décomposition (3.59) de la variable  $u$  est alors remplacée dans l'équation continue (3.1). La projection sur la base des modes POD conduit alors à

$$\begin{aligned} \ddot{a}_i &= c^2 \sum_{j=1}^q \left\langle \frac{d^2 \varphi_h^{(j)}}{dx^2}, \varphi_h^{(i)} \right\rangle a_j + c^2 \left\langle \frac{d^2 \bar{u}_h}{dx^2}, \varphi_h^{(i)} \right\rangle \\ &\quad + c^2 \gamma_D(t) \left\langle \frac{d^2 u_c}{dx^2}, \varphi_h^{(i)} \right\rangle - \frac{d^2 \gamma_D(t)}{dt^2} \left\langle u_c, \varphi_h^{(i)} \right\rangle \end{aligned} \quad (3.61)$$

et le modèle d'ordre réduit conserve donc la même forme générique (3.48), mais les coefficients sont cette fois définis par les expressions

$$\begin{aligned}
 \mathcal{L}_{\text{MFC},ij} &= c^2 \sum_{j=1}^q \left\langle \frac{d^2 \varphi_h^{(j)}}{dx^2}, \varphi_h^{(i)} \right\rangle, \\
 \mathcal{K}_{\text{MFC},i} &= c^2 \left\langle \frac{d^2 \bar{u}_h}{dx^2}, \varphi_h^{(i)} \right\rangle, \\
 \mathcal{S}_{\text{MFC},i} &= c^2 \gamma_D(t) \left\langle \frac{d^2 u_c}{dx^2}, \varphi_h^{(i)} \right\rangle - \frac{d^2 \gamma_D(t)}{dt^2} \langle u_c, \varphi_h^{(i)} \rangle.
 \end{aligned} \tag{3.62}$$

Les opérateurs linéaire et constant sont les mêmes que ceux donnés à l'équation (3.35), à la différence près qu'ils sont cette fois définis pour les modes POD et la partie moyenne *homogènes*. Le terme source nécessite quant à lui d'évaluer la dérivée seconde de la fonction de contrôle, ce qui est délicat lorsqu'une distribution de Dirac est choisie.

La méthode des fonctions de contrôle tire son nom de son application dans le domaine du contrôle actif pour tenir compte des conditions d'injection ou d'aspiration d'air sur un profil par exemple. La méthode a en particulier été adoptée par Park et Kim [175], Ravindran [185], Tang *et al.* [226], Vigo [240] ou encore Cordier et Bergmann [49]. Gunzburger *et al.* [101] ont par ailleurs développé et comparé plusieurs méthodes pour tenir compte de conditions aux limites non-homogènes dans le modèle d'ordre réduit lorsque le nombre de paramètres devient important. En effet, la méthode des fonctions de contrôle n'a été présentée ici que pour un seul degré de liberté dans la condition limite et son extension à un nombre important de degrés de liberté n'est pas immédiate.

## 3.5 Évaluation du modèle d'ordre réduit POD-Galerkin du système linéaire non-autonome

La formulation POD-Galerkin qui avait été évaluée dans le cas du système autonome au paragraphe 3.3 est à présent mise en œuvre dans le cas d'un système rendu non-autonome en imposant une condition limite de Dirichlet non-homogène. Les différentes techniques développées dans la section 3.4 pour tenir compte de ce type de condition limite dans le modèle d'ordre réduit basé sur la formulation POD-Galerkin sont alors confrontées et comparées notamment en terme d'erreur de reconstruction. On insiste dans un premier temps sur la forme des modes POD et sur les conditions aux limites qu'ils vérifient.

### 3.5.1 Aspect des modes POD non-homogènes et homogénéisés

Les clichés  $u^{(m)}$  calculés comme les solutions du système dynamique non réduit vérifient la condition limite de Dirichlet non-homogène au niveau du déplacement imposé. Parmi les techniques présentées dans la section précédente, la méthode des pénalités et la méthode des multiplicateurs de Lagrange reposent sur des modes POD non-homogènes qui ont été calculés à partir des clichés centrés non-homogènes  $s^{(m)} = u^{(m)} - \bar{u}$ . La méthode des fonctions de contrôle se distingue quant à elle par le fait que les modes POD sont calculés pour les clichés centrés  $s_h^{(m)} = u_h^{(m)} - \bar{u}_h$  rendus homogènes en soustrayant à chaque instant  $t_m$  la partie non-homogène de sorte que  $u_h^{(m)} = u^{(m)} - u_\gamma(x, t_m)$ .

L'allure des différents modes POD — non-homogènes et rendus homogènes au moyen des fonctions de contrôle (3.60) — est représentée sur les graphiques de la figure 3.9. Dans un premier temps les modes POD  $\varphi^{(j)}$  calculés à partir des clichés non-homogènes peuvent être comparés à ceux calculés à partir de la réponse du système autonome où les clichés vérifiaient des conditions aux limites homogènes. On avait observé sur la figure 3.3 que les modes POD convergeaient dans ce cas vers les modes propres de la barre encastrée-libre. Des différences assez légères sont constatées ici, en particulier au niveau de l'extrémité libre puisque les modes POD ne satisfont plus la condition homogène  $d\varphi^{(i)}/dx|_{x=L} = 0$  mais une condition du type (2.42). Les modes POD  $\varphi_h^{(j)}$  rendus homogènes avec la distribution de Dirac  $u_c(x) = \delta(x - L)$  sont identiques aux modes POD non-homogènes, sauf pour le point  $x = L$  où ils s'annulent. Les modes arborent donc clairement une allure discontinue.

Au contraire, lorsque les modes POD  $\varphi_h^{(j)}$  sont rendus homogènes à l'aide de la solution statique du problème  $u_c(x) = x/L$ , ils deviennent continus sur l'intervalle  $[0; L]$ . Par ailleurs, on peut remarquer que ces modes correspondent aux modes propres de la barre *encastrée aux deux extrémités* comme le confirment les graphiques de la figure 3.9. La méthode des fonctions de contrôle présente alors certaines similarités avec la méthode de synthèse modale décrite par Craig et Bampton [53] puisque les modes POD homogénéisés sont le pendant des modes propres encastrés.

La figure 3.10 représente les valeurs propres associées aux modes POD homogénéisés avec différentes fonctions de contrôle. Le spectre conserve une allure proche de celui obtenu lorsque les modes POD sont calculés à partir des clichés homogènes analytiques de la réponse du système autonome. Cependant on remarque que lorsque les clichés sont rendus homogènes avec la solution statique, le spectre décroît de façon moins prononcée, ce qui signifie que la réponse du système est plus complexe.

### 3.5.2 Réponse du modèle d'ordre réduit non-autonome

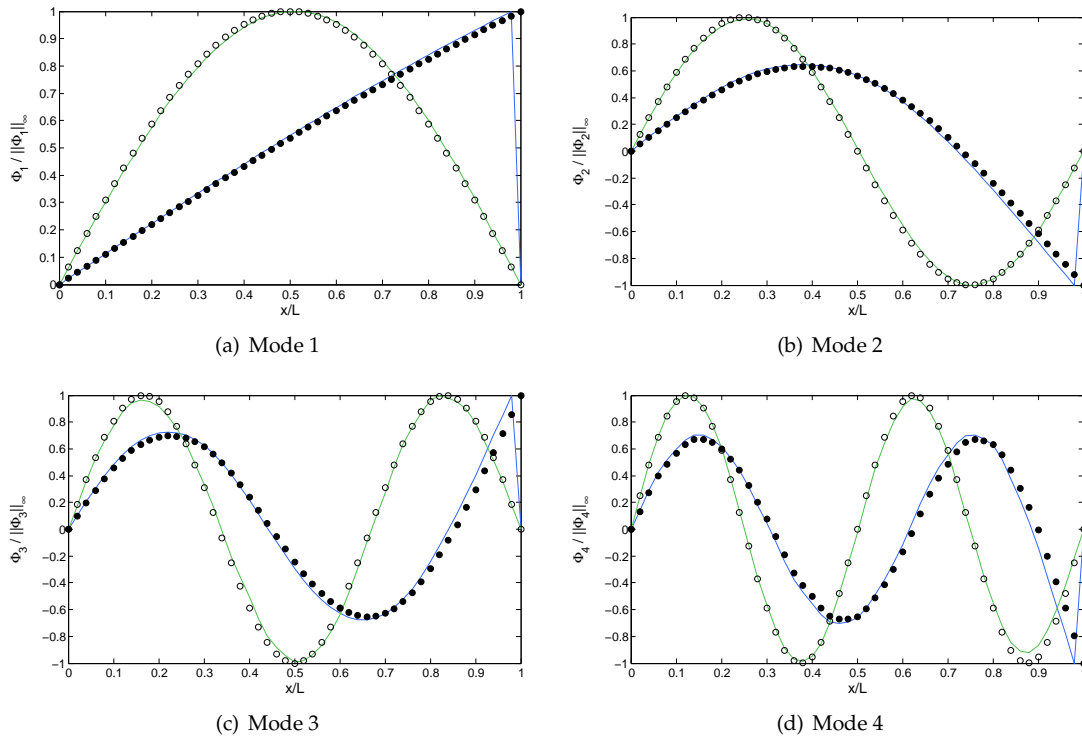
Les modèles d'ordre réduit basés sur les différentes techniques développées au paragraphe 3.4 pour tenir compte de la condition limite sont à présent mis à l'épreuve afin d'évaluer la qualité de la réponse reconstruite pour différents paramètres de l'excitation. Celle-ci est supposée être gouvernée par la loi sinusoïdale

$$\gamma_D(t) = U_e \sin(2\pi f_e t), \quad (3.63)$$

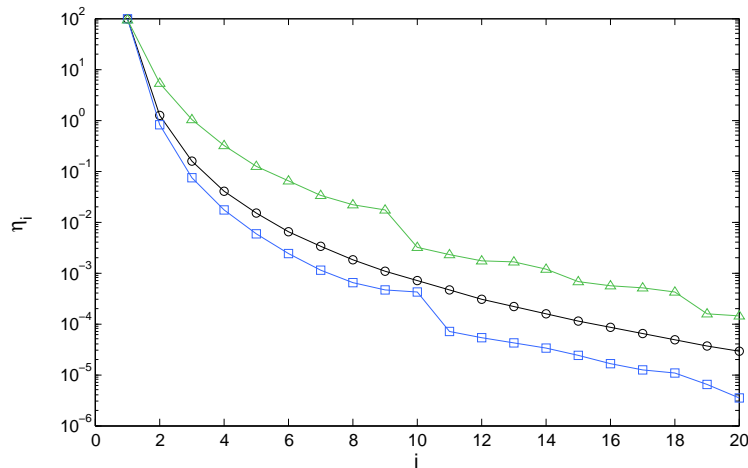
où  $U_e$  représente l'amplitude maximale des oscillations et  $f_e$  est la fréquence d'excitation.

La figure 3.11 présente les réponses reconstruites avec  $q = 5$  modes POD à l'aide du modèle d'ordre réduit basé sur la méthode des fonctions de contrôle dans laquelle la solution du problème statique est employée pour homogénéiser les clichés. Les réponses tracées sur la figure 3.11 correspondent au déplacement du point central de la barre en  $x = L/2$  puisque la réponse du point à l'extrémité est imposée et son





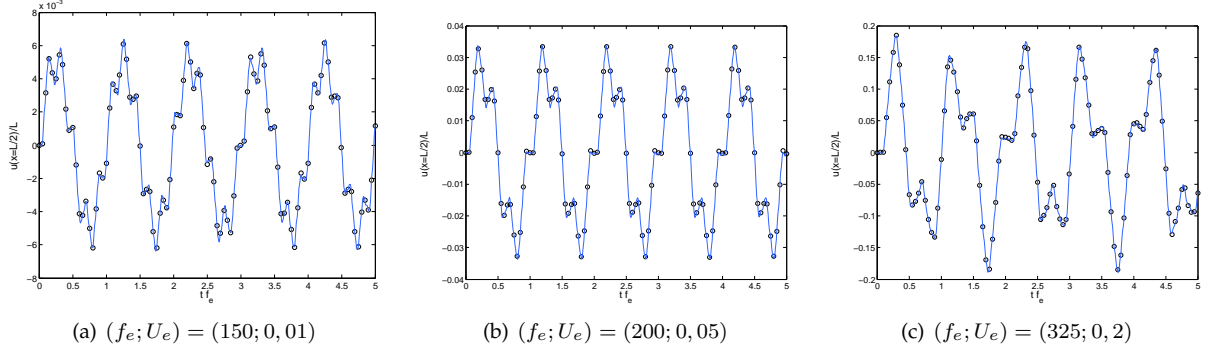
**Figure 3.9 :** Comparaison des quatre premiers modes POD de la barre avec un déplacement imposée à l'extrémité libre. Les courbes (—) correspondent aux modes POD calculés à partir des clichés rendus homogènes à l'aide d'un Dirac  $u_c(x) = \delta(x - L)$ , tandis que les courbes (—) représentent les modes obtenus avec des clichés homogénéisés avec la solution statique du problème  $u_c(x) = x/L$ . Les premiers sont quasiment identiques aux modes POD non-homogènes (●) calculés avec un déplacement imposé à l'extrémité libre, alors que les seconds coïncident avec les modes propres  $\psi_n^{EE} = \sin(n\pi/L x)$  de la barre encastree aux deux extrémités (○).



**Figure 3.10 :** Comparaison du spectre des valeurs propres pour différentes fonctions de contrôle : les carrés (□) correspondent aux valeurs propres associées aux modes POD rendus homogènes avec un Dirac et les triangles (△) à celles associées aux modes POD rendus homogènes avec la solution statique du problème. Les cercles (○) correspondent quant à eux aux valeurs propres associées aux modes POD calculés à partir des clichés homogènes analytiques de la réponse du système autonome.

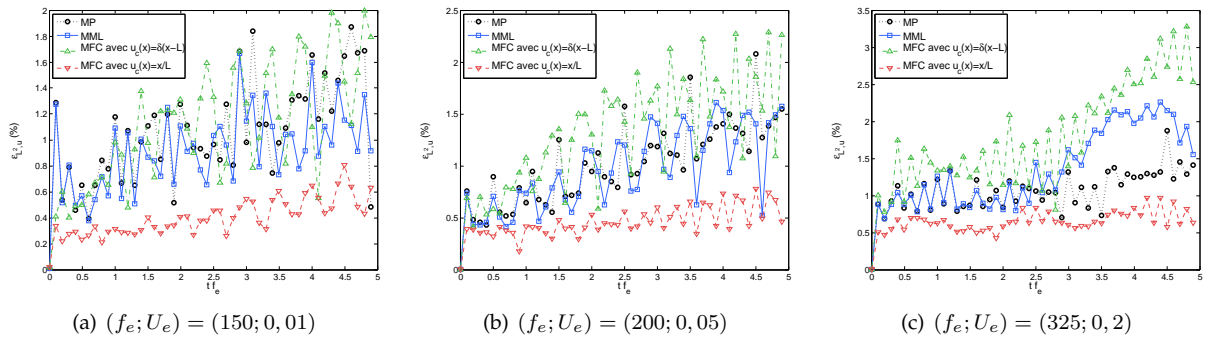
### 3.5. Évaluation du modèle d'ordre réduit POD-Galerkin du système linéaire non-autonome

évolution ne présente donc aucun intérêt. Pour les trois réponses, les modes POD utilisés ont été calculés à partir de clichés issus de la résolution du problème pour une sollicitation  $\gamma_D$  caractérisée par  $f_e = 150$  et  $U_e = 0,01$ . Les modes obtenus permettent alors de reconstruire presque parfaitement les réponses pour d'autres jeux de paramètres, tels que  $(f_e; U_e) = (200; 0,05)$  ou  $(f_e; U_e) = (325; 0,2)$ .



**Figure 3.11 :** Comparaison de la réponse reconstruite par le modèle d'ordre réduit avec différents déplacements imposés à l'extrémité de la barre. Les courbes (—) correspondent à la réponse reconstruite alors que les cercles (o) indiquent les valeurs de la solution de référence.

L'allure de la réponse obtenue avec les autres techniques n'est pas tracée puisque qu'elle est sensiblement identique. Pour comparer la qualité de la réponse obtenue avec les différentes techniques, il est plus judicieux de tracer en fonction du temps l'erreur de reconstruction  $\varepsilon_{L^2, u}$  évaluée au sens de la norme  $L^2$  par rapport à la solution de référence. Celle-ci est donc évaluée pour le modèle d'ordre réduit utilisant la méthode de la fonction de contrôle (MFC) avec un Dirac ou la solution du problème statique, ainsi que pour la méthode des pénalités (MP) et la méthode des multiplicateurs de Lagrange (MML). Le résultat est tracé sur la figure 3.12 pour les trois jeux de paramètres considérés précédemment. À première vue, il ressort que l'erreur augmente en fonction du temps, quelle que soit la formulation employée. Cet accroissement de l'erreur varie néanmoins selon la technique adoptée pour tenir compte de la condition limite dans le modèle d'ordre réduit. L'erreur la plus faible est toujours commise lorsque la solution est reconstruite pour le même jeu de paramètres que celui employé pour générer la base de projection et l'écart avec la solution de référence augmente d'autant plus que les paramètres utilisés pour la reconstruction s'écartent de ceux employés pour calculer les clichés.

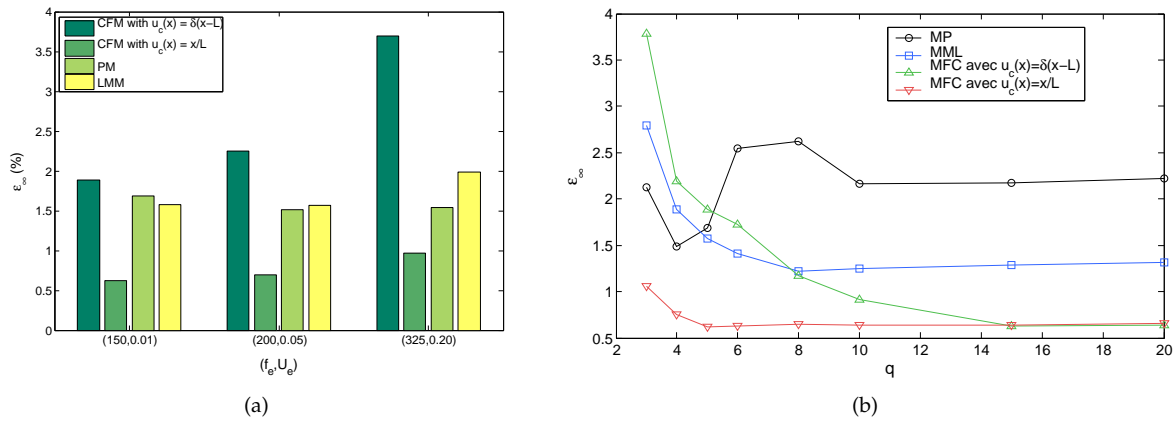


**Figure 3.12 :** Comparaison de l'erreur  $\varepsilon_{L^2, u}$  de reconstruction engendrée par les réponses reconstruites par le modèle d'ordre réduit dans lequel la condition limite est prise en compte par différentes techniques.

Dans l'exemple traité ici, il apparaît que la méthode des fonctions de contrôle utilisant la solution du problème statique fournit les meilleurs résultats : l'erreur de reconstruction est globalement plus faible et augmente moins rapidement qu'avec les autres formulations. Au contraire, lorsque cette formulation

est employée avec un Dirac pour assurer le rôle de la fonction de contrôle, l'erreur augmente considérablement. Ceci peut s'expliquer simplement par le fait que les modes POD homogénéisés avec la solution statique possèdent une forme continue régulière qui se prête bien aux opérations de dérivation et d'intégration numériques nécessaires pour construire le modèle d'ordre réduit. Entre ces deux extrêmes se situent la méthode des pénalités et la méthode des multiplicateurs de Lagrange dont les performances sont à peu près équivalentes. Le dernier jeu de paramètres laisse apparaître une légère différence entre ces deux méthodes en faveur de la méthode des pénalités.

La figure 3.13(a) résume l'évolution de l'erreur infinie de reconstruction  $\varepsilon_\infty$  calculée pour les différents jeux de paramètres avec les trois techniques de prise en compte de la condition limite de Dirichlet. L'erreur est calculée sur cinq périodes d'oscillations forcées avec des modèles d'ordre réduit utilisant  $q = 5$  modes POD dans la base de projection. L'évolution observée confirme celle des graphiques de la figure 3.12 puisque la méthode des fonctions de contrôle reste la plus efficace lorsque la solution statique est employée pour homogénéiser les clichés, tandis que l'emploi du Dirac fournit les résultats les moins bons.



**Figure 3.13 :** Le graphique de la figure (a) représente l'évolution de l'erreur infinie  $\varepsilon_\infty$  calculée avec les différentes techniques de prise en compte des conditions aux limites pour chaque jeu de paramètres  $(f_e; U_e)$ . Le graphique (b) donne pour  $(f_e; U_e) = (150; 0, 01)$  la convergence de l'erreur infinie de reconstruction  $\varepsilon_\infty$  en fonction du nombre de modes POD pour les différents modèles d'ordre réduit avec déplacement imposé.

Pour terminer, il est important de contrôler que chaque technique respecte bien la propriété de convergence de l'erreur de reconstruction lorsque le nombre  $q$  de modes POD retenus dans la base modale augmente. Le graphique 3.13(b) présente pour chaque formulation les courbes de convergence pour  $q = \llbracket 3; 20 \rrbracket$ . En dehors de la méthode des pénalités qui présente un comportement irrégulier, l'erreur de reconstruction converge de manière monotone vers une valeur asymptotique lorsque le nombre de modes POD devient grand. Ces résultats peuvent s'expliquer par le fait qu'avec la méthode des pénalités, la condition limite n'est pas exactement vérifiée mais est d'autant mieux approchée que le facteur de pénalité est grand<sup>11</sup>. Le graphique permet de confirmer le fait que la méthode des fonctions de contrôle dans laquelle la solution du problème statique est employée est la plus performante, tant au niveau de la valeur de l'erreur asymptotique que de la rapidité avec laquelle elle est atteinte. Lorsqu'un Dirac est employé à la place de la solution statique, la technique devient satisfaisante seulement lorsque le nombre de modes POD est élevé. Les erreurs relevées sur le graphique de la figure 3.13(a) pour  $q = 5$  modes POD avaient conduit à la conclusion que cette formulation était la moins performante, mais en réalité elle peut atteindre le même degré de précision qu'avec la solution statique. Enfin, la méthode des multiplicateurs de Lagrange produit une erreur intermédiaire.

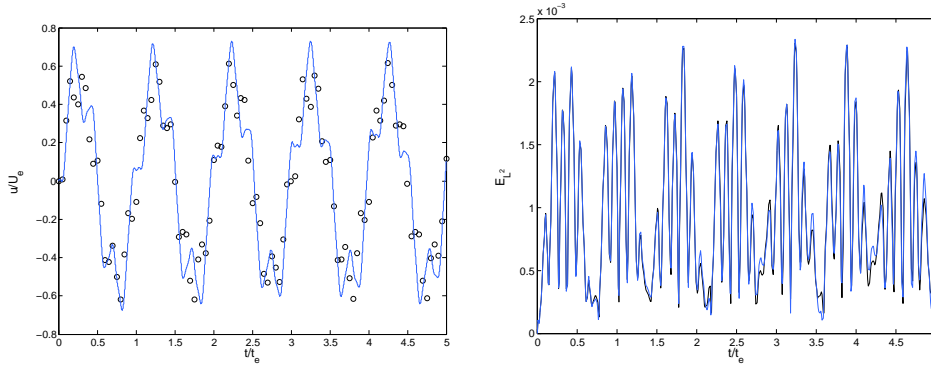
Un dernier point important concerne le type de réponse pouvant être reconstruit en fonction de la nature des modes POD employés. La question se pose en effet de savoir si il est possible — dans le cas d'un déplacement imposé sur l'extrémité de la barre — d'employer une base de projection dans laquelle les modes

<sup>11</sup>En pratique cela n'est pas toujours vrai et la détermination du facteur de pénalité adéquat est un problème difficile sur lequel se sont penchés Sirisup et Karniadakis [215].

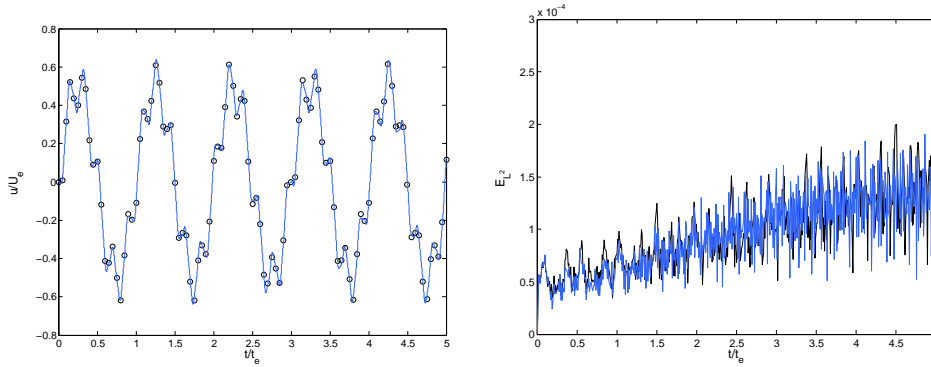
### 3.5. Évaluation du modèle d'ordre réduit POD-Galerkin du système linéaire non-autonome

POD ont été calculés avec des clichés solutions du problème autonome par exemple, ou encore des clichés solutions d'un problème avec une condition limite non-homogène de type Neumann au lieu de Dirichlet.

Ceci devient problématique lorsqu'il s'agit de travailler avec la méthode des fonctions de contrôle puisque les modes POD doivent être homogénéisés afin d'être compatibles avec le modèle d'ordre réduit. Supposons par exemple que les modes POD soient calculés avec les clichés provenant du système non-autonome avec une condition limite non-homogène de Neumann imposée à l'extrémité. Il s'agit alors de déterminer l'expression du terme de forçage  $u_\gamma(x, t) = \gamma_D(t) u_c(x)$  qui convient pour homogénéiser les modes, sachant que la modulation temporelle  $\gamma_D(t)$  est inconnue puisque le déplacement de l'extrémité est gouverné par une condition limite de Neumann. Par ailleurs, si les modes POD sont calculés à partir des clichés du système autonome, ils sont déjà homogènes et dans ce cas la décomposition (3.59) ne peut être employée puisque le terme de forçage est absent.



(a) Utilisation des modes POD du système autonome



(b) Utilisation des modes POD du système non-autonome vérifiant une condition limite de Neumann

**Figure 3.14 :** Reconstruction de la solution pour  $(f_e; U_e) = (150; 0, 01)$  à l'aide des modes POD de diverse nature : les graphiques de la figure (a) donnent la réponse et l'erreur commise lorsque les modes POD sont calculés à partir des clichés du système autonome ; sur la figure (b) les modes POD sont calculés avec les clichés du système non-autonome satisfaisant une condition limite non-homogène de Neumann. Dans les deux cas, la réponse de référence ( $\circ$ ) est comparée à celle obtenue avec la méthode des multiplicateurs de Lagrange ( $\text{—}$ ). Sur les graphiques d'erreur à droite, la courbe ( $\text{—}$ ) correspond à la méthode des multiplicateurs de Lagrange tandis que la méthode des pénalités est représentée par ( $\text{—}$ ).

En revanche, ces difficultés disparaissent avec la méthode des pénalités et la méthode des multiplicateurs de Lagrange puisqu'il n'existe pas d'étape d'homogénéisation des modes. Les graphiques de la figure 3.14 présentent la réponse obtenue avec ces techniques ainsi que l'évolution de l'erreur de reconstruction  $\varepsilon_{L^2, u}$  lorsque le modèle d'ordre réduit est construit avec les modes POD du système autonome ou du système non-autonome avec une condition limite non-homogène de Neumann imposée sur l'extrémité libre. On peut vérifier que l'utilisation des modes POD du système autonome permet certes de retrouver l'allure globale de la réponse, mais l'erreur devient visible. L'erreur infinie  $\varepsilon_\infty$  stagne autour de 20%, et ceci quel que soit le nombre de modes POD  $q$  employé. Au contraire, lorsque les modes POD du système non-autonome

avec une force imposée sont utilisés, la réponse est en bon accord avec celle attendue et l'erreur de reconstruction reste satisfaisante : ainsi, la méthode des pénalités (respectivement la méthode des multiplicateurs de Lagrange) produit une erreur infinie  $\varepsilon_\infty = 1,46\%$  (resp.  $\varepsilon_\infty = 1,62\%$ ) avec ce type de modes POD alors qu'elle s'élevait à  $1,69\%$  (resp.  $1,57\%$ ) avec les modes POD calculés à partir de clichés issus de la réponse du système non-autonome avec la condition limite de Dirichlet. Ces résultats s'expliquent simplement par le fait que les modes POD calculés avec les clichés satisfaisant la condition limite de Neumann vérifient aussi une condition limite non-homogène à l'extrémité libre, ce qui n'est pas le cas lorsque les modes sont ceux du système autonome.

En conclusion de cette section, il s'avère que la méthode de la fonction de contrôle est la plus performante en terme d'erreur de reconstruction lorsque la fonction de contrôle  $u_c$  est correctement choisie. La solution statique du problème avec une condition limite de Dirichlet unitaire paraît être un bon candidat pour homogénéiser les clichés à partir desquels les modes POD sont calculés. De leur côté, la méthode des pénalités et la méthode des multiplicateurs de Lagrange produisent une erreur de reconstruction légèrement plus importante, mais sont plus flexibles lorsqu'ils s'agit de construire un modèle d'ordre réduit avec des modes POD calculés à partir de clichés issus de réponses de nature diverse — telle que celle du système autonome ou du système non-autonome avec une condition limite non-homogène de Neumann par exemple. La méthode des multiplicateurs de Lagrange semble toutefois plus robuste puisque la convergence de l'erreur de reconstruction lorsque le nombre de modes POD augmente a pu être vérifiée. Une des limitations majeures de la méthode des pénalités provient du fait que la condition limite n'est pas exactement vérifiée, ce qui explique certainement pourquoi l'erreur asymptotique est plus élevée qu'avec les autres méthodes.

### 3.6 Synthèse du chapitre 3

Dans ce chapitre, nous nous sommes intéressés dans un premier temps au développement d'un modèle d'ordre réduit pour un système linéaire autonome. Le choix volontaire d'un système extrêmement simple a permis de développer puis de comparer trois formulations du modèle d'ordre réduit qui s'appuient respectivement sur l'existence d'une solution analytique pour la première d'entre elles, sur les équations discrétisées du système pour la formulation dite *discrète* et enfin sur les équations continues pour la formulation dite *POD-Galerkin*. Ces deux dernières formulations qui avaient été présentées au paragraphe 2.4 ont donc pu être comparées au modèle d'ordre réduit de référence construit à partir de la solution analytique et les résultats ont conduit à la conclusion que la formulation POD-Galerkin était particulièrement bien adaptée pour reproduire la réponse d'un tel système dynamique. La supériorité de la méthode par rapport à l'approche discrète peut être attribuée au fait que la formulation repose sur les équations continues du système dynamique.

Des résultats intéressants sur l'influence des paramètres de construction du modèle d'ordre réduit ont par ailleurs pu être établis et la convergence de l'erreur de reconstruction a été étudiée en fonction du nombre de modes POD conservés dans la base de projection. Par ailleurs, des résultats complémentaires concernant l'échantillonnage des clichés pour une réponse périodique et des remarques relatives à la convergence des modes POD vers les modes propres pour ce type particulier de système dynamique ont été ajoutés dans l'annexe C.

Dans un second temps, la formulation POD-Galerkin a été étendue pour reproduire la réponse du même système dynamique rendu non-autonome par l'introduction d'une condition limite de Dirichlet non-homogène. Trois techniques ont donc été développées pour prendre en compte explicitement ce type de condition limite dans le modèle d'ordre réduit. La méthode des pénalités et celle des multiplicateurs de Lagrange qui dérivent directement de techniques employées pour traiter ce type de condition limite dans les problèmes discrétisés par Éléments Finis ont été adaptées au modèle d'ordre réduit avec la formulation POD-Galerkin. Une troisième méthode qui présente certaines similarités avec les méthodes de relèvement statique et appelée ici méthode des fonctions de contrôle consiste à effectuer une décomposition orthogonale propre de clichés qui ont été au préalable homogénéisés par ladite fonction de contrôle.

Les comparaisons de ces trois techniques ont permis de conclure que la méthode des fonctions de contrôle était la plus efficace en terme d'erreur de reconstruction et de convergence lorsque le nombre de modes POD conservés dans la base de projection augmente. Ceci nécessite toutefois d'employer une fonction de contrôle adéquate, telle que la solution du problème statique par exemple. La méthode des multiplicateurs de Lagrange fournit des résultats intermédiaires alors que la méthode des pénalités semble souffrir du fait que la condition limite est seulement approchée et que le choix du facteur de pénalité peut être délicat. Ces deux formulations sont néanmoins plus souples lorsqu'il s'agit de construire un modèle d'ordre réduit avec des modes POD construits à partir de clichés issus d'un autre type de réponse.

Au vu des résultats obtenus dans ce chapitre, la formulation POD-Galerkin semble être un bon candidat pour le développement futur de modèles d'ordre réduit non-linéaires. La formulation, qui n'a été employée ici que dans le cas d'un système monovarié, est donc étendue dans la deuxième partie de ce manuscrit au cas d'un système dynamique non-linéaire multivarié qui gouverne le comportement aérodynamique d'un fluide compressible.



## **Deuxième partie**

# **Développement de modèles d'ordre réduit non-linéaires pour l'aéroélasticité**





# Détermination des équations du fluide en vue de la construction d'un modèle d'ordre réduit POD-Galerkin

## Sommaire

<b>4.1</b>	<b>Introduction</b>	<b>98</b>
<b>4.2</b>	<b>Modélisation d'un fluide compressible visqueux avec les variables conservatives</b>	<b>99</b>
4.2.1	Détermination des hypothèses relatives au modèle fluide	99
4.2.2	Caractérisation d'un fluide newtonien homogène	100
4.2.2.1	Descriptions cinématiques du fluide	100
4.2.2.2	Bilans de conservation des grandeurs physiques	102
4.2.2.3	Loi de comportement du fluide et équation d'état	104
4.2.3	Équations de Navier-Stokes en formulation ALE avec les variables conservatives	105
4.2.3.1	Expression des équations dans le repère absolu	105
4.2.3.2	Expression des équations dans le repère entraîné	106
<b>4.3</b>	<b>Problématique de l'aéroélasticité et méthodologie de résolution</b>	<b>108</b>
4.3.1	Principe de l'aéroélasticité	108
4.3.2	Formulations du problème de couplage	109
4.3.2.1	Approche monolithique vs approche couplée	109
4.3.2.2	Couplage fort vs couplage faible	110
4.3.3	Définition du problème aéroélastique général	110
<b>4.4</b>	<b>Modélisation du fluide avec les variables primitives modifiées</b>	<b>114</b>
4.4.1	Variables aérodynamiques pour l'obtention d'un système quadratique	114
4.4.2	Définition du produit scalaire	115
4.4.3	Équations de Navier-Stokes en formulation ALE avec les variables primitives modifiées	117
4.4.3.1	Formulation dans le repère absolu	117
4.4.3.2	Formulation dans le repère entraîné	117
<b>4.5</b>	<b>Synthèse du chapitre 4</b>	<b>120</b>

## 4.1 Introduction

Dans la partie précédente, le modèle d'ordre réduit d'un système dynamique linéaire a été développé au moyen de diverses formulations afin de révéler le potentiel de la formulation POD-Galerkin. D'après les remarques effectuées au paragraphe 2.4.3, cette formulation s'impose comme le meilleur choix lorsqu'il s'agit de construire un modèle d'ordre réduit dans lequel les non-linéarités sont préservées. L'objectif, dans cette seconde partie, est donc de parvenir à construire le modèle d'ordre réduit d'un système aérodynamique non-linéaire pour les écoulements instationnaires d'un fluide compressible visqueux.

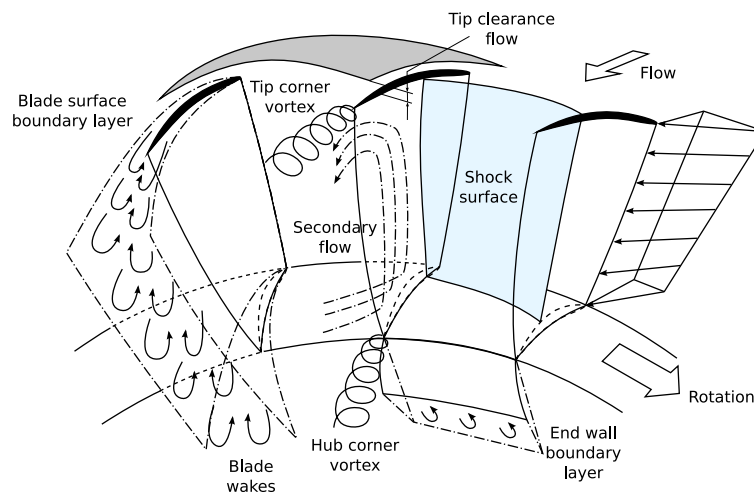
Ce premier chapitre constitue une étape préliminaire pour la construction d'un tel modèle d'ordre réduit. Le but est d'établir les équations continues qui serviront à modéliser le comportement non-linéaire d'un fluide compressible visqueux impliqué dans un problème de couplage fluide-structure. Pour cela, les équations du fluide sont présentées dans un premier temps sous la forme d'équations conservatives habituellement employées dans les codes de calcul pour la résolution numérique. Plusieurs formulations des équations sont données afin de tenir compte du mouvement d'une structure immergée dans le fluide, en se plaçant dans un repère absolu ou entraîné. La formulation des équations dans ce dernier cas se justifie en particulier lorsque le système est animé d'un mouvement de corps rigide, tel que la rotation des systèmes aubagés dans les turbomachines. Le problème aéroélastique est ensuite défini de manière générale et le principe du couplage est explicité.

Les équations du fluide écrites à l'aide des variables conservatives sont cependant inadaptées pour la construction d'un modèle d'ordre réduit avec l'approche POD-Galerkin puisque le système dynamique présente des non-linéarités qui ne sont pas polynomiales. La reformulation des équations locales du fluide avec les *variables primitives modifiées* est un moyen judicieux de ramener le système dynamique sous la forme souhaitée qui se prête bien à la construction du modèle d'ordre réduit. Les équations du fluide écrites avec les variables primitives modifiées dans un repère absolu et entraîné sont alors établies et constitueront dans les chapitres suivants les équations continues de base pour la projection de Galerkin.

## 4.2 Modélisation d'un fluide compressible visqueux avec les variables conservatives

### 4.2.1 Détermination des hypothèses relatives au modèle fluide

L'objectif de cette thèse est de parvenir à construire un modèle d'ordre réduit capable de reproduire les écoulements aérodynamiques au sein des turbomachines. Les équations continues du système dynamique qui serviront pour la projection de Galerkin doivent donc nécessairement avoir été développées pour être en mesure de reproduire le maximum de phénomènes impliqués dans la réponse du fluide<sup>1</sup>. On rappelle dans ce premier paragraphe les principaux types de phénomènes qui caractérisent la réponse de l'écoulement. À partir de là, les hypothèses relatives aux principaux phénomènes sont conservées pour établir les équations continues. Même si ces hypothèses proviennent de l'analyse des phénomènes relatifs aux turbomachines, les équations employées restent très générales comme on le verra par la suite et les modèles d'ordre réduit pourront donc être employés pour reproduire des réponses de nature assez diverse.



**Figure 4.1 :** Schématisation des principaux phénomènes aérodynamiques au sein des turbomachines (d'après W.D. Mc Nally, Review of experimental work on transonic flow in turbomachinery, 1977).

L'écoulement au sein d'une turbomachine est particulièrement complexe en raison de la géométrie des pièces mécaniques composant le système, du mouvement relatif des différents éléments les uns par rapport aux autres et des conditions de fonctionnement extrêmes. La lecture du AGARD Manual vol.1 [3], du rapport de la NASA de Verdon *et al.* [237], des chapitres 8 et 9 de l'ouvrage de Dowell *et al.* [58] ou encore de l'article de Marshall et Imregun [161] fournira un aperçu assez exhaustif des principales caractéristiques de l'écoulement qui sont résumées sur le schéma de la figure 4.1. Ici, on retiendra en particulier les aspects suivants de l'écoulement :

- *Présence de chocs* : sur l'ensemble de la plage de fonctionnement d'une turbomachine, l'écoulement évolue du régime subsonique vers le supersonique, en passant par un régime transsonique. Le fluide est par conséquent *compressible* et des zones de discontinuité correspondant à des chocs dont la position peut osciller apparaissent.
- *Écoulement séparé* : dans certaines plages de fonctionnement, les couches de fluide se séparent et donnent éventuellement naissance à des détachements de vortex dans le sillage des profils. Le caractère *visqueux* du fluide à l'origine de la séparation a une influence significative sur l'amortissement aérodynamique du système — et donc sur la limite de stabilité — ou encore sur la fréquence d'oscillation des ondes de choc en régime transsonique [97].

<sup>1</sup>Dans l'idéal, le modèle d'ordre élevé employé pour générer les clichés doit vérifier (au moins) les mêmes hypothèses.

- *Écoulement turbulent* : au vu des conditions opérationnelles, le nombre de Reynolds qui traduit le caractère turbulent de l'écoulement est en général très élevé. La modélisation de la turbulence est un enjeu fondamental puisqu'elle conditionne la précision de la reproduction des effets visqueux.
- *Interactions stator/rotor* : l'écoulement arrivant sur les aubes d'un stator dans une turbomachine à plusieurs étages est fortement influencé par la proximité de l'étage amont. Trois types d'interactions peuvent être mis en évidence : une interaction sillage-rotor provoquée par le sillage des aubes de l'étage amont [84], une interaction potentielle-rotor due au champ potentiel de l'écoulement des étages amont et aval, et enfin une interaction provoquée par les retours de pression fluctuante.
- *Décrochage en rotation* : un dernier phénomène mentionné ici est le décrochage en rotation — ou en propagation — qui correspond à un écoulement circonférentiel asymétrique dans les compresseurs axiaux. Il se manifeste par une ou plusieurs régions tournant autour de l'axe du compresseur avec une vitesse d'écoulement faible voire inversée. La formation des structures caractéristiques de ce type de décrochage a été étudiée notamment dans [96, 174].

Outre les phénomènes mentionnés ci-dessus, il faut signaler l'influence des effets tridimensionnels tels que les effets de bords étudiés par exemple par Bell et He [27] et analysés expérimentalement par Yu et Liu [250]. Enfin, les non-linéarités de l'écoulement sont à l'origine de phénomènes tels que l'apparition de cycles limites d'oscillation qui sont particulièrement importants dans le régime transsonique [40, 58, 133].

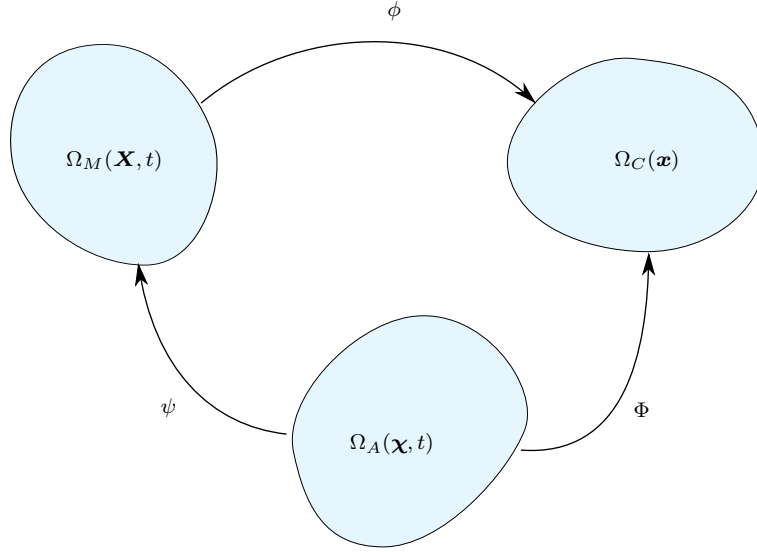
Les modèles physiques du fluide les plus élaborés sont capables de reproduire la plupart des phénomènes décrits ci-dessus, même si la modélisation des effets turbulents, entre autres, reste difficile. L'objectif étant ici de construire un modèle d'ordre réduit, il ne s'agit pas de reproduire l'ensemble des phénomènes, mais seulement les principaux à partir desquels une estimation correcte de la fréquence et de l'amortissement des efforts aérodynamiques peut être obtenue. On se restreint donc à l'étude d'un seul étage et même d'un seul secteur d'une roue aubagée de turbomachine, de sorte que les phénomènes d'interaction stator/rotor ou le décrochage en rotation ne sont pas considérés. Par ailleurs, puisque la modélisation de la turbulence reste un enjeu d'actualité, on négligera en première approximation ces effets. En conclusion, le modèle fluide devra être construit de façon à ce que le caractère *non-linéaire, visqueux, compressible* et éventuellement *tridimensionnel* de l'écoulement puisse être reproduit.

## 4.2.2 Caractérisation d'un fluide newtonien homogène

### 4.2.2.1 Descriptions cinématiques du fluide

On considère un domaine spatial  $\Omega \subset \mathbb{R}^d$  de frontière  $\partial\Omega$  et de normale  $\mathbf{n}$  dirigée vers l'extérieur, dans lequel le fluide est décrit sur l'intervalle de temps de simulation  $I = [t_0; t_0 + t_f]$ . L'approche *lagrangienne* fournit la description cinématique d'une quantité physique associée à une certaine quantité de matière contenue un domaine  $\Omega_M$  tel que  $\Omega_M(t_0) = \Omega$ . Si les particules parcourent de longues distances sur l'intervalle de temps considéré, la description lagrangienne n'est pas adaptée puisque le domaine  $\Omega_M(t)$  se déplace et se déforme considérablement pour continuer à suivre la quantité de matière considérée à l'instant initial. Un autre point de vue consiste à considérer un domaine de contrôle  $\Omega_C$  invariable au cours du temps et sur lequel on suit l'évolution d'une quantité physique qui est cette fois associée à un certain domaine spatial. Cette description dite *eulérienne* est en particulier adaptée lorsqu'il s'agit d'étudier le comportement du fluide dans une certaine région confinée de l'espace.

Cependant, lorsqu'une structure mobile et/ou déformable est placée dans l'écoulement, une partie de la frontière  $\partial\Omega$  n'est plus invariante au cours du temps si bien qu'il n'est plus possible de définir un domaine spatial  $\Omega_C$  indépendant du temps qui couvre l'ensemble du domaine d'intérêt. La solution consiste alors à introduire un domaine *arbitraire* noté  $\Omega_A$  tel que la portion de frontière fixe coïncide avec une description eulérienne, tandis que la portion de l'interface mobile entre le fluide et la structure suit le mouvement lagrangien de la structure. Les points à l'intérieur du domaine sont alors disposés de manière arbitraire, d'où le nom de description *Arbitrairement Lagrangienne-Eulérienne*, communément abrégée en ALE. Une brève présentation des différentes descriptions est fournie ci-dessous. Pour plus de détails, le lecteur pourra se reporter à [57] par exemple.



**Figure 4.2 :** Représentation du domaine matériel  $\Omega_M(\mathbf{X}, t)$  de la description lagrangienne, du domaine de contrôle  $\Omega_C(\mathbf{x})$  de la description eulérienne et du domaine arbitraire  $\Omega_A(\boldsymbol{\chi}, t)$  de la description ALE. Les trois domaines sont reliés entre eux par les applications  $\phi$ ,  $\psi$  et  $\Phi$ .

On introduit pour commencer les coordonnées matérielles  $\mathbf{X}$  associées au domaine matériel  $\Omega_M(\mathbf{X}, t)$  de la description lagrangienne, ainsi que les coordonnées spatiales  $\mathbf{x}$  correspondant quant à elles au domaine de contrôle  $\Omega_C(\mathbf{x})$  de la description eulérienne. Par analogie, on introduit finalement les coordonnées arbitraires  $\boldsymbol{\chi}$  associées au domaine arbitraire  $\Omega_A(\boldsymbol{\chi}, t)$  de la description ALE. Les trois domaines sont reliés entre eux au moyen des applications  $\phi$ ,  $\psi$  et  $\Phi$ . La figure 4.2 représente schématiquement les trois domaines et indique les relations définies par les applications précédentes.

Le lien entre les coordonnées matérielles  $\mathbf{X}$  et les coordonnées spatiales  $\mathbf{x}$  est décrit par l'application

$$\begin{aligned} \phi: \Omega_M \times I &\longrightarrow \Omega_C \times I \\ (\mathbf{X}, t) &\longmapsto \phi(\mathbf{X}, t) = (\mathbf{x}, t), \end{aligned} \quad (4.1)$$

qui définit la loi de mouvement  $\mathbf{x} = \mathbf{x}(\mathbf{X}, t)$  à partir de laquelle on introduit la vitesse matérielle — ou vitesse lagrangienne — notée

$$\mathbf{u}(\mathbf{X}, t) = \left. \frac{\partial \mathbf{x}}{\partial t} \right|_{\mathbf{X}}. \quad (4.2)$$

La vitesse précédente est donc celle de la particule  $\mathbf{X}$  qu'on suit dans son mouvement. La vitesse eulérienne est quant à elle notée simplement

$$\mathbf{u} = \mathbf{u}(\mathbf{x}, t) \quad (4.3)$$

et correspond à la vitesse en un point  $\mathbf{x}$  de l'espace. Les équations décrivant le comportement du fluide découlent de lois de bilans formulées sur le domaine  $\Omega$  et détaillent l'évolution au cours du temps de différentes quantités physiques. Il est donc nécessaire d'établir une relation entre la dérivée temporelle dans le domaine  $\Omega_M$  d'une quantité physique notée  $f(\mathbf{X}, t)$  et la dérivée dans le domaine spatial  $\Omega_C$  de la même quantité physique notée  $f(\mathbf{x}, t)$ . Les règles de composition des fonctions permettent d'écrire  $f(\mathbf{X}, t) = f(\mathbf{x}, t) \circ \phi(\mathbf{X}, t)$  et la dérivation par rapport au temps de la relation précédente donne

$$\left. \frac{\partial f(\mathbf{X}, t)}{\partial t} \right|_{\mathbf{X}} = \left. \frac{\partial f(\mathbf{x}, t)}{\partial t} \right|_{\mathbf{x}} + \mathbf{u}(\mathbf{x}, t) \cdot \nabla_{\mathbf{x}} f(\mathbf{x}, t). \quad (4.4)$$

L'expression ci-dessus fait apparaître l'opérateur  $\nabla$  introduit à l'équation (D.13) et qui permet ici de définir le gradient  $\nabla_{\mathbf{x}} f(\mathbf{x}, t)$  de la fonction par rapport au système de coordonnées  $\mathbf{x}$ . La dérivée temporelle

$\partial \cdot / \partial t|_{\mathbf{X}}$ , parfois notée simplement  $d \cdot / dt$  ou encore  $D \cdot / Dt$ , correspond à la *dérivée matérielle*. Pour des quantités intégrales, la relation (4.4) est transformée pour aboutir à l'expression

$$\frac{\partial}{\partial t} \Big|_{\mathbf{X}} \int_{\Omega_M(t)} f(\mathbf{x}, t) d\Omega = \int_{\Omega_C(t)} \frac{\partial f(\mathbf{x}, t)}{\partial t} \Big|_{\mathbf{x}} d\Omega + \oint_{\partial\Omega_C(t)} f(\mathbf{x}, t) \mathbf{u}(\mathbf{x}, t) \cdot \mathbf{n} d\partial\Omega \quad (4.5)$$

qui correspond au *théorème de Reynolds*. Le même type de relation peut être établi entre le domaine arbitraire  $\Omega_A$  et le domaine spatial  $\Omega_C$  à partir de l'application

$$\begin{aligned} \Phi : \Omega_A \times I &\longrightarrow \Omega_C \times I \\ (\boldsymbol{\chi}, t) &\longmapsto \Phi(\boldsymbol{\chi}, t) = (\mathbf{x}, t). \end{aligned} \quad (4.6)$$

Par analogie, la vitesse de déplacement des points du domaine arbitraire par rapport aux coordonnées du domaine spatial est définie par

$$\mathbf{s}(\boldsymbol{\chi}, t) = \frac{\partial \mathbf{x}}{\partial t} \Big|_{\boldsymbol{\chi}}. \quad (4.7)$$

Lorsque le domaine arbitraire est discrétisé, cette vitesse représente la vitesse des nœuds du maillage. La relation entre les dérivées temporelles dans le domaine arbitraire et spatial s'obtient en dérivant la relation  $\phi = \Phi \circ \psi^{-1}$  par rapport à  $(\mathbf{X}, t)$ , ce qui entraîne l'égalité

$$\mathbf{c} = \mathbf{u} - \mathbf{s} = \nabla_{\mathbf{X}} \mathbf{x} \cdot \mathbf{w} \quad \text{avec} \quad \mathbf{w}(\mathbf{X}, t) = \frac{\partial \boldsymbol{\chi}}{\partial t} \Big|_{\mathbf{X}}. \quad (4.8)$$

La vitesse de convection  $\mathbf{c}$  représente dans ce cas la vitesse relative entre la vitesse des points matériels et celle des points du domaine arbitraire qui correspond à la vitesse de déplacement du maillage. Si le domaine est fixe, la vitesse du maillage est nulle ( $\mathbf{s} = \mathbf{0}$ ) et par conséquent la vitesse de convection est égale à celle des points matériels ( $\mathbf{c} = \mathbf{u}$ ) : on retrouve alors la formulation eulérienne classique. Au contraire, si le domaine arbitraire bouge en suivant les particules, alors la vitesse de maillage est égale à celle des points matériels ( $\mathbf{s} = \mathbf{u}$ ) si bien que la vitesse de convection est nulle ( $\mathbf{c} = \mathbf{0}$ ) : on retrouve alors la formulation lagrangienne. La formulation ALE permet donc de passer d'une formulation à l'autre en fonction de la vitesse attribuée aux nœuds du maillage.

Finalement, des relations analogues à (4.4) et (4.5) sont établies pour relier la dérivée temporelle d'une quantité exprimée dans le domaine arbitraire  $\Omega_A$  à celle exprimée dans le domaine spatial  $\Omega_C$ . Lorsqu'il s'agit d'une grandeur  $f$  locale la relation (4.4) devient

$$\frac{\partial f(\boldsymbol{\chi}, t)}{\partial t} \Big|_{\boldsymbol{\chi}} = \frac{\partial f(\mathbf{x}, t)}{\partial t} \Big|_{\mathbf{x}} + \mathbf{s} \cdot \nabla_{\mathbf{x}} f(\mathbf{x}, t), \quad (4.9)$$

alors que dans le cas d'une grandeur intégrale, le théorème de Reynolds (4.5) se transforme en

$$\frac{\partial}{\partial t} \Big|_{\mathbf{X}} \int_{\Omega_M(t)} f(\mathbf{x}, t) d\Omega = \int_{\Omega_A(t)} \frac{\partial f(\mathbf{x}, t)}{\partial t} \Big|_{\mathbf{x}} d\Omega + \oint_{\partial\Omega_A(t)} f(\mathbf{x}, t) \mathbf{s} \cdot \mathbf{n} d\partial\Omega. \quad (4.10)$$

#### 4.2.2.2 Bilans de conservation des grandeurs physiques

Le système dynamique qui caractérise le comportement d'un fluide provient de l'écriture d'un certain nombre de bilans de conservation qui décrivent l'évolution de différentes grandeurs physiques et qui peuvent être résumés sous la forme générique  $dF(\mathbf{X}, t)/dt = Q_F(\mathbf{X}, t)$ , où la dérivée temporelle est la dérivée matérielle puisque le bilan est écrit pour une certaine quantité de matière. La grandeur physique  $F(\mathbf{X}, t)$  est en général associée à une grandeur  $f(\mathbf{x}, t)$  telle que  $F(\mathbf{X}, t) = \int_{\Omega} \rho f(\mathbf{x}, t) d\Omega$  où  $\rho$  désigne la masse volumique du fluide. Ainsi, le bilan générique s'exprime sous la forme *intégrale*

$$\int_{\Omega} \frac{\partial \rho f}{\partial t} d\Omega + \oint_{\partial\Omega} \rho f \mathbf{u} \cdot \mathbf{n} d\partial\Omega = Q_F \quad (4.11)$$

écrite sur le domaine spatial avec la description eulérienne en utilisant le théorème de Reynolds (4.5) et en omettant les notations relatives au domaine pour plus de clarté. On considère alors d'une ma-

nière générale que le terme source  $Q_F$  provient de contributions volumiques et surfaciques telles que  $Q_F = \int_{\Omega} q_{F,v} d\Omega + \oint_{\partial\Omega} \mathbf{q}_{F,s} \cdot \mathbf{n} d\partial\Omega$ . Le théorème de Green-Ostrogradski (D.28) permet alors de transformer les intégrales surfaciques en intégrales volumiques de sorte que le bilan soit formulé comme l'intégrale d'une somme de termes qui doit être nulle. On déduit alors le bilan *local* eulérien

$$\frac{\partial \rho f}{\partial t} + \operatorname{div}(\rho f \mathbf{u}) = q_{F,v} + \operatorname{div}(\mathbf{q}_{F,s}) \quad (4.12)$$

puisque la relation reste valable pour un domaine  $\Omega$  infiniment petit. Pour décrire complètement le comportement du fluide, on écrit en général un bilan pour la conservation de la masse, un autre pour la conservation de la quantité de mouvement et enfin un dernier pour la conservation de l'énergie. Les détails relatifs à l'obtention de ces bilans peuvent être consultés dans les ouvrages de Batchelor [25] et Ferziger et Perić [76] par exemple. En l'absence de terme source ( $Q_F = 0$ ), le bilan de conservation de la masse s'écrit pour  $f = 1$  sous la forme intégrale

$$\int_{\Omega} \frac{\partial \rho}{\partial t} d\Omega + \oint_{\partial\Omega} \rho \mathbf{u} \cdot \mathbf{n} d\partial\Omega = 0, \quad (4.13)$$

et de manière équivalente sous la forme locale

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{u}) = 0. \quad (4.14)$$

La conservation de la quantité de mouvement s'écrit pour le vecteur des composantes de la vitesse eulérienne  $f = \mathbf{u}$ . Dans ce cas, les écritures (4.11) et (4.12) doivent être adaptées puisqu'il s'agit d'une quantité vectorielle : le terme  $\rho f \mathbf{u} \cdot \mathbf{n}$  se transforme alors en  $\rho \mathbf{u} \otimes \mathbf{u} \cdot \mathbf{n}$ . Par ailleurs, le terme source volumique est un vecteur qui représente une force volumique  $\mathbf{q}_{F,v} = \rho \mathbf{f}_{\text{vol}}$ , comme la gravité par exemple ( $\mathbf{f}_{\text{vol}} = \mathbf{g}$ ), tandis que le terme surfacique provient du tenseur des contraintes de Cauchy  $\boldsymbol{\sigma} \in \mathcal{M}(d, d)$  en lieu et place de  $\mathbf{q}_{F,s}$ . Le bilan intégral de conservation de la quantité de mouvement prend donc la forme

$$\int_{\Omega} \frac{\partial \rho \mathbf{u}}{\partial t} d\Omega + \oint_{\partial\Omega} \rho \mathbf{u} \otimes \mathbf{u} \cdot \mathbf{n} d\partial\Omega = \int_{\Omega} \rho \mathbf{f}_{\text{vol}} d\Omega + \oint_{\partial\Omega} \boldsymbol{\sigma} \cdot \mathbf{n} d\partial\Omega, \quad (4.15)$$

qui devient localement

$$\frac{\partial \rho \mathbf{u}}{\partial t} + \operatorname{div}(\rho \mathbf{u} \otimes \mathbf{u}) = \rho \mathbf{f}_{\text{vol}} + \operatorname{div} \boldsymbol{\sigma}. \quad (4.16)$$

Enfin, le bilan de conservation de l'énergie peut être établi à partir du premier principe de la thermodynamique. Dans ce cas le bilan porte sur l'énergie interne volumique  $f = e_I$ . Le terme source volumique correspond au travail des forces volumiques et surfaciques  $\rho \mathbf{u} \cdot \mathbf{f}_{\text{vol}}$  et  $\operatorname{div}(\boldsymbol{\sigma} \mathbf{u})$  respectivement, mais une partie de ce travail doit être soustrait puisqu'il participe à l'énergie cinétique et non à l'énergie interne. Le terme subsistant est alors  $\boldsymbol{\sigma} : \nabla \mathbf{u}$  [25]. Par ailleurs le terme surfacique dans le bilan provient du flux de chaleur sur la frontière qui découle de la loi empirique de Fourier

$$\mathbf{q}_{\theta} = -k_{\theta} \nabla \theta, \quad (4.17)$$

où  $k_{\theta}$  est la constante de conductivité thermique et  $\theta$  désigne la température. Le terme source surfacique est alors  $\mathbf{q}_{F,s} = -\mathbf{q}_{\theta}$  pour être cohérent avec l'orientation de la normale vers l'extérieur du domaine. Finalement le bilan de conservation de l'énergie interne s'écrit sous forme intégrale et locale :

$$\int_{\Omega} \frac{\partial \rho e_I}{\partial t} d\Omega + \oint_{\partial\Omega} \rho e_I \mathbf{u} \cdot \mathbf{n} d\partial\Omega = \int_{\Omega} \boldsymbol{\sigma} : \nabla \mathbf{u} d\Omega - \oint_{\partial\Omega} \mathbf{q}_{\theta} \cdot \mathbf{n} d\partial\Omega, \quad (4.18)$$

$$\frac{\partial \rho e_I}{\partial t} + \operatorname{div}(\rho e_I \mathbf{u}) = \boldsymbol{\sigma} : \nabla \mathbf{u} - \operatorname{div} \mathbf{q}_{\theta}. \quad (4.19)$$

Le bilan précédent est souvent reformulé avec une grandeur physique différente de l'énergie interne. Ainsi, l'énergie totale  $e_T = e_I + e_C$  définie comme la somme de l'énergie interne et de l'énergie cinétique est souvent employée. Le bilan tient alors compte de toutes les contributions liées au travail des forces et s'écrit sous forme intégrale et locale :

$$\int_{\Omega} \frac{\partial \rho e_T}{\partial t} d\Omega + \oint_{\partial\Omega} \rho e_T \mathbf{u} \cdot \mathbf{n} d\partial\Omega = \int_{\Omega} \rho \mathbf{u} \cdot \mathbf{f}_{\text{vol}} d\Omega + \oint_{\partial\Omega} (\boldsymbol{\sigma} \mathbf{u} - \mathbf{q}_{\theta}) \cdot \mathbf{n} d\partial\Omega, \quad (4.20)$$



$$\frac{\partial \rho e_T}{\partial t} + \operatorname{div}(\rho e_T \mathbf{u}) = \rho \mathbf{u} \cdot \mathbf{f}_{\text{vol}} + \operatorname{div}(\boldsymbol{\sigma} \mathbf{u} - \mathbf{q}_\theta). \quad (4.21)$$

D'autres variables comme par exemple l'enthalpie, l'entropie, la pression ou encore la température sont parfois employées pour exprimer un bilan équivalent à celui de l'énergie [25, 76]. Les équations locales<sup>2</sup> de bilan (4.14), (4.16) et (4.19) ou (4.21) doivent encore être complétées et détaillées pour que le système soit fermé, *i. e.* que toutes les variables soient explicitement liées entre elles. Les variables  $\mathbf{q} = [\rho, \rho \mathbf{u}, \rho e_T]^\top$  portent habituellement le nom de *variables conservatives* puisqu'elles satisfont les équations des bilans de conservation.

#### 4.2.2.3 Loi de comportement du fluide et équation d'état

Les équations de bilan précédentes font intervenir en plus des variables conservatives le tenseur des contraintes de Cauchy  $\boldsymbol{\sigma}$  et la température  $\theta$  qui sont des inconnues supplémentaires<sup>3</sup>. Dans le cas général d'un fluide newtonien homogène, le tenseur des contraintes de Cauchy qui fournit la loi de comportement se décompose en une partie isotropique  $-p\mathbf{I}_d$  et une partie déviatorique  $\boldsymbol{\tau}$  telles que

$$\boldsymbol{\sigma} = -p\mathbf{I}_d + \boldsymbol{\tau}. \quad (4.22)$$

La partie isotropique fait intervenir la pression définie par  $p = -1/3\operatorname{Tr}(\boldsymbol{\sigma})$  qui peut être différente de la pression à l'équilibre thermodynamique notée  $p_e$ . La relation entre les deux pressions est donnée par  $p - p_e = -\kappa \operatorname{div} \mathbf{u}$ , où  $\kappa$  est la viscosité de dilatation. L'hypothèse de Stokes couramment effectuée consiste à supposer que les deux pressions sont égales ( $\kappa = 0$ ). Batchelor [25] explique alors que le tenseur déviatorique se simplifie en

$$\boldsymbol{\tau} = 2\mu \mathbf{D} + \lambda (\operatorname{div} \mathbf{u}) \mathbf{I}_d \quad \text{avec} \quad \lambda = -\frac{2}{3}\mu. \quad (4.23)$$

Les coefficients  $\mu$  et  $\lambda$  représentent respectivement le premier et le second coefficient de viscosité. En pratique, l'hypothèse de Stokes permet de considérer seulement  $\mu$  puisque dans ce cas  $\lambda = -2/3 \mu$ . La détermination précise de la valeur de la viscosité dépend de plusieurs paramètres, et notamment du régime — laminaire ou turbulent — de l'écoulement. Elle dépend par ailleurs de la température à laquelle elle est reliée via la loi de Sutherland  $\mu(\theta) = \beta_S \theta^{3/2}/(\theta + \alpha_S)$ , où  $\alpha_S$  et  $\beta_S$  sont des constantes égales à  $\alpha_S = 110,4$  et  $\beta_S = 1,45 \times 10^{-6}$  pour représenter les variations de viscosité de l'air sur une plage de température s'étalant environ de 170 K à 1900 K. Lorsque l'écoulement est turbulent, la viscosité dépend alors de manière complexe des variables conservatives et parfois de grandeurs annexes qui nécessitent d'introduire des équations de bilan supplémentaires pour fermer le système dynamique.

Le tenseur  $\mathbf{D}$  correspond au tenseur des taux de déformations introduit en décomposant le gradient des vitesses en une partie symétrique et une partie anti-symétrique telles que  $\nabla \mathbf{u} = \mathbf{D} + \boldsymbol{\Omega}$  avec

$$\mathbf{D} = \frac{1}{2}(\nabla \mathbf{u} + \nabla^\top \mathbf{u}) \quad \text{et} \quad \boldsymbol{\Omega} = \frac{1}{2}(\nabla \mathbf{u} - \nabla^\top \mathbf{u}). \quad (4.24)$$

Les équations de bilan (4.14), (4.16) et (4.21) ne sont toujours pas fermées puisque la température et la pression qui vient d'être introduite ne sont pas encore exprimées en fonction du reste des variables conservatives. Sous l'hypothèse d'un gaz parfait, l'équation d'état qui relie les grandeurs thermodynamiques  $p$ ,  $\rho$  et  $\theta$  entre elles est donnée par

$$p = \rho R_s \theta, \quad (4.25)$$

où  $R_s = R/M$  est la constante spécifique des gaz parfaits définie à partir de la constante universelle des gaz parfaits  $R = 8,314 \text{ J K}^{-1} \text{ mol}^{-1}$  et de la masse molaire moléculaire  $M$  du gaz. La constante spécifique de l'air vaut  $R_{s,\text{air}} = 287,06 \text{ J K}^{-1} \text{ mol}^{-1}$ . Cette constante vérifie la loi de Carnot  $c_p - c_v = R_s$  qui porte sur les coefficients de chaleur spécifique à volume et pression constants notés respectivement  $c_v$  et  $c_p$ . L'énergie interne est donc définie par la relation  $e_I = c_v \theta$  pour un gaz parfait et par conséquent l'équation d'état (4.25) se réécrit

$$p = (\gamma - 1)\rho e_I, \quad (4.26)$$

<sup>2</sup>Il en va de même pour les équations intégrales de bilan (4.13), (4.15) et (4.18) ou (4.20).

<sup>3</sup>En général, la force volumique  $\mathbf{f}_{\text{vol}}$  est explicitement connue. Elle ne dépend pas des variables conservatives si elle représente les effets de la gravité, mais dans les autres situations, l'expression de la force doit être explicitée pour fermer le problème.

en introduisant le rapport  $\gamma = c_p/c_v$  des coefficients de chaleur spécifique. Finalement, les bilans de conservation donnés dans le paragraphe précédent complétés par les expressions (4.22), (4.23) et (4.24) définissant le tenseur des contraintes de Cauchy puis l'équation d'état (4.25) ou (4.26) forment un système dynamique fermé dont la résolution permet de prédire le comportement non-linéaire d'un écoulement compressible et visqueux pour un fluide newtonien homogène.

### 4.2.3 Équations de Navier-Stokes en formulation ALE avec les variables conservatives

#### 4.2.3.1 Expression des équations dans le repère absolu

Les équations sont la plupart du temps exprimées dans un repère absolu  $\mathcal{R}_A$  muni d'une base  $\mathcal{B}_A$  qui est invariante au cours du temps. Les quantités vectorielles telles que la vitesse ou le tenseur des contraintes de Cauchy sont donc exprimées dans cette base de sorte qu'on note  $\mathbf{u}_A = \text{Mat}_{\mathcal{B}_A}(\mathbf{u})$  le vecteur des composantes de la vitesse exprimé dans la base absolue, ainsi que  $\boldsymbol{\sigma}_A$  pour le tenseur des contraintes de Cauchy. Par ailleurs, les opérateurs différentiels opèrent suivant les directions  $x_{A,i} = (x_A, y_A, z_A)$  associées à la base  $\mathcal{B}_A$  et on notera respectivement  $\text{div}_A$  et  $\nabla_A$  les opérateurs de divergence et de gradient par rapport aux directions de la base absolue.

La forme intégrale des équations de Navier-Stokes écrites avec les variables conservatives s'obtient à partir des bilans eulériens (4.13), (4.15) et (4.20). Ces bilans sont transportés dans le domaine arbitraire noté ici simplement  $\Omega = \Omega_A(t)$  au moyen de la relation (4.10) pour établir la formulation ALE des équations dans le repère absolu  $\mathcal{R}_A$  qui est donnée ci-dessous :

$$\begin{cases} \frac{\partial}{\partial t} \Big|_{\mathbf{x}} \int_{\Omega} \rho \, d\Omega + \oint_{\partial\Omega} \rho (\mathbf{u}_A - \mathbf{s}_A) \cdot \mathbf{n}_A \, d\partial\Omega = 0 \\ \frac{\partial}{\partial t} \Big|_{\mathbf{x}} \int_{\Omega} \rho \mathbf{u}_A \, d\Omega + \oint_{\partial\Omega} \rho \mathbf{u}_A \otimes (\mathbf{u}_A - \mathbf{s}_A) \mathbf{n}_A \, d\partial\Omega = \int_{\Omega} \rho \mathbf{f}_{\text{vol},A} \, d\Omega + \oint_{\partial\Omega} \boldsymbol{\sigma}_A \mathbf{n}_A \, d\partial\Omega \\ \frac{\partial}{\partial t} \Big|_{\mathbf{x}} \int_{\Omega} \rho e_T \, d\Omega + \oint_{\partial\Omega} \rho e_T (\mathbf{u}_A - \mathbf{s}_A) \cdot \mathbf{n}_A \, d\partial\Omega = \int_{\Omega} \rho \mathbf{f}_{\text{vol},A} \cdot \mathbf{u}_A \, d\Omega + \oint_{\partial\Omega} (\boldsymbol{\sigma}_A \mathbf{u}_A - \text{div} \mathbf{q}_{\theta,A}) \cdot \mathbf{n}_A \, d\partial\Omega. \end{cases} \quad (4.27)$$

Les équations précédentes sont écrites sous forme conservative et font apparaître des termes surfaciques dits de *flux*. La forme du système dynamique précédent se prête bien à la résolution numérique, notamment lorsqu'une discrétisation en Volumes Finis est envisagée. Toutefois, pour la construction du modèle d'ordre réduit il est préférable de disposer d'équations continues locales. Pour cela, les bilans locaux eulériens (4.14), (4.16) et (4.21) sont transformés avec la relation (4.9) et les termes sont remaniés au moyen des formules données dans l'annexe D.3 pour aboutir au système suivant :

$$\begin{cases} \frac{\partial \rho}{\partial t} \Big|_{\mathbf{x}} + (\mathbf{u}_A - \mathbf{s}_A) \cdot \nabla_A \rho = -\rho \text{div}_A \mathbf{u}_A \\ \rho \frac{\partial \mathbf{u}_A}{\partial t} \Big|_{\mathbf{x}} + \rho (\mathbf{u}_A - \mathbf{s}_A) \cdot \nabla_A \mathbf{u}_A = \text{div}_A \boldsymbol{\sigma}_A + \rho \mathbf{f}_{\text{vol},A} \\ \rho \frac{\partial e_T}{\partial t} \Big|_{\mathbf{x}} + \rho (\mathbf{u}_A - \mathbf{s}_A) \cdot \nabla_A e_T = \text{div}_A (\boldsymbol{\sigma}_A \mathbf{u}_A - \mathbf{q}_{\theta,A}) + \rho \mathbf{f}_{\text{vol},A} \cdot \mathbf{u}_A. \end{cases} \quad (4.28)$$

La formulation eulérienne des équations de Navier-Stokes s'obtient à partir des équations intégrales ou locales en posant simplement  $\mathbf{s}_A = \mathbf{0}$  puisque dans ce cas le domaine arbitraire est supposé invariant et confondu avec le domaine spatial de contrôle. On désignera la formulation ALE des équations dans le repère absolu par FALE-RAUA et la formulation eulérienne par FEUL-RAUA pour les différencier des autres formulations développées dans le paragraphe suivant. Dans la suite, on négligera pour simplifier le terme de force volumique  $\mathbf{f}_{\text{vol}}$  puisque aucun effet particulier n'a besoin d'être pris en compte par ce type de terme.

Le système dynamique obtenu est non-linéaire en raison de l'opérateur d'advection, ainsi que des dépendances des variables entre elles dans l'équation d'état et la loi de Sutherland gouvernant la viscosité. La construction du modèle d'ordre réduit nécessitera d'explicitier l'ensemble des expressions afin de déterminer le type de non-linéarités rencontrées.

#### 4.2.3.2 Expression des équations dans le repère entraîné

Lorsque le système considéré est entraîné dans un mouvement de corps rigide, les équations sont de préférence formulées dans le repère entraîné  $\mathcal{R}_E$  associé au mouvement. L'écriture des équations de Navier-Stokes dans un repère entraîné peut s'avérer judicieuse dans certains cas puisque l'écoulement peut devenir permanent dans le repère entraîné alors qu'il est instationnaire dans le repère absolu. C'est le cas notamment pour une roue isolée de turbomachine, pour un rotor d'hélicoptère en vol stationnaire ou pour une hélice en mouvement de translation axiale par exemple. Les simulations numériques de tels écoulements bénéficient alors des techniques d'accélération de la convergence développées pour les problèmes stationnaires.

Le changement de repère conduit à une nouvelle forme des équations dans laquelle des termes sources supplémentaires apparaissent pour modéliser les effets du mouvement du domaine relatif. Les équations peuvent alors être formulées au moyen des composantes de la vitesse *absolue* ou des composantes de la vitesse *relative*. Ces formulations sont désignées respectivement par FALE-REUE et FALE-REVE lorsque la description cinématique ALE est adoptée. Le principe d'obtention de ces équations est brièvement présenté ici mais une justification plus rigoureuse est donnée dans l'annexe E.1.

L'écriture des équations dans le repère entraîné nécessite de définir un certain nombre de relations de changement de base pour les quantités vectorielles, ainsi que pour les opérateurs de différentiation spatiale et temporelle. La position absolue d'un point  $M$  de coordonnées  $x$  est décomposée par la relation de Chasles en  $x = x_0 + \tilde{x}$  où  $x_0$  représente la position absolue de l'origine du repère entraîné et  $\tilde{x}$  est la position relative. Si on note  $\mathbf{R} = \mathbf{P}_{\mathcal{B}_A \rightarrow \mathcal{B}_E}$  la matrice de passage entre les bases  $\mathcal{B}_E$  et  $\mathcal{B}_A$  associées aux repères absolu et entraîné respectivement, la position du point  $M$  s'écrit  $\mathbf{x}_A = \mathbf{R} \mathbf{x}_E$ , ou encore

$$\mathbf{x}_A = \mathbf{x}_{0,A} + \tilde{\mathbf{x}}_A = \mathbf{x}_{0,A} + \mathbf{R} \tilde{\mathbf{x}}_E, \quad (4.29)$$

où  $\tilde{\mathbf{x}}_E$  est la position relative exprimée dans la base du repère entraîné. Si on considère que le changement de base correspond seulement à un ensemble de rotations, la matrice  $\mathbf{R}$  est orthogonale et unitaire ( $\mathbf{R}^{-1} = \mathbf{R}^\top$ ), ce qui signifie au passage que la relation de changement de base de  $\mathcal{B}_A$  vers  $\mathcal{B}_E$  s'écrit  $\mathbf{x}_E = \mathbf{R}^\top \mathbf{x}_A$ . On associe à  $\mathbf{R}$  la matrice de la vitesse angulaire  $\boldsymbol{\Omega}_E = \mathbf{R}^\top d\mathbf{R}/dt$  qui est antisymétrique. Le vecteur vitesse angulaire  $\boldsymbol{\omega}_E$  est alors défini tel que pour tout vecteur  $\mathbf{r}_E$ ,  $\boldsymbol{\Omega}_E \mathbf{r}_E = \boldsymbol{\omega}_E \wedge \mathbf{r}_E$ .

La vitesse absolue des points du domaine arbitraire est alors exprimée comme la dérivée de la position absolue (4.29) qui correspond à  $\mathbf{s}_A(M) = d\mathbf{x}_A/dt$ . Cependant, puisque les équations doivent être formulées dans le repère entraîné, on écrit de préférence  $\mathbf{s}_E = \mathbf{R}^\top \mathbf{s}_A$  en projetant dans la base  $\mathcal{B}_E$ . Après réorganisation des termes, on parvient à l'expression

$$\mathbf{s}_E(M) = \mathbf{s}_{e,E}(M) + \mathbf{s}_{d,E}(M) \quad \text{avec} \quad \begin{cases} \mathbf{s}_{e,E}(M) = \mathbf{s}_{0,E} + \boldsymbol{\omega}_E \wedge \tilde{\mathbf{x}}_E \\ \mathbf{s}_{d,E}(M) = \frac{d\tilde{\mathbf{x}}_E}{dt}. \end{cases} \quad (4.30)$$

La relation précédente fait intervenir la vitesse d'entraînement  $\mathbf{s}_{e,E}$  et la vitesse de déformation  $\mathbf{s}_{d,E}$  exprimées dans la base  $\mathcal{B}_E$ . La vitesse  $\mathbf{s}_{0,E}$  correspond à la vitesse d'entraînement en translation qui est définie par  $\mathbf{R}^\top d(\mathbf{R}\mathbf{x}_{0,E})/dt$  alors que le terme  $\boldsymbol{\omega}_E \wedge \tilde{\mathbf{x}}_E$  correspond à la vitesse d'entraînement en rotation.

En plus de ces relations, les opérateurs de différentiation spatiale sont modifiés par les relations de changement de base analogues à (A.9) et (A.11) lorsque la quantité résultant de l'action de l'opérateur est un vecteur ou une matrice.

Les équations intégrales de bilan (4.27) sont donc transformées en exprimant les quantités vectorielles ainsi que les opérateurs de différentiation spatiale dans la base  $\mathcal{B}_E$ . Le bilan de quantité de mouvement est en outre projeté sur la base  $\mathcal{B}_E$  puisqu'il s'agit d'une équation vectorielle. Après manipulation des termes (voir l'annexe E.1.2), on parvient aux équations suivantes données sous forme intégrale :

$$\begin{cases} \left. \frac{\partial}{\partial t} \right|_{\mathbf{x}} \int_{\Omega} \rho d\Omega + \oint_{\partial\Omega} \rho (\mathbf{u}_E - \mathbf{s}_E) \cdot \mathbf{n}_E d\partial\Omega = 0 \\ \left. \frac{\partial}{\partial t} \right|_{\mathbf{x}} \int_{\Omega} \rho \mathbf{u}_E d\Omega + \oint_{\partial\Omega} [\rho \mathbf{u}_E \otimes (\mathbf{u}_E - \mathbf{s}_E)] \mathbf{n}_E d\partial\Omega = \oint_{\partial\Omega} \boldsymbol{\sigma}(\mathbf{u}_E) \mathbf{n}_E d\partial\Omega - \boldsymbol{\omega}_E \wedge \int_{\Omega} \rho \mathbf{u}_E d\Omega \\ \left. \frac{\partial}{\partial t} \right|_{\mathbf{x}} \int_{\Omega} \rho e_T d\Omega + \oint_{\partial\Omega} \rho e_T (\mathbf{u}_E - \mathbf{s}_E) \cdot \mathbf{n}_E d\partial\Omega = \oint_{\partial\Omega} [\boldsymbol{\sigma}(\mathbf{u}_E) \mathbf{u}_E - \mathbf{q}_{\theta,E}] \cdot \mathbf{n}_E d\partial\Omega. \end{cases} \quad (4.31)$$



### 4.3 Problématique de l'aéroélasticité et méthodologie de résolution

#### 4.3.1 Principe de l'aéroélasticité

L'aéroélasticité statique concerne l'étude du couplage entre les forces aérodynamiques et élastiques. Lorsque les effets dynamiques entrent en ligne de compte, les forces inertielles participent en plus au couplage. Les analyses aéroélastiques s'imposent lorsque les interactions mutuelles entre les différentes forces influent significativement sur la réponse du système. Elles se justifient par exemple lorsque les vibrations d'une structure *in vacuum* sont amorties alors que la prise en compte du fluide peut conduire à des vibrations entretenues ou — dans le pire des cas — à des vibrations dont l'amplitude augmente jusqu'à un seuil critique pour l'intégrité de la structure. Dans ce cas, les échanges d'énergie entre les deux milieux, fluide et structure, doivent être modélisés afin de pouvoir analyser différents types de phénomènes aéroélastiques. Dowell *et al.* [58] ont notamment distingué le phénomène de *divergence* en aéroélasticité statique, le phénomène de *flottement* qui décrit la stabilité dynamique du système aéroélastique en réponse à une petite perturbation ou encore la *réponse dynamique* qui correspond à la réponse temporelle du système à une excitation externe telle qu'une rafale, un contrôle actif,... Seuls les deux derniers types d'analyses sont instationnaires et correspondent au cadre de cette étude.

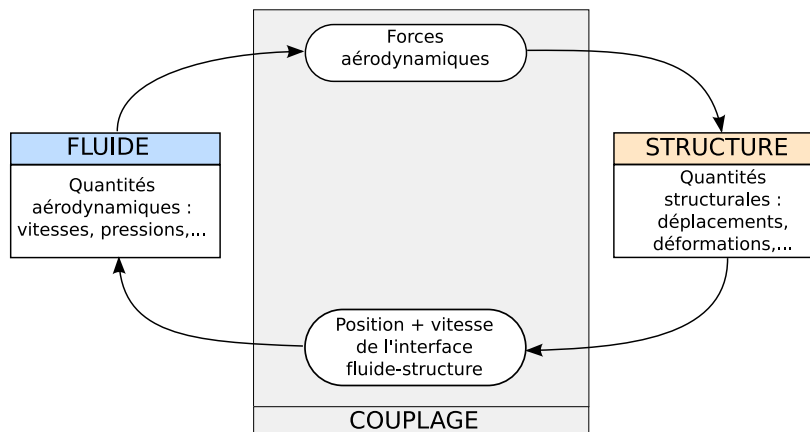


Figure 4.3 : Schéma de principe du cycle d'interaction entre le fluide et la structure.

Le principe du couplage peut être représenté par le cycle d'interaction de la figure 4.3 dont les étapes sont les suivantes :

1. les forces aérodynamiques exercent une charge sur la structure qui se déplace et/ou se déforme ;
2. la résolution du problème structural fournit les déplacements et/ou les déformations de la structure et en particulier ceux de la frontière entre le fluide et la structure en réponse aux efforts ;
3. la modification de la position et de la vitesse de l'interface fluide-structure redéfinit donc le domaine fluide dans lequel évolue le champ aérodynamique ;
4. la résolution du problème fluide pour la nouvelle topologie du domaine fournit alors le nouveau champ aérodynamique à partir duquel la force aérodynamique est calculée, et ainsi de suite.

La description précédente du cycle d'interaction indique que l'état d'équilibre du système est obtenu — à chaque instant pour la réponse dynamique — lorsque les forces aérodynamiques sont en adéquation avec la bonne position de l'interface entre le fluide et la structure. La résolution du cycle d'interaction en vue de déterminer cet équilibre fait l'objet d'un certain nombre de choix concernant notamment la manière de traiter les équations du système couplé (approche monolithique ou couplée) et de tenir compte du mouvement de la structure (couplage fort ou faible).

## 4.3.2 Formulations du problème de couplage

### 4.3.2.1 Approche monolithique vs approche couplée

En fonction des hypothèses émises pour chacun des deux milieux, les équations se prêtent plus ou moins bien à la formation d'un unique système couplé fluide-structure. Lorsque cela est possible, on parle d'une approche *monolithique* puisque les équations des deux milieux sont résolues simultanément. Dans le cas contraire, on parlera d'une approche *couplée* puisque les équations des deux milieux sont résolues successivement l'une après l'autre avec un certain décalage temporel.

L'approche monolithique est par exemple adéquate lorsqu'on s'intéresse à des problèmes de ballonnement d'un liquide dans une structure déformable ou pour des études de vibroacoustique dans une cavité. Le fluide, alors supposé incompressible, parfait et homogène, est modélisé par les équations d'un fluide potentiel qui sont linéarisées pour de petites perturbations. Il est ainsi possible de définir des modes de ballonnement en basse fréquence, ou encore des modes acoustiques en haute fréquence. La résolution du système fluide peut donc être conduite simultanément avec celle de la structure en utilisant par exemple un formalisme Éléments Finis. En outre, une approche modale peut être adoptée pour réduire la taille du système couplé. Pour plus de détails sur ces méthodes, le lecteur pourra consulter en particulier les travaux de Morand et Ohayon [168] ainsi que [204, 211]. Une extension de ces travaux pour des applications au ballonnement dans une structure déformable en microgravité est en cours de développement [67].

Lorsque le fluide incompressible est modélisé par les équations non-linéaires de Navier-Stokes, la méthode des domaines fictifs [88, 89, 176] peut être employée de manière à conserver une approche monolithique. En effet, une unique formulation variationnelle du problème couplé est obtenue par extension des équations du domaine fluide à la structure. Pour cela, la contrainte de rigidité sur le domaine fictif est imposée au moyen d'un coefficient de pénalisation.

L'approche monolithique réalise le couplage de façon intrinsèque, si bien que les schémas numériques sont en général plus stables et plus précis [166]. Toutefois, cette approche n'est pas adaptée pour formuler les équations du problème couplé aéroélastique qui motive cette étude. Kamakoti et Shyy [125] ont en effet relevé très peu de travaux basés sur une approche monolithique pour l'aéroélasticité, et ont signalé qu'il existait encore beaucoup de restrictions à leur emploi. Les principales raisons viennent de la différence de modélisation des deux milieux, de l'ordre de grandeur des pas de temps caractéristiques et des descriptions cinématiques employées. Par conséquent la résolution pour des systèmes de grande taille est pratiquement impossible, à moins de recourir à des méthodes de décomposition de domaine [102].

Afin de pouvoir jouir de certains développements spécifiques<sup>6</sup> à chacun des deux milieux, il est donc préférable de traiter le problème par une approche couplée avec des codes de calcul dédiés. Intuitivement, le principe revient à reproduire une à une les étapes du schéma de la figure 4.3. Si l'algorithme de couplage est explicite, le cycle d'interaction est calculé une seule fois par pas de temps. Au contraire, si l'algorithme est implicite, plusieurs itérations sur le cycle d'interaction sont effectuées pour le pas de temps en question, jusqu'à la convergence de l'état d'équilibre du système couplé. L'approche couplée nécessite donc de disposer d'un code de calcul pour le milieu fluide et d'un second code pour la structure, ainsi que d'un algorithme de couplage permettant de gérer le transfert d'informations entre les deux milieux (procédure de transfert des efforts et des déplacements, sous-itérations dans le cas d'un couplage implicite,...).

Les procédures décalées ont ainsi permis d'obtenir une description précise de phénomènes complexes tels que l'écoulement transsonique autour d'une aile 3D [71, 72], le comportement d'absorbeurs de choc hydrauliques [142], la réponse transitoire à des explosions [39] ou encore les phénomènes de vibrations induites par un détachement de vortex dans un écoulement incompressible [90, 179, 202]. De telles simulations sont progressivement employées dans l'industrie aérospatiale pour tenter de prédire les limites de stabilité des systèmes mécaniques. De nombreux développements ont ainsi été menés à l'Onera afin de prédire le comportement aéroélastique de voilures ou de fuselages d'avions, ou encore dans le but de déterminer la réponse aéroélastique des roues aubagées rencontrées dans les turbomachines [64, 65, 85, 86, 206].

Les méthodes couplées restent donc les plus appropriées pour traiter les problèmes d'aéroélasticité. La difficulté consiste toutefois à préserver la conservation des échanges d'énergie entre les deux milieux sans laquelle une production artificielle d'énergie à l'interface peut conduire à la divergence. Pour limiter ce risque, les simulations deviennent souvent très coûteuses en temps de calcul puisqu'un petit pas de temps est requis avec les algorithmes explicites et les algorithmes implicites sont itératifs.

<sup>6</sup>Par exemple les lois de comportement, la turbulence pour le fluide, ou encore les algorithmes de résolution.

### 4.3.2 Couplage fort vs couplage faible

Indépendamment de l'approche choisie pour formuler les équations, il existe deux manières de traiter le mouvement de la structure. On parle de *couplage fort* lorsque le cycle d'interaction est effectivement calculé, c'est-à-dire lorsque les vibrations de la structure sont provoquées par les charges aérodynamiques qui lui sont appliquées. Une autre technique consiste à pratiquer un *couplage faible*, c'est-à-dire un couplage dans lequel il n'y a pas de rétroaction du fluide sur la structure : le principe consiste à résoudre les équations du fluide pour certains mouvements imposés à la structure. Les forces aérodynamiques obtenues sont alors introduites a posteriori au second membre de l'équation gouvernant la structure, en supposant qu'elles dépendent des déplacements et des vitesses. La méthode est particulièrement intéressante lorsque la structure est linéaire puisqu'il est possible de pratiquer une décomposition sur la base des modes propres de manière à réduire le nombre de degrés de liberté dans le modèle structural. Les mouvements imposés à la structure lors de la résolution du problème fluide sont donc les déformées modales qui produisent les *forces aérodynamiques généralisées* — *i. e.* les forces associées à chaque mode excitateur de la structure. Sous l'hypothèse de linéarité des forces aérodynamiques généralisées en fonction des degrés de liberté structuraux, la réponse temporelle du système aéroélastique peut être déterminée à partir de l'équation structurale.

Le couplage faible est souvent employé pour les problèmes de stabilité en flottement. Dans ce cas, on suppose une petite perturbation harmonique des degrés de liberté structuraux et l'équation gouvernant la structure s'écrit sous la forme d'un problème aux valeurs propres non-linéaire puisque la matrice des forces aérodynamiques généralisées dépend des valeurs propres. Plusieurs méthodes itératives peuvent être employées pour parvenir à résoudre le problème (voir par exemple [58], paragraphe 3.6). La plus populaire d'entre elles est la méthode  $p - k$  [108] qui fournit en plus de la vitesse de flottement une estimation de l'amortissement du système couplé. La méthode s'applique aussi aux systèmes aubagés présents dans les turbomachines [104] et a été décrite en détail par Tran *et al.* [231]. La difficulté principale lors du processus itératif de la méthode  $p - k$  est l'estimation des forces aérodynamiques généralisées pour la valeur inconnue de la solution du problème aux valeurs propres du système aéroélastique. Pour cela, des méthodes d'interpolation plus ou moins élaborées [127, 183] sont employées. Le couplage faible simplifie considérablement les calculs puisque le couplage entre les deux milieux n'est pas directement effectué. Il permet des gains substantiels en temps de calcul mais nécessite d'interpoler les forces aérodynamiques généralisées lors de la résolution du problème de stabilité et repose en outre sur l'hypothèse que les forces aérodynamiques généralisées sont linéaires en fonction des degrés de liberté structuraux.

Que ce soit par le biais d'un couplage fort ou faible, la résolution du problème aéroélastique n'est pas complètement satisfaisante lorsqu'il s'agit de calculer la réponse pour un grand nombre de paramètres tels que le nombre de Mach, ou encore le déphasage interaube ou la vitesse de rotation pour les applications aux turbomachines. En effet, le couplage fort conduit à des temps de calcul prohibitifs, tandis que le couplage faible — lui aussi assez coûteux — nécessite de recourir à des procédés d'interpolation qui requièrent le réglage souvent délicat d'un certain nombre de paramètres [231]. En outre, le couplage faible repose sur l'hypothèse de linéarité des forces aérodynamiques généralisées en fonction des degrés de liberté structuraux qui n'est pas justifiée dans le cas général.

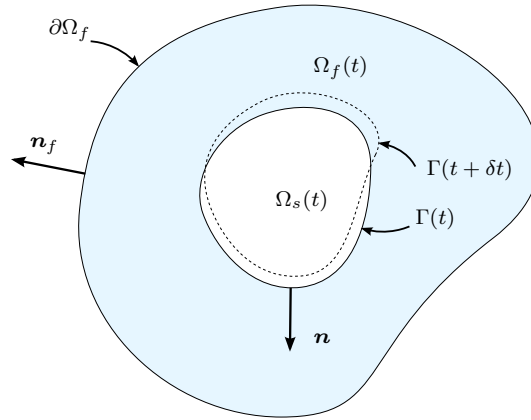
### 4.3.3 Définition du problème aéroélastique général

Soit  $\Omega \subset \mathbb{R}^d$  un domaine de l'espace de dimension  $d$  défini par la réunion disjointe d'un domaine fluide  $\Omega_f$  et d'un domaine solide  $\Omega_s$  tels que

$$\begin{aligned}\Omega &= \Omega_f \cup \Omega_s, \\ \Omega_f \cap \Omega_s &= \emptyset.\end{aligned}\tag{4.36}$$

On considère une structure présentant une interface commune avec le fluide au travers de laquelle aucun échange de matière n'a lieu. Le fluide est en outre supposé monophasique. À tout instant  $t \in I$ , on note  $\Gamma(t)$  la frontière commune entre le domaine fluide et le domaine solide, *i. e.* l'interface fluide-structure. Les autres frontières des domaines fluide et structure seront notées respectivement  $\partial\Omega_f$  et  $\partial\Omega_s$ <sup>7</sup>. La normale à l'interface fluide-structure  $\Gamma$  dirigée vers le fluide sera notée  $\mathbf{n}$ , tandis que les normales sur  $\partial\Omega_f$  et  $\partial\Omega_s$  seront indiquées respectivement par  $\mathbf{n}_f$  et  $\mathbf{n}_s$ . Le schéma de la figure 4.4 résume les différentes notations.

<sup>7</sup>Si la structure est entièrement immergée dans le fluide, la frontière du domaine solide correspond uniquement à  $\Gamma$ .



**Figure 4.4 :** Schématisation des domaines fluide et structure pour la résolution du problème couplé aéroélastique à un instant  $t$ . La frontière externe  $\partial\Omega_f$  est fixe, alors que l'interface fluide-structure  $\Gamma$  se déforme au cours du temps. Les traits pointillés représentent la position de l'interface fluide-structure à un instant ultérieur  $t + \delta t$ .

Dans l'optique d'une résolution du système aéroélastique par une approche couplée, le problème est décrit par un ensemble d'équations gouvernant le fluide dans le domaine  $\Omega_f(t)$  et le solide dans le domaine  $\Omega_s(t)$ , le tout étant complété par un ensemble de conditions de couplage sur l'interface fluide-structure  $\Gamma(t)$ .

**Modélisation du fluide** Le fluide, supposé homogène, compressible, visqueux et newtonien, est décrit par les équations de Navier-Stokes. Le domaine fluide  $\Omega_f(t)$  se déforme au cours du temps pour tenir compte du déplacement et/ou de la déformation de la structure sur l'interface  $\Gamma$  et par conséquent une formulation ALE des équations doit être adoptée. Le domaine  $\Omega_f$  correspond alors à un domaine arbitraire  $\Omega_A$  de l'espace, tel que le bord  $\partial\Omega_f$  soit fixe et que le bord  $\Gamma$  suive le mouvement de la structure.

Les systèmes d'équations (4.27), (4.31) et (4.33) établis au paragraphe 4.2.3 sont donc employés pour décrire le fluide dans le repère absolu ou relatif, avec les composantes de la vitesse absolue ou relative. Les équations sont ensuite complétées par une condition initiale portant sur le champ aérodynamique qui peut être initialisé par exemple avec le champ stationnaire. Par ailleurs, des conditions aux limites sur les frontières du domaine sont imposées : sur  $\partial\Omega_f$ , des conditions de non-réflexion sont par exemple imposées pour modéliser un domaine infini et les conditions de couplage spécifiques à l'interface fluide-structure  $\Gamma$  seront détaillées dans la suite de ce paragraphe.

**Modélisation de la structure** Dans la formulation classique d'un problème aéroélastique, la structure est décrite dans le cadre de l'élasticité linéaire pour un matériau isotrope<sup>8</sup>. La structure est supposée ne subir que de petites déformations qui sont correctement caractérisées au moyen de l'équation de bilan locale

$$\rho_s \frac{\partial^2 \mathbf{x}_s}{\partial t^2} \Big|_{\mathbf{X}} = \operatorname{div} \boldsymbol{\sigma}_s + \rho_s \mathbf{f}_{\text{vol},s} \quad (4.37)$$

gouvernant les déplacements  $\mathbf{x}_s$  en formulation lagrangienne dans  $\Omega_s(t)$ . Par ailleurs, le tenseur des déformations  $\boldsymbol{\varepsilon}_s = 1/2 (\nabla \mathbf{x}_s + \nabla^T \mathbf{x}_s)$  est relié au tenseur des contraintes de Cauchy  $\boldsymbol{\sigma}_s$  par la loi de comportement

$$\boldsymbol{\varepsilon}_s = \frac{1 + \nu_s}{E} \boldsymbol{\sigma}_s - \frac{\nu_s}{E} \operatorname{Tr}(\boldsymbol{\sigma}) \mathbf{I}_d, \quad (4.38)$$

qui nécessite la donnée du module de Young  $E$  et du coefficient de Poisson  $\nu_s$  caractéristiques du matériau employé. La densité  $\rho_s$  du solide est ici considérée comme constante dans  $\Omega_s$ . Les forces volumiques se réduisent en général à la pesanteur, si bien que  $\mathbf{f}_{\text{vol},s} = \mathbf{g}$  où  $\mathbf{g}$  est l'accélération de la pesanteur.

<sup>8</sup>Il est tout à fait possible d'employer des lois de comportement plus complexes et de dépasser ainsi le cadre de l'élasticité linéaire pour modéliser la structure, notamment lorsqu'elle est constituée par exemple d'un matériau composite.



De la même manière que pour le fluide, l'équation (4.37) est complétée par une condition initiale portant sur les déplacements de la structure ainsi que par des conditions limites. Si le solide est entièrement immergé dans le fluide, seules les conditions limites de couplage entre le fluide et la structure sur  $\Gamma$  s'appliquent. Si une partie de la structure n'est pas directement en contact avec le fluide, il est possible d'introduire d'autres conditions limites de type Dirichlet ou Neumann par exemple.

#### Remarque 4.1 : Mouvement de corps rigide

Pour simplifier le problème dans un premier temps, il est courant de faire l'hypothèse d'un corps rigide en mouvement dans le fluide. Le mouvement de l'interface n'est donc plus gouverné par la déformée élastique de la structure mais par un petit nombre de degrés de liberté n'excédant pas 6 lorsque l'espace est tridimensionnel (3 rotations et 3 translations). Dans ce cas, il ne s'agit plus au sens strict du terme d'une analyse aéroélastique puisque la structure est rigide.

Dans l'hypothèse d'un corps rigide, le domaine solide  $\Omega_s(t)$  se déduit du domaine à l'instant initial par un ensemble de translations et de rotations. La distance entre deux points distincts  $x$  et  $y$  de coordonnées  $\mathbf{x}$  et  $\mathbf{y}$  dans  $\Omega_s(t)$  reste donc constante à chaque instant, ce qui se traduit mathématiquement par

$$\forall t \in \mathbb{R}_+, \exists \mathbf{K} \in \mathbb{R}^3, \forall (x, y) \in \Omega_s(t), \quad \mathbf{v}_s(x, t) = \mathbf{v}_s(y, t) + j(\mathbf{K})(\mathbf{x} - \mathbf{y}), \quad (4.39)$$

où  $\mathbf{v}_s(x, t) = \dot{\mathbf{x}}$  est le champ de vitesse dans le domaine  $\Omega_s(t)$ . L'application  $j$  vérifie quant à elle

$$\forall \mathbf{K} = \text{Mat}_{\mathcal{B}}(\mathbf{K}) = \begin{bmatrix} K_1 \\ K_2 \\ K_3 \end{bmatrix} \in \mathbb{R}^3, \forall \mathbf{r} \in \mathbb{R}^3, \quad j(\mathbf{K})\mathbf{r} = \mathbf{K} \wedge \mathbf{r} \quad \text{avec} \quad j(\mathbf{K}) = \begin{bmatrix} 0 & -K_3 & K_2 \\ K_3 & 0 & -K_1 \\ -K_2 & K_1 & 0 \end{bmatrix}. \quad (4.40)$$

Par conséquent  $\nabla \mathbf{v}_s = j(\mathbf{K})$  et  $\nabla^\top \mathbf{v}_s = -j(\mathbf{K})$  donc le tenseur des taux de déformation pour un solide rigide est nul :  $\mathbf{D}(\mathbf{v}_s) = 1/2(\nabla \mathbf{v}_s + \nabla^\top \mathbf{v}_s) = \mathbf{0}_{d,d}$ . La réciproque est aussi vraie. Ainsi, l'équivalence suivante définit formellement un corps rigide :

$$\text{Un milieu continu est indéformable si et seulement si } \mathbf{D}(\mathbf{v}_s) = \frac{1}{2}(\nabla \mathbf{v}_s + \nabla^\top \mathbf{v}_s) = \mathbf{0}_{d,d}. \quad (4.41)$$

Puisque le solide est rigide, la position de n'importe quel point  $M$  du domaine solide de coordonnées  $\mathbf{x}_s$  se déduit de la position du centre de gravité  $G$  de coordonnées  $\mathbf{x}_G$  à l'aide de la loi de composition des vitesses (4.39) qui se transforme avec la relation (4.40) en

$$\mathbf{v}_s(M, t) = \mathbf{v}_G + \boldsymbol{\omega} \wedge (\mathbf{x}_s(M) - \mathbf{x}_G). \quad (4.42)$$

La vitesse  $\mathbf{v}_s(M, t) = \dot{\mathbf{x}}_s$  du point  $M$  — et par conséquent la position — s'exprime donc en fonction de celle du centre de gravité  $\mathbf{v}_G = \dot{\mathbf{x}}_G$  au moyen du vecteur vitesse angulaire  $\boldsymbol{\omega}$ <sup>9</sup>. Les vitesses de translation  $\mathbf{v}_G$  et de rotation  $\boldsymbol{\omega}$  sont gouvernées par le système d'équations différentielles ordinaires

$$\begin{cases} m \frac{d\mathbf{v}_G}{dt} = \mathbf{F}_a(t) \\ \mathbf{J} \frac{d\boldsymbol{\omega}}{dt} + \mathbf{J} \boldsymbol{\omega} \wedge \boldsymbol{\omega} = \mathbf{M}_a(t). \end{cases} \quad (4.43)$$

Dans les expressions précédentes,  $m$  désigne la masse du solide, et  $\mathbf{J}$  sa matrice d'inertie. Les termes du second membre représentent les résultantes de la force et du moment aérodynamiques dus à l'action du fluide sur la paroi du solide. Ils sont respectivement donnés par

$$\mathbf{F}_a(t) = \int_{\Gamma(t)} \boldsymbol{\sigma}_f \mathbf{n} \, d\Gamma, \quad (4.44a)$$

$$\mathbf{M}_a(t) = \int_{\Gamma(t)} (\mathbf{x}_s - \mathbf{x}_g) \wedge (\boldsymbol{\sigma}_f \mathbf{n}) \, d\Gamma. \quad (4.44b)$$

<sup>9</sup>Ce vecteur vitesse correspond au vecteur  $\mathbf{K}$  de la relation (4.39). Il est donc associé à l'application  $j$  dont la représentation matricielle est un tenseur antisymétrique, d'où l'interprétation en terme de rotation.

**Conditions de couplage sur l'interface fluide-structure** L'interface  $\Gamma(t)$  est le siège des transferts entre le fluide et la structure et assure donc le couplage entre les deux milieux. Ce dernier est assuré par deux types de conditions qui expriment (i) le fait qu'aucune particule ne peut traverser l'interface, et (ii) que les contraintes doivent être continues au passage de l'interface. La première condition est de type cinématique et revient à imposer la continuité des vitesses au passage de l'interface. La seconde est quant à elle dynamique et assure la continuité des contraintes normales à l'interface suivant le principe d'action-réaction. Pour un fluide visqueux, ces deux conditions s'écrivent

$$\begin{cases} \mathbf{u} = \dot{\mathbf{x}}_s \\ \boldsymbol{\sigma}_f \mathbf{n} = \boldsymbol{\sigma}_s \mathbf{n} \end{cases} \quad \text{sur } \Gamma(t). \quad (4.45)$$

Dans le cas d'un fluide parfait, seules les composantes normales de la vitesse sont couplées dans la condition cinématique car le fluide est libre de glisser sur la paroi du solide. En outre, la condition de couplage dynamique peut être simplifiée : la loi de comportement du fluide ne fait plus intervenir que la pression ( $\boldsymbol{\sigma}_f = -p \mathbf{I}_d$ ) puisque les termes visqueux sont éliminés. Les conditions de couplage sur l'interface deviennent alors

$$\begin{cases} \mathbf{u} \cdot \mathbf{n} = \dot{\mathbf{x}}_s \cdot \mathbf{n} \\ -p \mathbf{n} = \boldsymbol{\sigma}_s \mathbf{n} \end{cases} \quad \text{sur } \Gamma(t). \quad (4.46)$$

**Remarque 4.2 : Condition de paroi pour un solide fixe**

*La condition de couplage cinématique est simplement une extension de la condition de paroi habituellement employée lorsque la frontière est fixe. Dans ce cas, la vitesse du fluide est en effet supposée nulle ( $\mathbf{u} = \mathbf{0}$ ) si le fluide est visqueux ; pour un fluide parfait, on impose simplement que la vitesse normale soit nulle. Par conséquent il s'agit d'un cas particulier où  $\dot{\mathbf{x}}_s = \mathbf{0}$  pour les conditions de couplage cinématiques dans les équations (4.45) et (4.46).*

**Couplage fort et couplage faible** Si le problème aéroélastique à résoudre requiert un couplage fort, la solution est recherchée au moyen d'une approche couplée dans laquelle les équations du fluide établies au paragraphe 4.2.3 puis celles de la structure élastique (4.37) sont résolues successivement en vérifiant à chaque pas de temps les conditions de couplage (4.45) si le fluide est supposé visqueux. Si la structure est supposée rigide, les équations résolues sont simplement remplacées par (4.43). On obtient de cette manière la réponse temporelle du système aéroélastique qui est ensuite analysée, pour évaluer par exemple l'amortissement si l'objectif est de contrôler la stabilité du système couplé.

Dans le cas d'un couplage faible les équations du fluide sont résolues pour un mouvement de la structure imposé a priori. Par conséquent la position de l'interface  $\Gamma(t)$  est explicitement prescrite et ne requiert pas l'intégration en temps de l'équation (4.37) si le corps est élastique, ou des équations (4.43) si la structure est supposée rigide. Pour un solide rigide, il suffit d'imposer un mouvement au centre de gravité ainsi qu'une loi d'évolution pour la vitesse angulaire et la position de l'interface  $\Gamma$  se déduit immédiatement par la loi de composition des vitesses (4.42). Si la structure est élastique, alors on peut calculer les déformées associées aux premiers modes propres qui seront imposées avec une excitation harmonique à une certaine fréquence et avec une certaine amplitude. L'objectif est alors de calculer les forces aérodynamiques généralisées en réponse à l'excitation imposée, puis d'introduire ces forces au second membre de l'équation gouvernant la structure afin de mener par exemple une étude de stabilité.

## 4.4 Modélisation du fluide avec les variables primitives modifiées pour la construction du modèle d'ordre réduit

### 4.4.1 Variables aérodynamiques pour l'obtention d'un système quadratique

Les équations du fluide ont été établies dans les paragraphes précédents pour les variables conservatives qui sont habituellement employées lorsque la résolution est effectuée au moyen d'une discrétisation Éléments Finis ou Volumes Finis par exemple. L'utilisation de ces mêmes équations pour la construction d'un modèle d'ordre réduit non-linéaire avec l'approche POD-Galerkin décrite au paragraphe 2.4.3 pose néanmoins problème puisque le système dynamique ne s'écrit pas sous la forme d'un polynôme dont la projection permet de conserver simplement les non-linéarités. Pour s'en convaincre, il suffit d'expliquer la forme locale des équations (4.28) en faisant apparaître exclusivement les variables conservatives  $\mathbf{q} = [\rho, \rho\mathbf{u}, \rho e_T]^\top$ . Parmi les termes qui dépendent implicitement des variables, le tenseur des contraintes de Cauchy dont l'expression a été donnée à l'équation (4.22) fait intervenir la pression  $p$ . Celle-ci s'exprime en fonction de l'énergie interne  $e_I$  par la relation (4.26) sous l'hypothèse d'un gaz parfait et par conséquent en fonction de l'énergie totale puisque  $e_T = e_I + e_C$ . En développant l'énergie cinétique, la pression satisfait l'expression

$$p = \rho(\gamma - 1) \left( e_T - \frac{\mathbf{u}^\top \mathbf{u}}{2} \right) = (\gamma - 1) \left( \rho e_T - \frac{(\rho\mathbf{u})^\top (\rho\mathbf{u})}{2\rho} \right) \quad (4.47)$$

qui ne dépend que des variables conservatives. Cependant, l'apparition d'une fraction rationnelle dans l'expression conduit à des termes difficiles à traiter lors de la projection de Galerkin après remplacement des variables conservatives par leur décomposition sur la base des modes POD. La même difficulté survient pour le terme de température dans le flux de chaleur  $\mathbf{q}_\theta$ , et plus généralement dès qu'un terme isolé  $\mathbf{u}$  ou  $e_T$  apparaît dans les équations du système (4.28). Ce problème, déjà signalé par Vigo [240] et Rowley [192], provient principalement du terme de masse volumique  $\rho$  qui doit être introduit comme facteur multiplicatif dans les termes isolés  $\mathbf{u}$  et  $e_T$  pour faire apparaître les variables conservatives  $\rho\mathbf{u}$  et  $\rho e_T$ .

La décomposition orthogonale propre des variables « quasi-conservatives »  $\mathbf{q} = [\rho, \mathbf{u}, e_T]^\top$  ne règle pas totalement le problème puisque le système dynamique devient certes polynomial mais s'écrit sous la forme implicite  $\mathbf{f}(\mathbf{q}, \mathbf{q}, t) = \mathbf{0}$ . D'autres jeux de variables peuvent alors être envisagés pour tenter de remédier au problème. En effet, l'équation d'état liant les variables thermodynamiques permet de transformer les équations de bilan pour obtenir un système d'équations équivalent dont on espère que la forme deviendra polynomiale et permettra de conserver une forme explicite  $\dot{\mathbf{q}} = \mathbf{f}(\mathbf{q}, t)$  pour les équations aux dérivées partielles. Un certain nombre de jeux de variables est envisagé dans l'annexe E.2 et il s'avère que le choix des *variables primitives modifiées*  $\mathbf{q} = [\vartheta, \mathbf{u}, p]^\top$  remplit toutes les conditions souhaitées grâce à l'introduction du *covolume*  $\vartheta = 1/\rho$ . Ce jeu de variables a notamment été employé par Vigo [239, 240] pour développer ensuite un modèle d'ordre réduit selon une approche POD-Galerkin. Le même modèle a ensuite été repris par Iollo *et al.* [117] et plus récemment par Gloerfelt [87] et Bourguet *et al.* [34]. Rowley [192] et Moreno *et al.* [169] ont quant à eux préféré émettre l'hypothèse supplémentaire que le fluide était isentropique, afin de pouvoir ramener les équations de Navier-Stokes à un système quadratique gouverné seulement par la vitesse locale du son  $a = \sqrt{\gamma p / \rho}$  et la vitesse du fluide  $\mathbf{u}$ .

Le jeu de variables primitives modifiées s'impose comme le plus prometteur puisqu'il permet d'obtenir une forme polynomiale des équations de Navier-Stokes sans faire l'hypothèse d'un fluide isentropique. Il est toutefois important de préciser que pour y parvenir, il est nécessaire de supposer que la viscosité est constante. Dans le cas contraire, si la viscosité suit la loi de Sutherland habituellement adoptée pour les fluides compressibles, la dépendance avec la température est telle que le système dynamique perd sa forme polynomiale. Par ailleurs, si l'écoulement est turbulent il devient difficile de traiter les variations de viscosité turbulente dans le système tout en conservant une forme polynomiale. Vigo [240] a tenté de modéliser les effets turbulents au moyen du modèle de Spalart-Allmaras [219], mais des approximations drastiques doivent être effectuées afin de maintenir la forme polynomiale. Le champ aérodynamique doit donc être reconstruit à chaque itération pour pouvoir calculer les termes non polynomiaux mais le gain en temps de calcul est alors perdu. Vigo s'est donc contenté de conserver le premier mode POD de viscosité turbulente qui correspond à la viscosité turbulente moyenne en temps.

En général, les variations de viscosité sont négligées en première instance pour la construction du modèle d'ordre réduit, mais il s'agit d'une hypothèse relativement forte dont les conséquences peuvent être

désastreuses pour la stabilité du modèle d'ordre réduit. Ainsi, les modèles d'ordre réduit sont la plupart du temps corrigés afin de pallier les erreurs liées aux simplifications effectuées pour la modélisation. On reviendra sur ce problème par la suite lors des applications numériques.

**Remarque 4.3 : Construction du modèle d'ordre réduit d'un écoulement incompressible**

Lorsque le fluide est supposé incompressible, le bilan de quantité de mouvement suffit seul à décrire la dynamique de l'écoulement dans le cas laminaire. La masse volumique et la viscosité sont alors traitées comme des termes constants si bien que la réponse du système se déduit d'un ensemble d'équations polynomiales quadratiques portant sur les composantes de la vitesse  $\mathbf{q} = [u, v, w]^T$ . L'approche POD vectorielle (2.83) employée pour décomposer les variables se combine donc bien avec la projection de Galerkin des équations du fluide et produit un modèle d'ordre réduit polynomial quadratique. Dans ce cas, le jeu de variables employé pour la construction du modèle d'ordre réduit est identique à celui habituellement utilisé pour résoudre les équations de Navier-Stokes avec une approche Volumes Finis ou Éléments Finis.

#### 4.4.2 Définition du produit scalaire

L'approche POD-Galerkin envisagée pour construire le modèle d'ordre réduit nécessite d'effectuer la projection selon le produit scalaire de l'espace de Hilbert considéré. L'espace  $H = L^2(\Omega)$  des fonctions de carré intégrable est un bon candidat puisqu'on peut raisonnablement considérer que le champ aérodynamique solution des équations de Navier-Stokes vérifie à chaque instant les propriétés d'appartenance à l'espace  $L^2$ . Le produit scalaire associé à cet espace avait été défini par la relation (2.91) pour des vecteurs de fonctions de dimension  $n_v$ . Le vecteur  $\mathbf{q} = [\vartheta, \mathbf{u}, p]^T$  composé des variables primitives modifiées sera donc considéré comme un élément de  $H = (L^2(\Omega))^{n_v}$  avec  $n_v = 2 + d$  où  $d$  est la dimension du domaine spatial tel que  $\Omega \subset \mathbb{R}^d$ .

L'application donnée par la relation (2.91) définit bien un produit scalaire de manière cohérente au sens mathématique. Toutefois, la signification physique est remise en question puisque les différentes composantes du vecteur  $\mathbf{q}$  ont des unités physiques diverses dont la somme n'a aucun sens. L'idéal consisterait à travailler avec un produit scalaire tel que la norme induite par le produit scalaire soit reliée à une certaine énergie physique  $\mathcal{E}$  du système par la relation

$$\|\mathbf{q}\|^2 = \langle \mathbf{q}, \mathbf{q} \rangle = \mathcal{E}. \quad (4.48)$$

Ceci justifierait ainsi l'assertion que la décomposition orthogonale propre soit basée sur la recherche des structures contenant le plus d'énergie du système puisque le terme général de l'opérateur de corrélation  $\mathbf{R}^*$  dont on recherche les valeurs propres est proportionnel à  $\langle \mathbf{q}^{(i)} - \bar{\mathbf{q}}, \mathbf{q}^{(j)} - \bar{\mathbf{q}} \rangle$ .

Pour un fluide incompressible, la définition (2.91) est cohérente puisque le vecteur des inconnues ne contient que les composantes de la vitesse :  $\mathbf{q} = [u, v, w]^T$ . Par conséquent, la norme induite par le produit scalaire standard est liée à l'énergie cinétique volumique  $e_C$  du système :  $\|\mathbf{q}\|^2 = \int_{\Omega} (u^2 + v^2 + w^2) d\Omega = \int_{\Omega} 2e_C d\Omega$ . Le terme sous l'intégrale s'écrit donc  $\mathbf{q}^T \mathbf{P}_{\mathcal{E}} \mathbf{q}$  avec  $\mathbf{P}_{\mathcal{E}} = \mathbf{I}_3$ . En choisissant plutôt  $\mathbf{P}_{\mathcal{E}} = \rho/2 \mathbf{I}_3$ , la norme induite par le produit scalaire correspond exactement à l'énergie cinétique puisque

$$\|\mathbf{q}\|^2 = \int_{\Omega} \mathbf{q}^T \mathbf{P}_{\mathcal{E}} \mathbf{q} d\Omega = \int_{\Omega} \frac{\rho}{2} (u^2 + v^2 + w^2) d\Omega = \int_{\Omega} \rho e_C d\Omega = \mathcal{E}_C. \quad (4.49)$$

Par analogie, le produit scalaire associé aux variables d'un écoulement compressible devrait pour sa part être recherché de manière à ce que la norme induite représente l'énergie totale du système, puisque les variables thermodynamiques associées à l'énergie interne du système participent aussi à la dynamique. Il serait donc souhaitable de trouver une expression de  $\mathbf{P}_{\mathcal{E}}$  telle que  $\mathbf{q}^T \mathbf{P}_{\mathcal{E}} \mathbf{q} = \rho e_T$ . Dans ce cas, la norme induite par le produit scalaire correspondrait à l'énergie totale du système ce qui fournirait une description cohérente de l'opérateur de corrélation, tant d'un point de vue dimensionnel qu'énergétique. L'énergie totale volumique du système s'écrit

$$\rho e_T = \rho e_I + \rho e_C = \frac{p}{\gamma - 1} + \frac{1}{2\vartheta} \mathbf{u}^T \mathbf{u} \quad (4.50)$$

en fonction des variables primitives modifiées  $\mathbf{q} = [\vartheta, \mathbf{u}, p]^T$ . Le problème est qu'il ne s'agit pas d'une forme quadratique puisque l'énergie interne est linéaire en fonction de  $p$  alors que l'énergie cinétique s'exprime à

l'aide d'une fraction rationnelle. Ce problème est commun à la plupart des jeux de variables envisagés dans l'annexe E.2. Il est alors possible de définir la norme de manière à ce qu'elle corresponde à l'énergie totale volumique  $e_T$ , bien qu'il ne s'agisse pas exactement d'une quantité conservée. Dans ce cas, les variables primitives modifiées  $\mathbf{q} = [\vartheta, \mathbf{u}, p]^T$  fournissent une forme quadratique définie par la matrice

$$\mathbf{P}_{\mathcal{E}} = \begin{bmatrix} 0 & \mathbf{0}_{1,d} & \frac{1}{2(\gamma-1)} \\ \mathbf{0}_{d,1} & 1/2 \mathbf{I}_d & \mathbf{0}_{d,1} \\ \frac{1}{2(\gamma-1)} & \mathbf{0}_{1,d} & 0 \end{bmatrix}. \quad (4.51)$$

Toutefois, une condition nécessaire pour que l'application qui au vecteur  $\mathbf{q}$  associe  $\int_{\Omega} \mathbf{q}^T \mathbf{P}_{\mathcal{E}} \mathbf{q} \, d\Omega$  soit un produit scalaire est que la matrice  $\mathbf{P}_{\mathcal{E}}$  soit définie positive [223]. Or cette propriété n'est pas vérifiée puisque la matrice en question possède la valeur propre négative  $-1/(2(\gamma-1))$ .

Rowley [192, 194] a attribué certaines sources d'instabilité des modèles d'ordre réduit POD-Galerkin à une mauvaise définition du produit scalaire et insistait donc sur la nécessité d'une définition cohérente qui préserve l'énergie du système dynamique. Sous l'hypothèse d'un fluide isentropique émise dans [192, 194], l'énergie totale volumique  $e_T$  a été écrite sous la forme  $e_T = a^2/(\gamma-1) + 1/2 \mathbf{u}^T \mathbf{u}$  qui est bien quadratique en fonction du vecteur de variables  $\mathbf{q} = [\mathbf{u}, a]^T$  qui suffit à décrire un écoulement isentropique. Dans ce cas, la matrice diagonale  $\mathbf{P}_{\mathcal{E}} = \text{diag}(1/2, 1/(\gamma-1))$  associée au produit scalaire est bien définie positive. Puisque le produit scalaire est basé sur l'énergie du système, Rowley *et al.* [194] ont montré que la projection de Galerkin préserve la stabilité du système, au moins pour les points d'équilibre autour de l'origine. Cependant, la stabilité des autres points d'équilibre et celle des cycles limites n'est pas assurée.

En dehors de l'hypothèse d'un fluide compressible isentropique, il est difficile de construire un produit scalaire énergétique cohérent. Une solution alternative consiste donc à définir le produit scalaire de façon à satisfaire les propriétés mathématiques, et à adimensionner les variables afin de remédier aux problèmes d'unité. Le produit scalaire (2.91) ainsi modifié est alors défini par

$$\langle \mathbf{q}, \mathbf{r} \rangle_{L^2} = \int_{\Omega} \mathbf{q}^T \mathbf{P}_{\text{adim}}^T \mathbf{P}_{\text{adim}} \mathbf{r} \, d\Omega, \quad (4.52)$$

où la matrice  $\mathbf{P}_{\text{adim}}$  contient des coefficients de pondération choisis de manière à rendre chaque variable  $q_i$  et  $r_i$  adimensionnelle. Il suffit alors de choisir une densité de référence  $\rho_0$  et une vitesse de référence  $U_0$  et de construire la matrice de pondération

$$\mathbf{P}_{\text{adim}} = \begin{bmatrix} \rho_0 & \mathbf{0}_{1,d} & 0 \\ \mathbf{0}_{d,1} & \frac{1}{U_0} \mathbf{I}_{d,d} & \mathbf{0}_{d,1} \\ 0 & \mathbf{0}_{1,d} & \frac{1}{\rho_0 U_0^2} \end{bmatrix} \quad (4.53)$$

associée aux variables primitives modifiées  $\mathbf{q} = [\vartheta, \mathbf{u}, p]^T$ . Rowley [192] a cependant signalé que la procédure d'adimensionnement n'est pas transparente et altère l'optimalité des modes POD obtenus puisque les projections effectuées dépendent du choix relativement arbitraire des valeurs prises par les grandeurs d'adimensionnement. Bourguet *et al.* [34] ont quant à eux proposé d'utiliser un adimensionnement de chaque variable par la variance statistique temporelle. Le produit  $\mathbf{P}_{\text{adim}}^T \mathbf{P}_{\text{adim}}$  a alors été remplacé par une matrice diagonale dont les termes sont donnés par  $1/(\sigma_i^2 + \varepsilon)$ , où  $\varepsilon$  est une petite constante positive introduite pour éviter les indéterminations sur les bords du domaine. La variance statistique temporelle  $\sigma_i^2$  est quant à elle définie pour chaque variable  $q_i$  pour  $i = 1, \dots, n_v$  par

$$\sigma_i^2 = \frac{1}{T_e} \int_{t_0}^{t_0+T_e} (q_i(t) - \bar{q}_i)^2 \, dt. \quad (4.54)$$

Le produit scalaire obtenu de cette manière relève néanmoins toujours d'un processus d'adimensionnement susceptible d'affecter l'optimalité des modes POD.

Pour terminer, il est important de mentionner que pour des raisons de stabilité plutôt que par un souci de cohérence dimensionnelle du produit des variables, Kirby [132] a proposé de définir le produit scalaire sur l'espace de Sobolev  $W^{n,p} = \{u \in L^p(\Omega); D^\alpha u \in L^p(\Omega)\}$ , où  $D^\alpha$  désigne une dérivée partielle de  $u$  telle que  $0 \leq |\alpha| \leq n$  et  $L^p$  est l'espace de Lebesgue pour lequel les puissances d'ordre  $p$  de  $u$  sont intégrables.

De cette façon, le calcul des modes POD tient compte des corrélations du champ, du flux, de la dissipation, et ainsi de suite. L'application de ce produit scalaire sur l'équation de Kuramoto-Sivashinsky a révélé que la prise en compte des dérivées d'ordre deux dans le calcul des modes POD a un effet stabilisateur pour la réponse du modèle d'ordre réduit. Iollo *et al.* [117] ont pour leur part redéfini le produit scalaire sur l'espace de Sobolev  $H^1 = \{u \in L^2(\Omega); D^1 u \in L^2(\Omega)\}$  de façon à tenir compte des gradients. Dans ce cas, le produit scalaire est transformé en

$$\langle \mathbf{q}, \mathbf{r} \rangle_{H^1} = \int_{\Omega} \sum_{i=1}^{n_v} q_i r_i d\Omega + \mu_{H^1} \int_{\Omega} \sum_{i=1}^{n_v} \nabla q_i \cdot \nabla r_i d\Omega = \sum_{i=1}^{n_v} \int_{\Omega} \left[ q_i r_i + \mu_{H^1} \sum_{l=1}^d \frac{\partial q_i}{\partial x_l} \frac{\partial r_i}{\partial x_l} \right] d\Omega, \quad (4.55)$$

où  $\mu_{H^1}$  est un paramètre de pondération qui doit être ajusté — de manière empirique — pour stabiliser correctement le modèle d'ordre réduit. D'après les travaux existants, aucun produit scalaire véritablement cohérent n'a pu être établi dans le cas général des équations de Navier-Stokes pour un fluide compressible. L'adimensionnement des variables semble donc être pour l'instant la seule solution envisageable pour effectuer les produits scalaires sans violer la cohérence dimensionnelle des produits effectués. Bien que ce choix ne constitue pas la solution idéale, il s'inscrit naturellement dans le processus de calcul du code *elsA* qui fournit des clichés sous forme adimensionnelle.

### 4.4.3 Équations locales de Navier-Stokes en formulation ALE avec les variables primitives modifiées

D'après les remarques effectuées dans les deux paragraphes précédents, il s'avère que l'écriture des équations avec les variables primitives modifiées est adéquate pour la construction d'un modèle d'ordre réduit POD-Galerkin. Toutefois, les équations ont été établies dans l'annexe E.2 dans le cadre d'une formulation eulérienne pour simplifier la présentation. Dans l'optique de développer un modèle d'ordre réduit en aéroélasticité, il est nécessaire d'étendre les équations locales (E.98) à une description ALE formulée dans le repère absolu  $\mathcal{R}_A$  ou entraîné  $\mathcal{R}_E$ .

#### 4.4.3.1 Formulation dans le repère absolu

La description ALE se déduit immédiatement des équations (E.98) en notant que la dérivée temporelle qui apparaît dans le membre de gauche correspond à la dérivée  $\partial \cdot / \partial t|_{\mathbf{x}}$  puisqu'il s'agit d'une description eulérienne. Par conséquent, la relation (4.9) est mise à profit pour transformer les équations locales et obtenir ainsi une description ALE dans laquelle intervient la vitesse de déplacement  $\mathbf{s}$  des points du domaine arbitraire. Lorsque les quantités vectorielles sont décrites dans le repère absolu  $\mathcal{R}_A$ , les équations de Navier-Stokes prennent la forme

$$\begin{cases} \left. \frac{\partial \vartheta}{\partial t} \right|_{\mathbf{x}} + (\mathbf{u}_A - \mathbf{s}_A) \cdot \nabla_A \vartheta & = \vartheta \operatorname{div}_A \mathbf{u}_A \\ \left. \frac{\partial \mathbf{u}_A}{\partial t} \right|_{\mathbf{x}} + (\mathbf{u}_A - \mathbf{s}_A) \cdot \nabla_A \mathbf{u}_A & = -\vartheta \nabla_A p + \vartheta \operatorname{div}_A \boldsymbol{\tau}(\mathbf{u}_A) \\ \left. \frac{\partial p}{\partial t} \right|_{\mathbf{x}} + (\mathbf{u}_A - \mathbf{s}_A) \cdot \nabla_A p & = -\gamma p \operatorname{div}_A \mathbf{u}_A + (\gamma - 1) \boldsymbol{\tau}(\mathbf{u}_A) : \nabla_A \mathbf{u}_A + \operatorname{div}_A \left[ \frac{k_{\theta}}{c_v} \nabla_A (\vartheta p) \right]. \end{cases} \quad (4.56)$$

#### 4.4.3.2 Formulation dans le repère entraîné

Les équations écrites avec les variables primitives modifiées dans le repère entraîné s'obtiennent soit à partir des équations intégrales de Navier-Stokes fournies aux équations (4.31) et (4.33) dans le repère entraîné, soit à partir des équations locales (4.56) écrites en formulation ALE dans le repère absolu. La première alternative nécessite d'effectuer le travail de changement de variables et de ramener les équations intégrales à des équations locales alors que la seconde requiert seulement une transformation de repère. On opte pour la seconde alternative pour bénéficier des résultats déjà établis dans l'annexe E.1.

**Formulation en vitesse absolue** Pour établir les équations locales, on a besoin de relations de passage entre la base absolue  $\mathcal{B}_A$  et la base du repère entraîné  $\mathcal{B}_E$  pour l'opérateur  $\mathbf{u} \cdot \nabla$ . Il suffit pour cela de remarquer que pour un scalaire  $s$  et un vecteur  $\mathbf{c}$ ,  $(\mathbf{c} \cdot \nabla)s = \mathbf{c} \cdot (\nabla s)$ . De même, pour un vecteur  $\mathbf{r}$ ,  $(\mathbf{c} \cdot \nabla)\mathbf{r} = (\nabla \mathbf{r}) \mathbf{c}$ . Ainsi, le changement de base pour l'opérateur d'advection s'écrit

$$\mathbf{c}_A \cdot \nabla_A s = \mathbf{c}_E \cdot \nabla_E s \quad \text{et} \quad (4.57a)$$

$$\mathbf{c}_A \cdot \nabla_A \mathbf{r}_A = \mathbf{R} \mathbf{c}_E \cdot \nabla_E \mathbf{r}_E \quad (4.57b)$$

à l'aide des formules (E.27) et (E.29). L'équation de conservation du covolume se déduit donc de la première équation du système (4.56) avec les relations (4.57a) et (E.32) qui conduisent à

$$\left. \frac{\partial \vartheta}{\partial t} \right|_{\mathbf{x}} + (\mathbf{u}_E - \mathbf{s}_E) \cdot \nabla_E \vartheta = \vartheta \operatorname{div}_E \mathbf{u}_E. \quad (4.58)$$

Pour l'équation de conservation de la quantité de mouvement, on suit le même procédé que celui employé au paragraphe E.1.2.2 : la formule de dérivation en repère mobile (E.22) est appliquée cette fois à  $\mathbf{r}_A = \mathbf{u}_A$  puis les équations sont projetées dans le repère entraîné en multipliant par  $\mathbf{R}^\top$  ce qui donne

$$\mathbf{R}^\top \left. \frac{\delta \mathbf{u}_A}{\delta t} \right|_{\mathbf{x}} + \mathbf{R}^\top \boldsymbol{\omega}_A \wedge \mathbf{u}_A + \mathbf{R}^\top (\mathbf{u}_A - \mathbf{s}_A) \cdot \nabla_A \mathbf{u}_A = -\mathbf{R}^\top \vartheta \nabla_A p + \mathbf{R}^\top \vartheta \operatorname{div}_A \boldsymbol{\tau}(\mathbf{u}_A). \quad (4.59)$$

En explicitant la dérivée  $\delta \cdot / \delta t$  définie par l'équation (E.18), le premier terme s'écrit  $\partial \mathbf{u}_E / \partial t$ . Le second terme est modifié en introduisant le vecteur vitesse angulaire défini à l'équation (E.40) :  $\mathbf{R}^\top \boldsymbol{\omega}_A \wedge \mathbf{u}_A = \mathbf{R}^\top \boldsymbol{\Omega}_A \mathbf{u}_A = \mathbf{R}^\top (\dot{\mathbf{R}} \mathbf{R}^\top) \mathbf{R} \mathbf{u}_E = \boldsymbol{\Omega}_E \mathbf{u}_E = \boldsymbol{\omega}_E \wedge \mathbf{u}_E$ . Le troisième et le quatrième terme sont simplement modifiés à l'aide des relations (4.57b) et (E.27) respectivement. Enfin, le dernier terme se déduit de la relation (E.34). Le bilan de quantité de mouvement a donc pour expression

$$\left. \frac{\partial \mathbf{u}_E}{\partial t} \right|_{\mathbf{x}} + (\mathbf{u}_E - \mathbf{s}_E) \cdot \nabla_E \mathbf{u}_E = -\vartheta \nabla_E p + \vartheta \operatorname{div}_E \boldsymbol{\tau}(\mathbf{u}_E) - \boldsymbol{\omega}_E \wedge \mathbf{u}_E. \quad (4.60)$$

Enfin le bilan de conservation de l'énergie se déduit de la dernière équation dans le système (4.56) à l'aide des relations (4.57a) et (E.32) pour transformer les termes  $(\mathbf{u}_A - \mathbf{s}_A) \cdot \nabla_A p$  en  $(\mathbf{u}_E - \mathbf{s}_E) \cdot \nabla_E p$  et  $\gamma p \operatorname{div}_A \mathbf{u}_A$  en  $\gamma p \operatorname{div}_E \mathbf{u}_E$  respectivement.

Les deux derniers termes  $\nabla_A \mathbf{u}_A : \boldsymbol{\tau}(\mathbf{u}_A)$  et  $\operatorname{div}_A [k_\theta / c_v \nabla_A (p\vartheta)]$  sont invariants par changement de repère. En effet, le double produit de contraction peut s'écrire  $(\mathbf{R} \mathbf{A} \mathbf{R}^\top) : (\mathbf{R} \mathbf{B} \mathbf{R}^\top)$  en utilisant les relations (E.29) et (E.43) et en notant  $\mathbf{A} = \nabla_E \mathbf{u}_E$  et  $\mathbf{B} = \boldsymbol{\tau}(\mathbf{u}_E)$ . On a alors  $(\mathbf{R} \mathbf{A} \mathbf{R}^\top) : (\mathbf{R} \mathbf{B} \mathbf{R}^\top) = (R_{ij} A_{jk} R_{lk}) (R_{ij} B_{jk} R_{lk}) = (R_{ij} R_{ji}^\top) A_{jk} B_{jk} (R_{lk} R_{kl}^\top)$ . Or puisque la matrice  $\mathbf{R}$  est orthogonale, le double produit de contraction se réduit à  $A_{jk} B_{jk} = \nabla_E \mathbf{u}_E : \boldsymbol{\tau}(\mathbf{u}_E)$ .

Enfin, en supposant  $k_\theta / c_v$  constant, le dernier terme s'écrit  $k_\theta / c_v \operatorname{div}_A (\nabla_A (p\vartheta)) = k_\theta / c_v \Delta_A (p\vartheta)$ . Or puisque le laplacien s'applique à une quantité scalaire, on a  $\Delta_A (p\vartheta) = \Delta_E (p\vartheta)$  et le bilan pour la pression s'écrit alors

$$\left. \frac{\partial p}{\partial t} \right|_{\mathbf{x}} + (\mathbf{u}_E - \mathbf{s}_E) \cdot \nabla_E p = -\gamma p \operatorname{div}_E \mathbf{u}_E + (\gamma - 1) \nabla_E \mathbf{u}_E : \boldsymbol{\tau}(\mathbf{u}_E) + \frac{k_\theta}{c_v} \Delta_E (p\vartheta). \quad (4.61)$$

Finalement, la formulation ALE locale des équations de Navier-Stokes écrites avec les variables primitives modifiées dans le repère entraîné avec les composantes de la vitesse absolue s'écrit

$$\left\{ \begin{array}{l} \left. \frac{\partial \vartheta}{\partial t} \right|_{\mathbf{x}} + (\mathbf{u}_E - \mathbf{s}_E) \cdot \nabla_E \vartheta = \vartheta \operatorname{div}_E \mathbf{u}_E \\ \left. \frac{\partial \mathbf{u}_E}{\partial t} \right|_{\mathbf{x}} + (\mathbf{u}_E - \mathbf{s}_E) \cdot \nabla_E \mathbf{u}_E = -\vartheta \nabla_E p + \vartheta \operatorname{div}_E \boldsymbol{\tau}(\mathbf{u}_E) - \boldsymbol{\omega}_E \wedge \mathbf{u}_E \\ \left. \frac{\partial p}{\partial t} \right|_{\mathbf{x}} + (\mathbf{u}_E - \mathbf{s}_E) \cdot \nabla_E p = -\gamma p \operatorname{div}_E \mathbf{u}_E + (\gamma - 1) \nabla_E \mathbf{u}_E : \boldsymbol{\tau}(\mathbf{u}_E) + \frac{k_\theta}{c_v} \Delta_E (p\vartheta). \end{array} \right. \quad (4.62)$$

Les équations formulées dans le repère entraîné avec les variables primitives modifiées conservent une forme quadratique polynomiale en fonction des variables  $\mathbf{q}_E = [\vartheta, \mathbf{u}_E, p]^\top$  lorsque les composantes de la vitesse absolue exprimée dans le repère  $\mathcal{R}_E$  sont employées. Ainsi, il restera possible de formuler le modèle d'ordre réduit au moyen d'une projection de Galerkin sur la base constituée des modes POD.

**Formulation en vitesse relative** Enfin, il reste à établir les équations de Navier-Stokes dans le repère entraîné avec les variables primitives modifiées utilisant les composantes de la vitesse relative. Ici encore, il suffit de remplacer les composantes de la vitesse absolue  $\mathbf{u}_E$  par  $\mathbf{u}_E = \mathbf{v}_E + \mathbf{s}_{e,E}$  et de scinder la vitesse du maillage en deux contributions telles que  $\mathbf{s}_E = \mathbf{s}_{e,E} + \mathbf{s}_{d,E}$ . Puisque le champ de vitesse  $\mathbf{s}_{e,E}$  est solénoïdal, la propriété (E.52b) permet d'obtenir facilement le bilan de conservation du covolume

$$\left. \frac{\partial \vartheta}{\partial t} \right|_{\mathbf{x}} + (\mathbf{v}_E - \mathbf{s}_{d,E}) \cdot \nabla_E \vartheta = \vartheta \operatorname{div}_E \mathbf{v}_E \quad (4.63)$$

à partir de l'équation (4.58). Ensuite, le bilan de conservation de la vitesse absolue (4.60) est transformé afin d'obtenir un bilan pour la vitesse relative en utilisant la relation (E.53), d'où l'expression

$$\begin{aligned} \left. \frac{\partial \mathbf{v}_E}{\partial t} \right|_{\mathbf{x}} + \left. \frac{\partial \mathbf{s}_{e,E}}{\partial t} \right|_{\mathbf{x}} + (\mathbf{v}_E - \mathbf{s}_{d,E}) \cdot \nabla_E \mathbf{v}_E + (\mathbf{v}_E - \mathbf{s}_{d,E}) \cdot \nabla_E \mathbf{s}_{e,E} = \\ -\vartheta \nabla_E p + \vartheta \operatorname{div}_E \boldsymbol{\tau}(\mathbf{v}_E) - \boldsymbol{\omega}_E \wedge \mathbf{v}_E - \boldsymbol{\omega}_E \wedge \mathbf{s}_{e,E}. \end{aligned} \quad (4.64)$$

Le théorème de Reynolds local (4.9) utilisé avec la vitesse de déformation de maillage  $\mathbf{s}_{d,E}$  permet de transformer la seconde dérivée temporelle suivant l'égalité

$$\left. \frac{\partial \mathbf{s}_{e,E}}{\partial t} \right|_{\mathbf{x}} + (\mathbf{v}_E - \mathbf{s}_{d,E}) \cdot \nabla_E \mathbf{s}_{e,E} = \left. \frac{\partial \mathbf{s}_{e,E}}{\partial t} \right|_{\mathbf{x}} + \cancel{\mathbf{s}_{d,E} \cdot \nabla_E \mathbf{s}_{e,E}} + \mathbf{v}_E \cdot \nabla_E \mathbf{s}_{e,E} - \cancel{\mathbf{s}_{d,E} \cdot \nabla_E \mathbf{s}_{e,E}} \quad (4.65)$$

dans laquelle une partie des termes se simplifie. Enfin, on remarque que  $\mathbf{v}_E \cdot \nabla_E \mathbf{s}_{e,E} = (\nabla_E \mathbf{s}_{e,E}) \mathbf{v}_E = \boldsymbol{\omega}_E \wedge \mathbf{v}_E$  avec la propriété (E.52c). Par conséquent, en remplaçant l'expression (4.65) dans (4.64) on parvient au bilan

$$\left. \frac{\partial \mathbf{v}_E}{\partial t} \right|_{\mathbf{x}} + (\mathbf{v}_E - \mathbf{s}_{d,E}) \cdot \nabla_E \mathbf{v}_E = -\vartheta \nabla_E p + \vartheta \operatorname{div}_E \boldsymbol{\tau}(\mathbf{v}_E) + \mathbf{f}_{\text{Cor}} + \mathbf{f}_{\text{cen}} \quad (4.66)$$

dans lequel on a fait apparaître la force de Coriolis ainsi que la force centrifuge définies à l'équation (4.34). Le bilan de la pression s'obtient en modifiant le terme convectif comme pour le bilan du covolume, puis en utilisant la propriété (E.52b) pour le terme  $\gamma p \operatorname{div}_E \mathbf{u}_E$  avec  $\mathbf{u}_E = \mathbf{v}_E + \mathbf{s}_{e,E}$ . Enfin, le double produit de contraction s'écrit  $\nabla_E \mathbf{u}_E : \boldsymbol{\tau}(\mathbf{u}_E) = \nabla_E \mathbf{v}_E : \boldsymbol{\tau}(\mathbf{v}_E) + \nabla_E \mathbf{s}_{e,E} : \boldsymbol{\tau}(\mathbf{v}_E)$  avec la relation (E.53). Par ailleurs, puisque le tenseur  $\nabla_E \mathbf{s}_{e,E}$  est antisymétrique et que le tenseur  $\boldsymbol{\tau}(\mathbf{v}_E)$  est au contraire symétrique, le double produit de contraction  $\nabla_E \mathbf{s}_{e,E} : \boldsymbol{\tau}(\mathbf{v}_E)$  s'annule<sup>10</sup>. Au final, le bilan pour la pression s'écrit donc

$$\left. \frac{\partial p}{\partial t} \right|_{\mathbf{x}} + (\mathbf{v}_E - \mathbf{s}_{d,E}) \cdot \nabla_E p = -\gamma p \operatorname{div}_E \mathbf{v}_E + (\gamma - 1) \nabla_E \mathbf{v}_E : \boldsymbol{\tau}(\mathbf{v}_E) + \frac{k_\theta}{c_v} \Delta_E (p\vartheta), \quad (4.67)$$

et la formulation ALE locale des équations de Navier-Stokes écrites avec les variables primitives modifiées dans le repère entraîné avec les composantes de la vitesse relative s'écrit

$$\left\{ \begin{array}{l} \left. \frac{\partial \vartheta}{\partial t} \right|_{\mathbf{x}} + (\mathbf{v}_E - \mathbf{s}_{d,E}) \cdot \nabla_E \vartheta = \vartheta \operatorname{div}_E \mathbf{v}_E \\ \left. \frac{\partial \mathbf{v}_E}{\partial t} \right|_{\mathbf{x}} + (\mathbf{v}_E - \mathbf{s}_{d,E}) \cdot \nabla_E \mathbf{v}_E = -\vartheta \nabla_E p + \vartheta \operatorname{div}_E \boldsymbol{\tau}(\mathbf{v}_E) + \mathbf{f}_{\text{Cor}} + \mathbf{f}_{\text{cen}} \\ \left. \frac{\partial p}{\partial t} \right|_{\mathbf{x}} + (\mathbf{v}_E - \mathbf{s}_{d,E}) \cdot \nabla_E p = -\gamma p \operatorname{div}_E \mathbf{v}_E + (\gamma - 1) \nabla_E \mathbf{v}_E : \boldsymbol{\tau}(\mathbf{v}_E) + \frac{k_\theta}{c_v} \Delta_E (p\vartheta). \end{array} \right. \quad (4.68)$$

<sup>10</sup>Sur ce point, voir la remarque effectuée dans l'annexe E.1.3 p. 257.



## 4.5 Synthèse du chapitre 4

Dans ce chapitre, les formulations ALE des équations intégrales de Navier-Stokes d'un fluide compressible visqueux ont tout d'abord été rappelées lorsque le système est décrit dans un repère absolu ou entraîné avec les variables conservatives. Les différentes formulations sont résumées dans le tableau 4.1 en fonction des composantes de la vitesse employées. Ces équations, écrites à l'aide des variables conservatives, sont généralement celles résolues par les codes de calcul au moyen d'une discrétisation en Volumes Finis. Elles sont en particulier adaptées pour traiter les problèmes en aéroélasticité où une structure déformable et éventuellement animée d'un mouvement de corps rigide est considérée.

	$\mathcal{R}_A$		$\mathcal{R}_E$	
Composantes de vitesse	$\mathbf{u}_A$	$\mathbf{v}_A$	$\mathbf{u}_E$	$\mathbf{v}_E$
Formulation	FALE-RAUA	FALE-RAVA	FALE-REUE	FALE-REVE

**Tableau 4.1 :** Notations employées pour les différentes formulations ALE possibles des équations du fluide.

Toutefois, dans l'optique de construire ensuite un modèle d'ordre réduit POD-Galerkin, les formulations basées sur les variables conservatives sont inadéquates. Il s'avère alors judicieux d'employer les *variables primitives modifiées* comme l'a suggéré Vigo [240] puisque les équations prennent une forme polynomiale quadratique qui se prête bien à la projection de Galerkin intervenant au cours de la construction du modèle d'ordre réduit. Les équations locales formulées avec ce jeu de variables ont alors été étendues à une description ALE pour tenir compte d'une structure déformable placée dans le fluide, ce qui a conduit au développement de la formulation FALE-RAUA avec les variables primitives modifiées. Enfin, les équations ont été formulées dans le repère entraîné  $\mathcal{R}_E$  afin de tenir compte d'un éventuel mouvement de corps rigide imposé à la structure en plus de son caractère déformable. De cette manière, les formulations FALE-REUE et FALE-REVE ont pu être développées avec les variables primitives modifiées. Le point crucial est que ces formulations conservent une forme polynomiale quadratique — même dans le repère entraîné — si bien que la projection de Galerkin reste possible.

Dans les chapitres suivants, on s'intéresse à la construction de modèles d'ordre réduit POD-Galerkin sur la base des équations écrites avec les variables primitives modifiées. Dans un premier temps, un modèle d'ordre réduit dit *aérodynamique* est construit pour l'écoulement dans un domaine à parois fixes, puis on s'intéressera finalement au développement d'un modèle d'ordre réduit *aéroélastique* afin de tenir compte d'une structure mobile dans le domaine fluide.

# Construction et mise en œuvre d'un modèle d'ordre réduit POD-Galerkin non-linéaire et autonome pour l'aérodynamique

## Sommaire

<b>5.1</b>	<b>Introduction</b>	<b>122</b>
<b>5.2</b>	<b>Construction du modèle d'ordre réduit POD-Galerkin</b>	<b>123</b>
5.2.1	Projection de Galerkin des équations de Navier-Stokes d'un fluide compressible	123
5.2.2	Calcul des coefficients du modèle d'ordre réduit	125
5.2.2.1	Expression des coefficients quadratiques convectifs	126
5.2.2.2	Expression des coefficients quadratiques diffusifs	126
5.2.3	Intégration du modèle d'ordre réduit en temps	129
<b>5.3</b>	<b>Correction des instabilités des modèles d'ordre réduit POD-Galerkin</b>	<b>130</b>
5.3.1	Origine des instabilités et principe de la correction	130
5.3.2	Techniques de correction a priori des modèles d'ordre réduit POD-Galerkin	131
5.3.2.1	Correction des vecteurs de la base POD	131
5.3.2.2	Introduction d'amortissement dans le modèle d'ordre réduit	132
5.3.3	Techniques de correction a posteriori des modèles d'ordre réduit POD-Galerkin	134
5.3.3.1	Optimisation des coefficients du modèle d'ordre réduit	134
5.3.3.2	Stabilisation intrinsèque	136
5.3.4	Autres techniques de correction des modèles d'ordre réduit POD-Galerkin	138
<b>5.4</b>	<b>Calcul des modes POD de l'écoulement autour d'un profil NACA0012 fixe en forte incidence</b>	<b>139</b>
5.4.1	Simulation numérique de l'écoulement autour du profil NACA0012	139
5.4.2	Calcul des modes POD de l'écoulement autour du profil NACA0012	142
<b>5.5</b>	<b>Réponse du modèle d'ordre réduit de l'écoulement autour du profil NACA0012</b>	<b>147</b>
5.5.1	Réponse du modèle d'ordre réduit non corrigé	147
5.5.2	Introduction de termes correctifs dans le modèle d'ordre réduit	149
5.5.2.1	Correction du nombre de Reynolds	150
5.5.2.2	Correction de <i>Cazemier et al.</i>	151
5.5.2.3	Correction de <i>Kalb et Deane</i>	152
5.5.3	Calibrage des paramètres du modèle d'ordre réduit	154
5.5.3.1	Calibrage du nombre de Reynolds	154
5.5.3.2	Calibrage des coefficients du modèle d'ordre réduit	157
5.5.4	Comparaison des réponses corrigées	158
<b>5.6</b>	<b>Synthèse du chapitre 5</b>	<b>160</b>

## 5.1 Introduction

Maintenant que les équations du fluide ont été présentées dans le cas général d'un fluide newtonien homogène, il reste à décrire la forme du modèle d'ordre réduit basé sur les équations de Navier-Stokes d'un fluide compressible. Dans ce chapitre, un premier modèle d'ordre réduit POD-Galerkin est construit pour les équations autonomes non-linéaires de Navier-Stokes. Celles-ci sont formulées avec les variables primitives modifiées dans le repère absolu puisque les frontières du domaine spatial sont supposées fixes. La formulation FALE-RAUA se ramène donc dans ce cas à la formulation FEUL-RAUA basée sur une description eulérienne.

Dans un premier temps, les expressions analytiques des coefficients du modèle d'ordre réduit sont déterminées à partir de la projection de Galerkin. Le problème de la stabilité des modèles d'ordre réduit POD-Galerkin sera alors évoqué et un certain nombre de méthodes de correction seront présentées afin d'améliorer la stabilité et la précision de la réponse du modèle d'ordre réduit.

Le modèle d'ordre réduit sera alors évalué sur l'exemple d'un profil NACA0012 fortement incliné de sorte qu'un détachement périodique de vortex apparaisse dans le sillage. Les paramètres du fluide seront choisis de manière à obtenir un écoulement laminaire et faiblement compressible afin d'analyser la qualité du modèle d'ordre réduit en l'absence de modélisation des effets turbulents. Le modèle d'ordre réduit sera finalement corrigé au moyen de diverses méthodes afin de déterminer quelles sont les techniques les plus appropriées pour améliorer la précision de la réponse obtenue.

## 5.2 Construction du modèle d'ordre réduit POD-Galerkin

### 5.2.1 Projection de Galerkin des équations de Navier-Stokes d'un fluide compressible dans un domaine fixe

Le modèle d'ordre réduit est construit par projection de Galerkin des équations continues du système aérodynamique de manière à préserver les non-linéarités. Les équations de Navier-Stokes formulées avec les variables primitives modifiées  $\mathbf{q} = [\vartheta, \mathbf{u}, p]^T$  dans le chapitre précédent ont permis d'aboutir à un système explicite d'équations aux dérivées partielles dans lequel des non-linéarités polynomiales au plus quadratiques ont été mises en évidence. Cette forme se prête bien à la procédure de projection de Galerkin présentée au paragraphe 2.4.3 de manière générale.

On considère dans ce chapitre que l'écoulement est étudié dans un domaine  $\Omega$  dont les frontières  $\partial\Omega$  sont fixes. Ainsi, les équations de Navier-Stokes sont décrites au moyen de la formulation FEUL-RAUA qui se déduit de la formulation FALE-RAUA en supposant simplement que la vitesse des points du domaine arbitraire de la description ALE est nulle :  $\mathbf{s}_A = \mathbf{0}$ . Dans un premier temps, les équations intégrales (4.27) de Navier-Stokes sont résolues avec le logiciel *elsA* pour générer un ensemble de clichés  $\mathbf{q}^{(m)}$  à partir desquels les modes POD sont calculés. Les clichés ainsi obtenus sont adimensionnels et le vecteur vitesse  $\mathbf{u}$  est exprimé dans le repère absolu  $\mathcal{R}_A$ . En toute rigueur, il faudrait noter les clichés  $\mathbf{q}_A^{(m)}$  pour signifier que les vecteurs sont exprimés dans la base absolue, mais pour alléger la notation on omet l'indice  $A$  dans l'écriture des clichés ainsi que dans celle de la vitesse notée simplement  $\mathbf{u} = \mathbf{u}_A = \text{Mat}_{\mathcal{R}_A}(\mathbf{u})$ . Les modes POD sont alors déterminés à partir du problème aux valeurs propres (2.81) avec la relation (2.82). Les modes POD sont supposés être contenus dans l'espace de Hilbert  $H = (L^2(\Omega))^{n_v}$ . Le produit scalaire associé à cet espace est celui qui a été défini à l'équation (4.52) avec  $\mathbf{P}_{\text{adim}} = \mathbf{I}_{n_v}$  puisque les variables sont adimensionnelles. La base  $\Phi$  est constituée des  $q$  modes POD  $\varphi^{(i)}$  qui seront décomposés en  $n_v$  composantes relatives à chacune des variables :  $\varphi^{(i)} = [\varphi_{\vartheta}^{(i)}, \varphi_{\mathbf{u}}^{(i)}, \varphi_p^{(i)}]^T$ . Chaque contribution est une fonction de l'espace  $L^2(\Omega)$  mais ne correspond pas, lorsqu'elle est considérée seule, à un mode POD.

Avant de projeter les équations locales de Navier-Stokes données par les relations (4.56), il est préférable de les rendre adimensionnelles afin que le modèle d'ordre réduit s'inscrive dans la continuité du calcul aérodynamique effectué avec *elsA*. En outre, l'utilisation des équations adimensionnelles est rendue nécessaire par le choix du produit scalaire employé qui n'a de sens que si les variables ont toutes la même unité, ou, par la force des choses, si elles sont toutes adimensionnelles. L'adimensionnement est pratiqué à partir d'une longueur  $L_0$ , d'une vitesse  $U_0$ , d'une masse volumique  $\rho_0$  et d'une température  $\theta_0$  de références qui servent à définir un temps  $t_0 = L_0/U_0$ , une pression  $p_0 = \rho_0 U_0^2$ , et un covolume  $\vartheta_0 = 1/\rho_0$  pour l'adimensionnement. Finalement on introduit une viscosité de référence  $\mu_r$  et une conductivité thermique de référence  $k_{\theta,r}$  si bien que le processus d'adimensionnement fait apparaître les célèbres nombres de Reynolds  $\text{Re}$ , de Prandtl  $\text{Pr}$  et de Mach  $\text{Ma}$  respectivement définis par

$$\text{Re} = \frac{\rho_0 U_0 L_0}{\mu_r}, \quad \text{Pr} = \frac{\mu_r c_p}{k_{\theta,r}} \quad \text{et} \quad \text{Ma} = \frac{U_0}{a} \quad (5.1)$$

à partir des grandeurs de référence et de la célérité du son notée ici  $a$ . Les équations de Navier-Stokes d'un fluide compressible écrites avec les variables primitives modifiées prennent alors la forme adimensionnelle

$$\left\{ \begin{array}{l} \frac{\partial \vartheta^*}{\partial t^*} + \mathbf{u}^* \cdot \nabla_{\mathbf{x}^*} \vartheta^* = \vartheta^* \text{div}_{\mathbf{x}^*} \mathbf{u}^* \\ \frac{\partial \mathbf{u}^*}{\partial t^*} + \mathbf{u}^* \cdot \nabla_{\mathbf{x}^*} \mathbf{u}^* = -\vartheta^* \nabla_{\mathbf{x}^*} p^* + \frac{1}{\text{Re}} \vartheta^* \text{div}_{\mathbf{x}^*} \boldsymbol{\tau}_{\mu_r}(\mathbf{u}^*) \\ \frac{\partial p^*}{\partial t^*} + \mathbf{u}^* \cdot \nabla_{\mathbf{x}^*} p^* = -\gamma p^* \text{div}_{\mathbf{x}^*} \mathbf{u}^* + \frac{\gamma - 1}{\text{Re}} \nabla_{\mathbf{x}^*} \mathbf{u}^* : \boldsymbol{\tau}_{\mu_r}(\mathbf{u}^*) + \frac{\gamma}{\text{Re Pr}} \text{div}_{\mathbf{x}^*} \left[ \frac{k_{\theta}^*}{c_v^*} \nabla_{\mathbf{x}^*} (\vartheta^* p^*) \right]. \end{array} \right. \quad (5.2)$$

Les quantités étoilées dans l'expression précédente indiquent le caractère adimensionnel des différentes variables :  $\vartheta^* = \vartheta/\vartheta_0$ ,  $\mathbf{u}^* = \mathbf{u}/U_0$ ,  $p^* = p/p_0$ . Le processus d'adimensionnement concerne aussi les opérateurs de dérivation, ce qui nécessite l'emploi du temps adimensionnel  $t^* = t/t_0$  et des longueurs  $\mathbf{x}^* = \mathbf{x}/L_0$  dans les opérateurs de divergence et de gradient. Enfin, le tenseur des contraintes visqueuses  $\boldsymbol{\tau}(\mathbf{u})$  est rendu

adimensionnel en factorisant par la viscosité de référence  $\mu_r$ , ce qui d'après la définition donnée à l'équation (4.23) conduit à l'égalité

$$\boldsymbol{\tau}(\mathbf{u}) = \frac{\mu_r U_0}{L_0} \left( 2 \frac{\mu}{\mu_r} \mathbf{D}(\mathbf{x}^*, \mathbf{u}^*) + \frac{\lambda}{\mu_r} (\operatorname{div}_{\mathbf{x}^*} \mathbf{u}^*) \mathbf{I}_d \right) = \frac{\mu_r U_0}{L_0} \boldsymbol{\tau}_{\mu_r}^*(\mathbf{u}^*). \quad (5.3)$$

Le système d'équations (5.2) est simplifié en émettant quelques hypothèses supplémentaires concernant notamment les coefficients de viscosité. On suppose ainsi que l'hypothèse de Stokes est vérifiée, ce qui conduit à prendre  $\lambda = -2/3 \mu$  pour le second coefficient de viscosité. Une seconde hypothèse, beaucoup plus forte, consiste à considérer que le premier coefficient de viscosité  $\mu$  est constant et égal à  $\mu_r$ . Ceci conduit à négliger d'une part la loi de Sutherland qui fournit les variations de viscosité en fonction de la température, et d'autre part les éventuelles fluctuations de viscosité turbulente. Ces simplifications conduisent à  $\mu/\mu_r = 1$  et  $\lambda/\mu_r = -2/3$ , de sorte que l'expression du tenseur des contraintes visqueuses adimensionnel  $\boldsymbol{\tau}_{\mu_r}^*$  devient

$$\boldsymbol{\tau}_{\mu_r}^*(\mathbf{u}^*) = 2 \mathbf{D}(\mathbf{x}^*, \mathbf{u}^*) - \frac{2}{3} (\operatorname{div}_{\mathbf{x}^*} \mathbf{u}^*) \mathbf{I}_d. \quad (5.4)$$

On supposera enfin que le coefficient de conductivité thermique  $k_\theta$  est constant et égal à  $k_{\theta,r}$ , de sorte que  $k_\theta^* = 1$ . Par ailleurs, sous l'hypothèse d'un gaz parfait, il est possible de supposer que les coefficients de chaleur spécifique sont constants, si bien que  $c_v^* = 1$ . Finalement, les équations du fluide s'écrivent sous la forme du système dynamique

$$\dot{\mathbf{q}} = \mathbf{Q}^C(\mathbf{q}, \mathbf{q}) + \frac{1}{\operatorname{Re}} \mathbf{Q}^D(\mathbf{q}, \mathbf{q}), \quad (5.5)$$

où les termes quadratiques ont été scindés en deux contributions, l'une relative aux termes convectifs et l'autre aux termes diffusifs. Les notations étoilées sont à présent abandonnées pour alléger l'écriture, de même que l'indice  $\mu_r$  précisant l'adimensionnement du tenseur des contraintes; sauf mention contraire, les variables et autres quantités manipulées seront donc supposées adimensionnelles. En particulier, le vecteur  $\mathbf{q}$  fera référence au vecteur des variables adimensionnelles  $\mathbf{q}^* = [\vartheta^*, \mathbf{u}^*, p^*]$ . Les opérateurs associés à l'équation (5.2) sont alors définis par

$$\mathbf{Q}^C(\mathbf{q}, \mathbf{q}) = \begin{bmatrix} -\mathbf{u} \cdot \nabla \vartheta + \vartheta \operatorname{div} \mathbf{u} \\ -\mathbf{u} \cdot \nabla \mathbf{u} - \vartheta \nabla p \\ -\mathbf{u} \cdot \nabla p - \gamma p \operatorname{div} \mathbf{u} \end{bmatrix} \quad \text{et} \quad \mathbf{Q}^D(\mathbf{q}, \mathbf{q}) = \begin{bmatrix} 0 \\ \vartheta \operatorname{div} \boldsymbol{\tau}(\mathbf{u}) \\ (\gamma - 1) \nabla \mathbf{u} : \boldsymbol{\tau}(\mathbf{u}) + \frac{\gamma}{\operatorname{Pr}} \Delta(\vartheta p) \end{bmatrix}. \quad (5.6)$$

La forme générale du modèle d'ordre réduit se déduit immédiatement de l'équation (5.5) en y introduisant la décomposition POD vectorielle de l'équation (2.83). La bilinéarité des opérateurs quadratiques convectif  $\mathbf{Q}^C$  et diffusif  $\mathbf{Q}^D$  mène alors à l'expression

$$\begin{aligned} \sum_{j=1}^q \dot{a}_j \boldsymbol{\varphi}^{(j)} &= \mathbf{Q}^C(\bar{\mathbf{q}}, \bar{\mathbf{q}}) + \frac{1}{\operatorname{Re}} \mathbf{Q}^D(\bar{\mathbf{q}}, \bar{\mathbf{q}}) \\ &+ \sum_{j=1}^q \left[ \mathbf{Q}^C(\bar{\mathbf{q}}, \boldsymbol{\varphi}^{(j)}) + \mathbf{Q}^C(\boldsymbol{\varphi}^{(j)}, \bar{\mathbf{q}}) + \frac{1}{\operatorname{Re}} \mathbf{Q}^D(\bar{\mathbf{q}}, \boldsymbol{\varphi}^{(j)}) + \frac{1}{\operatorname{Re}} \mathbf{Q}^D(\boldsymbol{\varphi}^{(j)}, \bar{\mathbf{q}}) \right] a_j \\ &+ \sum_{j,k=1}^q \left[ \mathbf{Q}^C(\boldsymbol{\varphi}^{(j)}, \boldsymbol{\varphi}^{(k)}) + \frac{1}{\operatorname{Re}} \mathbf{Q}^D(\boldsymbol{\varphi}^{(j)}, \boldsymbol{\varphi}^{(k)}) \right] a_j a_k, \end{aligned} \quad (5.7)$$

dans laquelle les symboles de sommation ont pu être extraits. L'égalité précédente est ensuite projetée sur les  $q$  modes POD  $\boldsymbol{\varphi}^{(i)}$  à l'aide du produit scalaire défini à l'équation (4.52), pour lequel les modes POD sont orthonormaux. Le modèle d'ordre réduit est alors caractérisé par l'ensemble d'équations non-linéaires

$$\dot{a}_i = \mathcal{K}_i + \sum_{j=1}^q \mathcal{L}_{ij} a_j + \sum_{j,k=1}^q \mathcal{Q}_{jik} a_j a_k. \quad (5.8)$$

Bien que les équations du fluide écrites en variables primitives modifiées soient composées uniquement de termes quadratiques, des termes constants et linéaires apparaissent dans le modèle d'ordre réduit en

raison de la partie moyenne  $\bar{\mathbf{q}}$  introduite dans la décomposition POD affine (2.83) adoptée. Les coefficients du modèle d'ordre réduit sont donc donnés par les expressions

$$\begin{aligned} \mathcal{K}_i &= \mathcal{K}_i^C + \frac{1}{\text{Re}} \mathcal{K}_i^D & \text{avec} & \begin{cases} \mathcal{K}_i^C = \langle \mathbf{Q}^C(\bar{\mathbf{q}}, \bar{\mathbf{q}}), \varphi^{(i)} \rangle \\ \mathcal{K}_i^D = \langle \mathbf{Q}^D(\bar{\mathbf{q}}, \bar{\mathbf{q}}), \varphi^{(i)} \rangle \end{cases}, \\ \mathcal{L}_{ij} &= \mathcal{L}_{ij}^C + \frac{1}{\text{Re}} \mathcal{L}_{ij}^D & \text{avec} & \begin{cases} \mathcal{L}_{ij}^C = \langle \mathbf{Q}^C(\bar{\mathbf{q}}, \varphi^{(j)}) + \mathbf{Q}^C(\varphi^{(j)}, \bar{\mathbf{q}}), \varphi^{(i)} \rangle \\ \mathcal{L}_{ij}^D = \langle \mathbf{Q}^D(\bar{\mathbf{q}}, \varphi^{(j)}) + \mathbf{Q}^D(\varphi^{(j)}, \bar{\mathbf{q}}), \varphi^{(i)} \rangle \end{cases}, \\ \mathcal{Q}_{ijk} &= \mathcal{Q}_{ijk}^C + \frac{1}{\text{Re}} \mathcal{Q}_{ijk}^D & \text{avec} & \begin{cases} \mathcal{Q}_{ijk}^C = \langle \mathbf{Q}^C(\varphi^{(j)}, \varphi^{(k)}), \varphi^{(i)} \rangle \\ \mathcal{Q}_{ijk}^D = \langle \mathbf{Q}^D(\varphi^{(j)}, \varphi^{(k)}), \varphi^{(i)} \rangle \end{cases}. \end{aligned} \quad (5.9)$$

Le modèle d'ordre réduit est donc entièrement défini par les coefficients précédents dont le nombre s'élève à  $N_c = q + q^2 + q^3$ . Toutefois, seuls  $q^2(q+1)/2$  coefficients quadratiques sont réellement nécessaires puisqu'une partie des termes est redondante. Par conséquent le modèle d'ordre réduit est décrit par  $N'_c = q(q+1)(q+2)/2$  coefficients.

Avant de poursuivre, le modèle d'ordre réduit POD-Galerkin est écrit sous forme tensorielle afin d'alléger l'écriture et de simplifier la mise en œuvre de l'intégration temporelle qui est effectuée de façon vectorielle. Le terme quadratique qui implique un tenseur d'ordre trois peut s'écrire comme le produit tensoriel  $(\underline{\mathcal{Q}} \otimes \bar{\mathbf{a}}) \otimes \bar{\mathbf{a}}$ . En effet, d'après les relations fournies dans l'annexe D.1.3 concernant le produit tensoriel généralisé, on obtient la suite d'égalités

$$\begin{aligned} (\underline{\mathcal{Q}} \otimes \bar{\mathbf{a}}) \otimes \bar{\mathbf{a}} &= [\mathcal{Q}_{ijk} (\delta_i \otimes \delta_j \otimes \delta_k) \bar{a}_l \delta_l] \otimes a_m \delta_m \\ &= [\mathcal{Q}_{ijk} a_l \delta_{kl} (\delta_i \otimes \delta_j)] \otimes a_m \delta_m \\ &= \mathcal{Q}_{ijk} a_k a_m (\delta_i \otimes \delta_j) \otimes \delta_m \\ &= \mathcal{Q}_{ijk} a_k a_m \delta_{jm} \delta_i \\ &= \mathcal{Q}_{ijk} a_k a_j \delta_i. \end{aligned} \quad (5.10)$$

Le produit tensoriel  $(\underline{\mathcal{Q}} \otimes \bar{\mathbf{a}}) \otimes \bar{\mathbf{a}}$  est donc un vecteur dont chaque composante est donnée par  $\mathcal{Q}_{ijk} a_k a_j$ , où la convention d'Einstein sur les indices répétés est utilisée. Par conséquent les composantes de ce vecteur correspondent bien aux contributions quadratiques établies à l'équation (5.8). En toute rigueur, le modèle d'ordre réduit devrait s'écrire sous la forme

$$\dot{\mathbf{a}} = \underline{\mathcal{K}} + \underline{\mathcal{L}} \otimes \bar{\mathbf{a}} + (\underline{\underline{\mathcal{Q}}} \otimes \bar{\mathbf{a}}) \otimes \bar{\mathbf{a}} \quad (5.11)$$

en adoptant les symboles  $\otimes$  pour le produit tensoriel et en désignant l'ordre du tenseur par autant de barres sous l'opérateur. Cependant, on adoptera de préférence par la suite l'écriture simplifiée

$$\dot{\mathbf{a}} = \underline{\mathcal{K}} + \underline{\mathcal{L}} \mathbf{a} + (\underline{\underline{\mathcal{Q}}} \otimes \bar{\mathbf{a}}) \mathbf{a}. \quad (5.12)$$

Le dernier symbole du produit tensoriel contracté entre  $(\underline{\underline{\mathcal{Q}}} \otimes \bar{\mathbf{a}})$  et  $\mathbf{a}$  peut être omis puisqu'il s'agit simplement d'un produit matrice-vecteur, au même titre que celui opérant entre  $\underline{\mathcal{L}}$  et  $\mathbf{a}$ .

## 5.2.2 Calcul des coefficients du modèle d'ordre réduit

Le détail des expressions des coefficients du modèle d'ordre réduit n'est fourni que pour les termes quadratiques  $\mathcal{Q}_{ijk}$ , puisque les termes constants et linéaires sont aussi construits à partir de la projection de l'opérateur  $\mathbf{Q} = \mathbf{Q}^C + \mathbf{Q}^D/\text{Re}$  appliqué non plus au couple  $(\varphi^{(j)}, \varphi^{(k)})$  mais aux couples  $(\bar{\mathbf{q}}, \bar{\mathbf{q}})$ ,  $(\varphi^{(j)}, \bar{\mathbf{q}})$  et  $(\bar{\mathbf{q}}, \varphi^{(j)})$  respectivement.

Afin de clarifier l'exposé, les opérateurs quadratiques introduits à l'équation (5.6) sont découpés en trois parties relatives aux différentes variables en jeu, au même titre que pour les modes POD décomposés en  $\varphi^{(i)} = [\varphi_\vartheta^{(i)}, \varphi_{\mathbf{u}}^{(i)}, \varphi_p^{(i)}]^\top$ . On notera ainsi  $\mathbf{Q}^C(\mathbf{q}, \mathbf{q}) = [Q_\vartheta^C, \mathbf{Q}_{\mathbf{u}}^C, Q_p^C]^\top$  pour l'opérateur convectif, et  $\mathbf{Q}^D(\mathbf{q}, \mathbf{q}) = [Q_\vartheta^D, \mathbf{Q}_{\mathbf{u}}^D, Q_p^D]^\top$  pour l'opérateur diffusif.

### 5.2.2.1 Expression des coefficients quadratiques convectifs

Les coefficients quadratiques  $\mathcal{Q}_{ijk}^C$  ont été définis à l'équation (5.9) comme la projection de la contribution convective  $\mathbf{Q}^C(\varphi^{(j)}, \varphi^{(k)})$  sur le mode POD  $\varphi^{(i)}$ . Le développement du produit scalaire donne donc

$$\mathcal{Q}_{ijk}^C = \int_{\Omega} Q_{\vartheta}^C(\varphi^{(j)}, \varphi^{(k)}) \varphi_{\vartheta}^{(i)} d\Omega + \int_{\Omega} \mathbf{Q}_{\mathbf{u}}^C(\varphi^{(j)}, \varphi^{(k)}) \cdot \varphi_{\mathbf{u}}^{(i)} d\Omega + \int_{\Omega} Q_p^C(\varphi^{(j)}, \varphi^{(k)}) \varphi_p^{(i)} d\Omega, \quad (5.13)$$

puis en explicitant les opérateurs quadratiques, l'expression de la contribution convective devient

$$\boxed{\begin{aligned} \mathcal{Q}_{ijk}^C &= - \int_{\Omega} \left( \varphi_{\mathbf{u}}^{(j)} \cdot \nabla \varphi_{\vartheta}^{(k)} - \varphi_{\vartheta}^{(k)} \operatorname{div} \varphi_{\mathbf{u}}^{(j)} \right) \varphi_{\vartheta}^{(i)} d\Omega \\ &\quad - \int_{\Omega} \left( \varphi_{\mathbf{u}}^{(j)} \cdot \nabla \varphi_{\mathbf{u}}^{(k)} + \varphi_{\vartheta}^{(j)} \nabla \varphi_p^{(k)} \right) \cdot \varphi_{\mathbf{u}}^{(i)} d\Omega \\ &\quad - \int_{\Omega} \left( \varphi_{\mathbf{u}}^{(j)} \cdot \nabla \varphi_p^{(k)} + \gamma \varphi_p^{(j)} \operatorname{div} \varphi_{\mathbf{u}}^{(k)} \right) \varphi_p^{(i)} d\Omega \end{aligned}}. \quad (5.14)$$

Une forme encore plus détaillée des coefficients  $\mathcal{Q}_{ijk}^C$  s'obtient en développant les opérateurs tensoriels dans la base cartésienne orthonormée directe  $\mathcal{B}_A$  associée aux directions  $x, y$  et  $z$ . Les coefficients sont alors donnés par l'expression

$$\boxed{\begin{aligned} \mathcal{Q}_{ijk}^C &= - \int_{\Omega} \left[ \varphi_u^{(j)} \frac{\partial \varphi_{\vartheta}^{(k)}}{\partial x} + \varphi_v^{(j)} \frac{\partial \varphi_{\vartheta}^{(k)}}{\partial y} + \varphi_w^{(j)} \frac{\partial \varphi_{\vartheta}^{(k)}}{\partial z} - \varphi_{\vartheta}^{(k)} \left( \frac{\partial \varphi_u^{(j)}}{\partial x} + \frac{\partial \varphi_v^{(j)}}{\partial y} + \frac{\partial \varphi_w^{(j)}}{\partial z} \right) \right] \varphi_{\vartheta}^{(i)} d\Omega \\ &\quad - \int_{\Omega} \left[ \varphi_u^{(j)} \frac{\partial \varphi_u^{(k)}}{\partial x} + \varphi_v^{(j)} \frac{\partial \varphi_u^{(k)}}{\partial y} + \varphi_w^{(j)} \frac{\partial \varphi_u^{(k)}}{\partial z} + \varphi_{\vartheta}^{(j)} \frac{\partial \varphi_p^{(k)}}{\partial x} \right] \varphi_{\mathbf{u}}^{(i)} d\Omega \\ &\quad - \int_{\Omega} \left[ \varphi_u^{(j)} \frac{\partial \varphi_v^{(k)}}{\partial x} + \varphi_v^{(j)} \frac{\partial \varphi_v^{(k)}}{\partial y} + \varphi_w^{(j)} \frac{\partial \varphi_v^{(k)}}{\partial z} + \varphi_{\vartheta}^{(j)} \frac{\partial \varphi_p^{(k)}}{\partial y} \right] \varphi_v^{(i)} d\Omega \\ &\quad - \int_{\Omega} \left[ \varphi_u^{(j)} \frac{\partial \varphi_w^{(k)}}{\partial x} + \varphi_v^{(j)} \frac{\partial \varphi_w^{(k)}}{\partial y} + \varphi_w^{(j)} \frac{\partial \varphi_w^{(k)}}{\partial z} + \varphi_{\vartheta}^{(j)} \frac{\partial \varphi_p^{(k)}}{\partial z} \right] \varphi_w^{(i)} d\Omega \\ &\quad - \int_{\Omega} \left[ \varphi_u^{(j)} \frac{\partial \varphi_p^{(k)}}{\partial x} + \varphi_v^{(j)} \frac{\partial \varphi_p^{(k)}}{\partial y} + \varphi_w^{(j)} \frac{\partial \varphi_p^{(k)}}{\partial z} + \gamma \varphi_p^{(j)} \left( \frac{\partial \varphi_u^{(k)}}{\partial x} + \frac{\partial \varphi_v^{(k)}}{\partial y} + \frac{\partial \varphi_w^{(k)}}{\partial z} \right) \right] \varphi_p^{(i)} d\Omega \end{aligned}}. \quad (5.15)$$

Le calcul des coefficients quadratiques convectifs nécessite donc seulement de disposer d'une méthode de calcul des gradients des modes POD  $\varphi^{(i)}$  et de la partie stationnaire  $\bar{\mathbf{q}} = [\bar{\vartheta}, \bar{\mathbf{u}}, \bar{p}]^T$  qui intervient dans les coefficients constants et linéaires. Le calcul des gradients et des intégrales volumiques s'effectuera en suivant une approche Volumes Finis cohérente avec celle employée dans *elsA* pour discrétiser les champs aérodynamiques (voir l'annexe E.3.2 p.265).

### 5.2.2.2 Expression des coefficients quadratiques diffusifs

Comme précédemment, les coefficients quadratiques diffusifs  $\mathcal{Q}_{ijk}^D$  définis à l'équation (5.9) comme la projection de la contribution diffusive  $\mathbf{Q}^D(\varphi^{(j)}, \varphi^{(k)})$  sur le mode POD  $\varphi^{(i)}$  sont détaillés en développant le produit scalaire en

$$\mathcal{Q}_{ijk}^D = \int_{\Omega} Q_{\vartheta}^D(\varphi^{(j)}, \varphi^{(k)}) \varphi_{\vartheta}^{(i)} d\Omega + \int_{\Omega} \mathbf{Q}_{\mathbf{u}}^D(\varphi^{(j)}, \varphi^{(k)}) \cdot \varphi_{\mathbf{u}}^{(i)} d\Omega + \int_{\Omega} Q_p^D(\varphi^{(j)}, \varphi^{(k)}) \varphi_p^{(i)} d\Omega, \quad (5.16)$$

puis en explicitant les opérateurs quadratiques pour aboutir à l'expression

$$\boxed{\begin{aligned} \mathcal{Q}_{ijk}^D &= \int_{\Omega} \varphi_{\vartheta}^{(j)} \operatorname{div} \boldsymbol{\tau}(\varphi_{\mathbf{u}}^{(k)}) \cdot \varphi_{\mathbf{u}}^{(i)} d\Omega \\ &\quad + \int_{\Omega} \left[ (\gamma - 1) \nabla \varphi_{\mathbf{u}}^{(j)} : \boldsymbol{\tau}(\varphi_{\mathbf{u}}^{(k)}) + \frac{\gamma}{\operatorname{Pr}} \operatorname{div} \left( \nabla \left( \varphi_{\vartheta}^{(j)} \varphi_p^{(k)} \right) \right) \right] \varphi_p^{(i)} d\Omega \end{aligned}}. \quad (5.17)$$

Enfin, le développement des opérateurs tensoriels dans la base  $\mathcal{B}_A$  conduit à l'expression

$$\begin{aligned}
 \mathcal{Q}_{ijk}^D = & \int_{\Omega} \varphi_{\vartheta}^{(j)} \left[ \frac{4}{3} \frac{\partial^2 \varphi_u^{(k)}}{\partial x^2} + \frac{\partial^2 \varphi_u^{(k)}}{\partial y^2} + \frac{\partial^2 \varphi_u^{(k)}}{\partial z^2} + \frac{1}{3} \frac{\partial^2 \varphi_v^{(k)}}{\partial x \partial y} + \frac{1}{3} \frac{\partial^2 \varphi_w^{(k)}}{\partial x \partial z} \right] \varphi_u^{(i)} \, d\Omega \\
 & + \int_{\Omega} \varphi_{\vartheta}^{(j)} \left[ \frac{\partial^2 \varphi_v^{(k)}}{\partial x^2} + \frac{4}{3} \frac{\partial^2 \varphi_v^{(k)}}{\partial y^2} + \frac{\partial^2 \varphi_v^{(k)}}{\partial z^2} + \frac{1}{3} \frac{\partial^2 \varphi_u^{(k)}}{\partial y \partial x} + \frac{1}{3} \frac{\partial^2 \varphi_w^{(k)}}{\partial y \partial z} \right] \varphi_v^{(i)} \, d\Omega \\
 & + \int_{\Omega} \varphi_{\vartheta}^{(j)} \left[ \frac{\partial^2 \varphi_w^{(k)}}{\partial x^2} + \frac{\partial^2 \varphi_w^{(k)}}{\partial y^2} + \frac{4}{3} \frac{\partial^2 \varphi_w^{(k)}}{\partial z^2} + \frac{1}{3} \frac{\partial^2 \varphi_u^{(k)}}{\partial z \partial x} + \frac{1}{3} \frac{\partial^2 \varphi_v^{(k)}}{\partial z \partial y} \right] \varphi_w^{(i)} \, d\Omega \\
 & + (\gamma - 1) \int_{\Omega} \left[ \frac{2}{3} \left( 2 \frac{\partial \varphi_u^{(k)}}{\partial x} - \frac{\partial \varphi_v^{(k)}}{\partial y} - \frac{\partial \varphi_w^{(k)}}{\partial z} \right) \frac{\partial \varphi_u^{(j)}}{\partial x} + \left( \frac{\partial \varphi_u^{(k)}}{\partial y} + \frac{\partial \varphi_v^{(k)}}{\partial x} \right) \frac{\partial \varphi_u^{(j)}}{\partial y} + \left( \frac{\partial \varphi_u^{(k)}}{\partial z} + \frac{\partial \varphi_w^{(k)}}{\partial x} \right) \frac{\partial \varphi_u^{(j)}}{\partial z} \right] \varphi_p^{(i)} \, d\Omega \\
 & + (\gamma - 1) \int_{\Omega} \left[ \frac{2}{3} \left( 2 \frac{\partial \varphi_v^{(k)}}{\partial y} - \frac{\partial \varphi_u^{(k)}}{\partial x} - \frac{\partial \varphi_w^{(k)}}{\partial z} \right) \frac{\partial \varphi_v^{(j)}}{\partial y} + \left( \frac{\partial \varphi_v^{(k)}}{\partial x} + \frac{\partial \varphi_u^{(k)}}{\partial y} \right) \frac{\partial \varphi_v^{(j)}}{\partial x} + \left( \frac{\partial \varphi_v^{(k)}}{\partial z} + \frac{\partial \varphi_w^{(k)}}{\partial y} \right) \frac{\partial \varphi_v^{(j)}}{\partial z} \right] \varphi_p^{(i)} \, d\Omega \\
 & + (\gamma - 1) \int_{\Omega} \left[ \frac{2}{3} \left( 2 \frac{\partial \varphi_w^{(k)}}{\partial z} - \frac{\partial \varphi_u^{(k)}}{\partial x} - \frac{\partial \varphi_v^{(k)}}{\partial y} \right) \frac{\partial \varphi_w^{(j)}}{\partial z} + \left( \frac{\partial \varphi_w^{(k)}}{\partial x} + \frac{\partial \varphi_u^{(k)}}{\partial z} \right) \frac{\partial \varphi_w^{(j)}}{\partial x} + \left( \frac{\partial \varphi_w^{(k)}}{\partial y} + \frac{\partial \varphi_v^{(k)}}{\partial z} \right) \frac{\partial \varphi_w^{(j)}}{\partial y} \right] \varphi_p^{(i)} \, d\Omega \\
 & + \frac{\gamma}{\text{Pr}} \int_{\Omega} \varphi_{\vartheta}^{(j)} \left( \frac{\partial^2 \varphi_p^{(k)}}{\partial x^2} + \frac{\partial^2 \varphi_p^{(k)}}{\partial y^2} + \frac{\partial^2 \varphi_p^{(k)}}{\partial z^2} \right) \varphi_p^{(i)} \, d\Omega \\
 & + \frac{\gamma}{\text{Pr}} \int_{\Omega} \varphi_p^{(k)} \left( \frac{\partial^2 \varphi_{\vartheta}^{(j)}}{\partial x^2} + \frac{\partial^2 \varphi_{\vartheta}^{(j)}}{\partial y^2} + \frac{\partial^2 \varphi_{\vartheta}^{(j)}}{\partial z^2} \right) \varphi_p^{(i)} \, d\Omega \\
 & + \frac{\gamma}{\text{Pr}} \int_{\Omega} 2 \left( \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial x} \frac{\partial \varphi_p^{(k)}}{\partial x} + \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial y} \frac{\partial \varphi_p^{(k)}}{\partial y} + \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial z} \frac{\partial \varphi_p^{(k)}}{\partial z} \right) \varphi_p^{(i)} \, d\Omega
 \end{aligned} \tag{5.18}$$

L'expression des termes diffusifs fait apparaître des dérivées secondes caractéristiques des processus de diffusion, ainsi que des dérivées croisées. Il est donc judicieux de réduire l'ordre des dérivées apparaissant dans le calcul des coefficients en procédant à des intégrations par parties. Seuls deux groupes de contributions nécessitent d'évaluer des dérivées secondes et mixtes des champs vectoriels. Ces dérivées proviennent précisément de la projection du terme  $\mathbf{Q}_u^D$  sur la contribution du mode POD  $\varphi_u^{(i)}$  associée, et de la projection de  $Q_p^D$  sur  $\varphi_p^{(i)}$ . On note alors  $\tilde{Q}_u^D = \int_{\Omega} \varphi_{\vartheta}^{(j)} \operatorname{div} \boldsymbol{\tau}(\boldsymbol{\varphi}_u^{(k)}) \cdot \boldsymbol{\varphi}_u^{(i)} \, d\Omega$  la contribution du terme de vitesse,  $\tilde{Q}_{p,1}^D = \int_{\Omega} \nabla \varphi_u^{(j)} : \boldsymbol{\tau}(\boldsymbol{\varphi}_u^{(k)}) \varphi_p^{(i)} \, d\Omega$  la première contribution de la projection du terme de pression et enfin  $\tilde{Q}_{p,2}^D = \int_{\Omega} \Delta(\varphi_{\vartheta}^{(j)} \varphi_p^{(k)}) \varphi_p^{(i)} \, d\Omega$  la seconde contribution. Ainsi, le coefficient diffusif défini à l'équation (5.17) se décompose en  $\mathcal{Q}_{ijk}^D = \tilde{Q}_u^D + (\gamma - 1) \tilde{Q}_{p,1}^D + \gamma/\text{Pr} \tilde{Q}_{p,2}^D$ .

En supposant que le domaine  $\Omega$  possède une frontière  $\partial\Omega$  assez régulière pour qu'en tout point de  $\partial\Omega$ , la normale extérieure  $\mathbf{n} = [n_x, n_y, n_z]^T$  puisse être définie, il est alors possible d'utiliser les théorèmes d'intégration par parties établis dans l'annexe D.4. Il suffit en effet d'appliquer la formule d'intégration par parties (D.32) pour le terme  $\tilde{Q}_u^D$  avec  $p = \varphi_{\vartheta}^{(j)}$ ,  $\mathbf{D} = \boldsymbol{\tau}(\boldsymbol{\varphi}_u^{(k)})$  et  $\mathbf{u} = \boldsymbol{\varphi}_u^{(i)}$ , ce qui conduit à

$$\begin{aligned}
 \tilde{Q}_u^D &= \int_{\Omega} \varphi_{\vartheta}^{(j)} \operatorname{div} \boldsymbol{\tau}(\boldsymbol{\varphi}_u^{(k)}) \cdot \boldsymbol{\varphi}_u^{(i)} \, d\Omega \\
 &= \oint_{\partial\Omega} \left[ \varphi_{\vartheta}^{(j)} (\boldsymbol{\varphi}_u^{(i)})^T \boldsymbol{\tau}(\boldsymbol{\varphi}_u^{(k)}) \right] \mathbf{n} \, d\partial\Omega \\
 &\quad - \int_{\Omega} \varphi_{\vartheta}^{(j)} \nabla \boldsymbol{\varphi}_u^{(i)} : \boldsymbol{\tau}(\boldsymbol{\varphi}_u^{(k)}) \, d\Omega - \int_{\Omega} (\boldsymbol{\varphi}_u^{(i)} \otimes \nabla \varphi_{\vartheta}^{(j)}) : \boldsymbol{\tau}(\boldsymbol{\varphi}_u^{(k)}) \, d\Omega.
 \end{aligned} \tag{5.19}$$

Par analogie, la formule d'intégration par parties (D.34) s'applique au terme  $\tilde{Q}_{p,2}^D$  en notant que pour tout scalaire  $s$ ,  $\Delta s = \operatorname{div}(\nabla s)$ . En posant  $p = \varphi_p^{(i)}$ ,  $q = \varphi_{\vartheta}^{(j)}$  et  $r = \varphi_p^{(k)}$  dans (D.34) on obtient

$$\begin{aligned}
 \tilde{Q}_{p,2}^D &= \int_{\Omega} \varphi_p^{(i)} \operatorname{div} \left[ \nabla (\varphi_{\vartheta}^{(j)} \varphi_p^{(k)}) \right] \, d\Omega \\
 &= \oint_{\partial\Omega} \varphi_p^{(i)} \nabla (\varphi_{\vartheta}^{(j)} \varphi_p^{(k)}) \cdot \mathbf{n} \, d\partial\Omega - \int_{\Omega} \varphi_{\vartheta}^{(j)} \nabla \varphi_p^{(k)} \cdot \nabla \varphi_p^{(i)} \, d\Omega - \int_{\Omega} \varphi_p^{(k)} \nabla \varphi_{\vartheta}^{(j)} \cdot \nabla \varphi_p^{(i)} \, d\Omega.
 \end{aligned} \tag{5.20}$$



Les relations précédentes ne font plus apparaître maintenant que des dérivées premières qui interviennent dans des intégrales surfaciques et volumiques. Le terme diffusif est donc de préférence scindé en deux parties, l'une notée  $\mathcal{Q}_{\text{Vol}}^D$  comprenant les termes d'intégration sur le domaine  $\Omega$ , et l'autre notée  $\mathcal{Q}_{\text{Surf}}^D$  regroupant les termes intégrés sur le bord  $\partial\Omega$  du domaine, de manière à ce que  $\mathcal{Q}_{ijk}^D = \mathcal{Q}_{\text{Vol}}^D + \mathcal{Q}_{\text{Surf}}^D$ . Après regroupement des différents termes, les contributions volumiques et surfaciques s'écrivent

$$\begin{aligned}
 \mathcal{Q}_{\text{Vol}}^D &= - \int_{\Omega} \varphi_{\vartheta}^{(j)} \nabla \varphi_{\mathbf{u}}^{(i)} : \boldsymbol{\tau}(\varphi_{\mathbf{u}}^{(k)}) \, d\Omega - \int_{\Omega} \left( \varphi_{\mathbf{u}}^{(i)} \otimes \nabla \varphi_{\vartheta}^{(j)} \right) : \boldsymbol{\tau}(\varphi_{\mathbf{u}}^{(k)}) \, d\Omega \\
 &\quad + (\gamma - 1) \int_{\Omega} \nabla \varphi_{\mathbf{u}}^{(j)} : \boldsymbol{\tau}(\varphi_{\mathbf{u}}^{(k)}) \varphi_p^{(i)} \, d\Omega \\
 &\quad - \frac{\gamma}{\text{Pr}} \int_{\Omega} \varphi_{\vartheta}^{(j)} \nabla \varphi_p^{(k)} \cdot \nabla \varphi_p^{(i)} \, d\Omega - \frac{\gamma}{\text{Pr}} \int_{\Omega} \varphi_p^{(k)} \nabla \varphi_{\vartheta}^{(j)} \cdot \nabla \varphi_p^{(i)} \, d\Omega, \\
 \mathcal{Q}_{\text{Surf}}^D &= \oint_{\partial\Omega} \left[ \varphi_{\vartheta}^{(j)} \left( \varphi_{\mathbf{u}}^{(i)} \right)^{\top} \boldsymbol{\tau}(\varphi_{\mathbf{u}}^{(k)}) \right] \mathbf{n} \, d\partial\Omega + \frac{\gamma}{\text{Pr}} \oint_{\partial\Omega} \varphi_p^{(i)} \nabla \left( \varphi_{\vartheta}^{(j)} \varphi_p^{(k)} \right) \cdot \mathbf{n} \, d\partial\Omega.
 \end{aligned} \tag{5.21}$$

La contribution volumique participant aux termes diffusifs se développe donc en

$$\begin{aligned}
 \mathcal{Q}_{\text{Vol}}^D &= - \int_{\Omega} \varphi_{\vartheta}^{(j)} \left[ \frac{2}{3} \left( 2 \frac{\partial \varphi_u^{(k)}}{\partial x} - \frac{\partial \varphi_v^{(k)}}{\partial y} - \frac{\partial \varphi_w^{(k)}}{\partial z} \right) \frac{\partial \varphi_u^{(i)}}{\partial x} + \left( \frac{\partial \varphi_u^{(k)}}{\partial y} + \frac{\partial \varphi_v^{(k)}}{\partial x} \right) \frac{\partial \varphi_u^{(i)}}{\partial y} + \left( \frac{\partial \varphi_u^{(k)}}{\partial z} + \frac{\partial \varphi_w^{(k)}}{\partial x} \right) \frac{\partial \varphi_u^{(i)}}{\partial z} \right] \, d\Omega \\
 &\quad - \int_{\Omega} \varphi_{\vartheta}^{(j)} \left[ \frac{2}{3} \left( 2 \frac{\partial \varphi_v^{(k)}}{\partial y} - \frac{\partial \varphi_u^{(k)}}{\partial x} - \frac{\partial \varphi_w^{(k)}}{\partial z} \right) \frac{\partial \varphi_v^{(i)}}{\partial y} + \left( \frac{\partial \varphi_v^{(k)}}{\partial x} + \frac{\partial \varphi_u^{(k)}}{\partial y} \right) \frac{\partial \varphi_v^{(i)}}{\partial x} + \left( \frac{\partial \varphi_v^{(k)}}{\partial z} + \frac{\partial \varphi_w^{(k)}}{\partial y} \right) \frac{\partial \varphi_v^{(i)}}{\partial z} \right] \, d\Omega \\
 &\quad - \int_{\Omega} \varphi_{\vartheta}^{(j)} \left[ \frac{2}{3} \left( 2 \frac{\partial \varphi_w^{(k)}}{\partial z} - \frac{\partial \varphi_u^{(k)}}{\partial x} - \frac{\partial \varphi_v^{(k)}}{\partial y} \right) \frac{\partial \varphi_w^{(i)}}{\partial z} + \left( \frac{\partial \varphi_w^{(k)}}{\partial x} + \frac{\partial \varphi_u^{(k)}}{\partial z} \right) \frac{\partial \varphi_w^{(i)}}{\partial x} + \left( \frac{\partial \varphi_w^{(k)}}{\partial y} + \frac{\partial \varphi_v^{(k)}}{\partial z} \right) \frac{\partial \varphi_w^{(i)}}{\partial y} \right] \, d\Omega \\
 &\quad - \int_{\Omega} \varphi_u^{(i)} \left[ \frac{2}{3} \left( 2 \frac{\partial \varphi_u^{(k)}}{\partial x} - \frac{\partial \varphi_v^{(k)}}{\partial y} - \frac{\partial \varphi_w^{(k)}}{\partial z} \right) \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial x} + \left( \frac{\partial \varphi_u^{(k)}}{\partial y} + \frac{\partial \varphi_v^{(k)}}{\partial x} \right) \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial y} + \left( \frac{\partial \varphi_u^{(k)}}{\partial z} + \frac{\partial \varphi_w^{(k)}}{\partial x} \right) \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial z} \right] \, d\Omega \\
 &\quad - \int_{\Omega} \varphi_v^{(i)} \left[ \frac{2}{3} \left( 2 \frac{\partial \varphi_v^{(k)}}{\partial y} - \frac{\partial \varphi_u^{(k)}}{\partial x} - \frac{\partial \varphi_w^{(k)}}{\partial z} \right) \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial y} + \left( \frac{\partial \varphi_v^{(k)}}{\partial x} + \frac{\partial \varphi_u^{(k)}}{\partial y} \right) \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial x} + \left( \frac{\partial \varphi_v^{(k)}}{\partial z} + \frac{\partial \varphi_w^{(k)}}{\partial y} \right) \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial z} \right] \, d\Omega \\
 &\quad - \int_{\Omega} \varphi_w^{(i)} \left[ \frac{2}{3} \left( 2 \frac{\partial \varphi_w^{(k)}}{\partial z} - \frac{\partial \varphi_u^{(k)}}{\partial x} - \frac{\partial \varphi_v^{(k)}}{\partial y} \right) \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial z} + \left( \frac{\partial \varphi_w^{(k)}}{\partial x} + \frac{\partial \varphi_u^{(k)}}{\partial z} \right) \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial x} + \left( \frac{\partial \varphi_w^{(k)}}{\partial y} + \frac{\partial \varphi_v^{(k)}}{\partial z} \right) \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial y} \right] \, d\Omega \\
 &\quad + (\gamma - 1) \int_{\Omega} \left[ \frac{2}{3} \left( 2 \frac{\partial \varphi_u^{(k)}}{\partial x} - \frac{\partial \varphi_v^{(k)}}{\partial y} - \frac{\partial \varphi_w^{(k)}}{\partial z} \right) \frac{\partial \varphi_u^{(j)}}{\partial x} + \left( \frac{\partial \varphi_u^{(k)}}{\partial y} + \frac{\partial \varphi_v^{(k)}}{\partial x} \right) \frac{\partial \varphi_u^{(j)}}{\partial y} + \left( \frac{\partial \varphi_u^{(k)}}{\partial z} + \frac{\partial \varphi_w^{(k)}}{\partial x} \right) \frac{\partial \varphi_u^{(j)}}{\partial z} \right] \varphi_p^{(i)} \, d\Omega \\
 &\quad + (\gamma - 1) \int_{\Omega} \left[ \frac{2}{3} \left( 2 \frac{\partial \varphi_v^{(k)}}{\partial y} - \frac{\partial \varphi_u^{(k)}}{\partial x} - \frac{\partial \varphi_w^{(k)}}{\partial z} \right) \frac{\partial \varphi_v^{(j)}}{\partial y} + \left( \frac{\partial \varphi_v^{(k)}}{\partial x} + \frac{\partial \varphi_u^{(k)}}{\partial y} \right) \frac{\partial \varphi_v^{(j)}}{\partial x} + \left( \frac{\partial \varphi_v^{(k)}}{\partial z} + \frac{\partial \varphi_w^{(k)}}{\partial y} \right) \frac{\partial \varphi_v^{(j)}}{\partial z} \right] \varphi_p^{(i)} \, d\Omega \\
 &\quad + (\gamma - 1) \int_{\Omega} \left[ \frac{2}{3} \left( 2 \frac{\partial \varphi_w^{(k)}}{\partial z} - \frac{\partial \varphi_u^{(k)}}{\partial x} - \frac{\partial \varphi_v^{(k)}}{\partial y} \right) \frac{\partial \varphi_w^{(j)}}{\partial z} + \left( \frac{\partial \varphi_w^{(k)}}{\partial x} + \frac{\partial \varphi_u^{(k)}}{\partial z} \right) \frac{\partial \varphi_w^{(j)}}{\partial x} + \left( \frac{\partial \varphi_w^{(k)}}{\partial y} + \frac{\partial \varphi_v^{(k)}}{\partial z} \right) \frac{\partial \varphi_w^{(j)}}{\partial y} \right] \varphi_p^{(i)} \, d\Omega \\
 &\quad - \frac{\gamma}{\text{Pr}} \int_{\Omega} \varphi_{\vartheta}^{(j)} \left( \frac{\partial \varphi_p^{(k)}}{\partial x} \frac{\partial \varphi_p^{(i)}}{\partial x} + \frac{\partial \varphi_p^{(k)}}{\partial y} \frac{\partial \varphi_p^{(i)}}{\partial y} + \frac{\partial \varphi_p^{(k)}}{\partial z} \frac{\partial \varphi_p^{(i)}}{\partial z} \right) \, d\Omega \\
 &\quad - \frac{\gamma}{\text{Pr}} \int_{\Omega} \varphi_p^{(k)} \left( \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial x} \frac{\partial \varphi_p^{(i)}}{\partial x} + \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial y} \frac{\partial \varphi_p^{(i)}}{\partial y} + \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial z} \frac{\partial \varphi_p^{(i)}}{\partial z} \right) \, d\Omega
 \end{aligned} \tag{5.22}$$

Pour parachever la définition des coefficients du modèle d'ordre réduit, il ne reste plus qu'à expliciter l'expression de la contribution surfacique participant aux termes diffusifs. Celle-ci se développe à l'aide du vecteur de la normale extérieure  $\mathbf{n} = [n_x, n_y, n_z]^{\top}$  en

$$\begin{aligned}
 Q_{\text{Surf}}^D = & \int_{\partial\Omega} \varphi_{\vartheta}^{(j)} \left[ \frac{2}{3} \left( 2 \frac{\partial \varphi_u^{(k)}}{\partial x} - \frac{\partial \varphi_v^{(k)}}{\partial y} - \frac{\partial \varphi_w^{(k)}}{\partial z} \right) n_x + \left( \frac{\partial \varphi_u^{(k)}}{\partial y} + \frac{\partial \varphi_v^{(k)}}{\partial x} \right) n_y + \left( \frac{\partial \varphi_u^{(k)}}{\partial z} + \frac{\partial \varphi_w^{(k)}}{\partial x} \right) n_z \right] \varphi_u^{(i)} d\partial\Omega \\
 & + \int_{\partial\Omega} \varphi_{\vartheta}^{(j)} \left[ \frac{2}{3} \left( 2 \frac{\partial \varphi_v^{(k)}}{\partial y} - \frac{\partial \varphi_u^{(k)}}{\partial x} - \frac{\partial \varphi_w^{(k)}}{\partial z} \right) n_y + \left( \frac{\partial \varphi_v^{(k)}}{\partial x} + \frac{\partial \varphi_u^{(k)}}{\partial y} \right) n_x + \left( \frac{\partial \varphi_v^{(k)}}{\partial z} + \frac{\partial \varphi_w^{(k)}}{\partial y} \right) n_z \right] \varphi_v^{(i)} d\partial\Omega \\
 & + \int_{\partial\Omega} \varphi_{\vartheta}^{(j)} \left[ \frac{2}{3} \left( 2 \frac{\partial \varphi_w^{(k)}}{\partial z} - \frac{\partial \varphi_u^{(k)}}{\partial x} - \frac{\partial \varphi_v^{(k)}}{\partial y} \right) n_z + \left( \frac{\partial \varphi_w^{(k)}}{\partial x} + \frac{\partial \varphi_u^{(k)}}{\partial z} \right) n_x + \left( \frac{\partial \varphi_w^{(k)}}{\partial y} + \frac{\partial \varphi_v^{(k)}}{\partial z} \right) n_y \right] \varphi_w^{(i)} d\partial\Omega \\
 & + \frac{\gamma}{\text{Pr}} \int_{\partial\Omega} \varphi_p^{(i)} \varphi_{\vartheta}^{(j)} \left( \frac{\partial \varphi_p^{(k)}}{\partial x} n_x + \frac{\partial \varphi_p^{(k)}}{\partial y} n_y + \frac{\partial \varphi_p^{(k)}}{\partial z} n_z \right) d\partial\Omega \\
 & + \frac{\gamma}{\text{Pr}} \int_{\partial\Omega} \varphi_p^{(i)} \varphi_p^{(k)} \left( \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial x} n_x + \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial y} n_y + \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial z} n_z \right) d\partial\Omega
 \end{aligned} \quad (5.23)$$

L'utilisation d'une intégration par parties a permis de réduire l'ordre des dérivées qu'il est nécessaire d'évaluer pour calculer les coefficients du modèle d'ordre réduit. Cette approche introduit un plus grand nombre de termes à calculer, mais l'avantage principal est que leurs expressions ne font appel qu'aux gradients des modes POD et de la partie moyenne des clichés qui peuvent être évalués de manière satisfaisante avec les méthodes de calcul proposées dans l'annexe E.3.2.

Le modèle d'ordre réduit de l'écoulement compressible est donc constitué de l'équation (5.8) qui reproduit la dynamique du système. Les opérateurs sont caractérisés par les coefficients (5.9), eux-mêmes étant entièrement définis par les expressions des termes convectifs (5.15) et des contributions volumiques (5.22) et surfaciques (5.23) participant aux termes diffusifs.

Signalons enfin qu'en fonction des conditions aux limites choisies dans le problème fluide, les termes de bord s'annulent si les quantités physiques ou leur gradient projeté sur la normale sortante sont nuls sur le bord  $\partial\Omega$  puisque les modes héritent des propriétés linéaires homogènes comme on l'a vu au paragraphe 2.2.3.4. Il est donc fort probable qu'une partie des termes de bord  $Q_{\text{Surf}}^D$  s'annule et c'est la raison pour laquelle Iollo *et al.* [117] ont choisi dans leur modèle de négliger les termes surfaciques issus de l'intégration par parties.

### 5.2.3 Intégration du modèle d'ordre réduit en temps

Le modèle d'ordre réduit POD-Galerkin s'écrit sous la forme d'un système d'équations différentielles ordinaires non-linéaires à coefficients constants d'ordre un en temps, défini par l'équation (5.12). Le système doit ensuite être complété par un ensemble de  $q$  conditions initiales portant sur les amplitudes modales  $\{a_i\}_{i=1}^q$  afin que le problème (5.12) soit bien posé. Le modèle d'ordre réduit est par exemple initialisé avec la solution stationnaire de l'écoulement ou encore avec le premier cliché puisque celui-ci est solution du problème instationnaire. Cette solution initiale du modèle non réduit notée  $\mathbf{q}_0$  est alors supposée être solution du modèle d'ordre réduit à l'instant initial  $t = t_0$  et se décompose alors sur la base des modes POD en  $\mathbf{q}_0 = \mathbf{q}(t = t_0) = \bar{\mathbf{q}} + \sum_{j=1}^q a_j(t = t_0) \varphi^{(j)}$ . Les amplitudes modales à l'instant initial s'obtiennent alors par projection de la relation précédente sur les modes POD et les propriétés d'orthogonalité des modes POD conduisent à

$$a_i(t = t_0) = \langle \mathbf{q}_0 - \bar{\mathbf{q}}, \varphi^{(i)} \rangle. \quad (5.24)$$

Le système (5.12) complété par les conditions initiales (5.24) est à présent bien posé puisque le modèle d'ordre réduit ne dépend que du temps. Le système dynamique réduit a donc la forme  $\dot{\mathbf{a}} = \hat{\mathbf{f}}(\mathbf{a}, t)$  et s'intègre au moyen du solveur `lsoda` de la librairie `ODEPACK` (voir la remarque 3.1 p.73). Il est important de noter que le nombre CFL imposé pour résoudre en temps le modèle d'ordre réduit peut être largement augmenté par rapport à celui du modèle original. La raison en est que la méthode POD peut être vue comme un filtre passe-bas [117] éliminant les hautes fréquences qui habituellement nécessitent de travailler avec un CFL petit.

## 5.3 Correction des instabilités des modèles d'ordre réduit POD-Galerkin

### 5.3.1 Origine des instabilités et principe de la correction

Malgré les précautions prises lors de la construction du modèle d'ordre réduit, la réponse souffre souvent d'un manque de stabilité qui se traduit sous diverses formes allant d'un faible excédent sur l'amplitude d'oscillations entretenues à la divergence rapide de la réponse. L'existence de termes non-linéaires dans le modèle d'ordre réduit est une des sources d'instabilité puisqu'une perturbation infinitésimale des coefficients des opérateurs du modèle d'ordre réduit est susceptible de conduire à des comportements très différents de la réponse. Le système dynamique de Lorenz [151] en est un bon exemple.

En ce qui concerne les modèles d'ordre réduit POD-Galerkin, le caractère instable a été observé et analysé en particulier par Rempfer [188] et Noack *et al.* [171]. Rempfer rappelait que pour fonctionner correctement, les fonctions de base — en l'occurrence ici les modes POD — employées dans la projection de Galerkin doivent (i) engendrer entièrement l'espace des phases de l'équation originale, (ii) être linéairement indépendantes pour assurer l'unicité de la solution et (iii) vérifier les conditions limites du problème. Par ailleurs Rempfer a aussi attiré l'attention sur le fait que le champ initial doit se trouver sur la trajectoire de l'écoulement original. Dans le contexte de la décomposition orthogonale propre, seule la condition (ii) est en pratique vérifiée. Rempfer a alors construit un exemple dans lequel l'ensemble des conditions est vérifié en dehors de celle portant sur les conditions aux limites et a montré que le non-respect de cette seule condition suffit à déstabiliser la réponse. Le même type d'analyse mené par Noack *et al.* [171] a aussi montré que le modèle d'ordre réduit conduisait à un mauvais attracteur. Enfin Iollo [116] s'est intéressé aussi aux problèmes de stabilité des systèmes réduits POD-Galerkin pour l'équation d'Euler et a montré que la stabilité peut être maintenue par rapport au système original non réduit en effectuant un choix judicieux pour les schémas de discrétisation des flux.

Pour résumer, l'instabilité régulièrement observée avec les modèles d'ordre réduit POD-Galerkin provient principalement des trois causes suivantes :

- la construction de la base des modes POD qui n'engendrent pas totalement l'espace des phases<sup>1</sup> et ne satisfont pas les mêmes conditions aux limites que celles du modèle initial ;
- les erreurs de modélisation qui résultent d'hypothèses simplificatrices émises pour la construction du modèle d'ordre réduit, comme par exemple le fait de supposer la viscosité constante ;
- les erreurs d'approximation numérique à l'origine de perturbations infinitésimales dans l'évaluation des coefficients et qui sont potentiellement suffisantes pour déstabiliser le système dynamique.

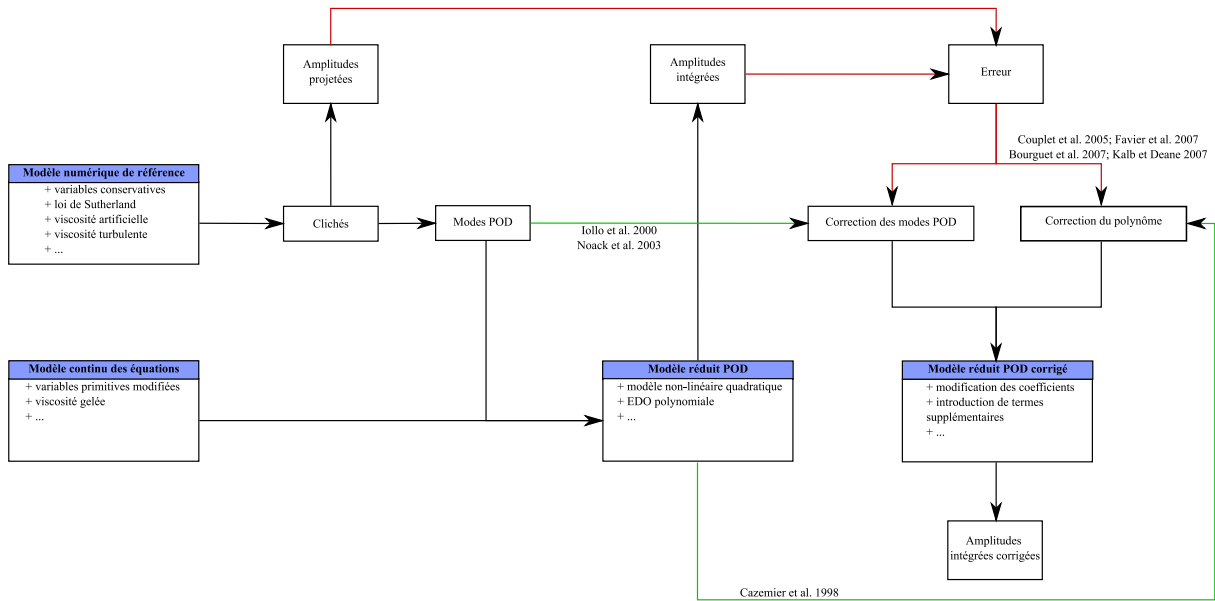
La question de la stabilité des modèles d'ordre réduit POD-Galerkin a été soulevée dès les premiers travaux et un grand nombre de techniques a vu le jour afin de tenter de la préserver. On peut distinguer d'une part les techniques visant à améliorer la structure et le contenu de la base tronquée des modes POD sans modifier la forme du modèle d'ordre réduit obtenu par projection de Galerkin, et d'autre part les techniques dans lesquelles le modèle d'ordre réduit est modifié soit en le complétant par l'introduction de termes supplémentaires, soit en corrigeant les coefficients. Le schéma de la figure 5.1 résume pour plus de clarté le principe de fonctionnement des différentes méthodes de correction. Il est aussi possible de distinguer comme on le fera dans la suite les techniques a priori et a posteriori avec lesquelles la correction s'appuie ou non sur la réponse produite par le modèle d'ordre réduit intégré en temps sans aucune correction.

Le système dynamique (5.12) construit à partir d'une base de modes POD  $\Phi$  est récrit pour la suite sous la forme

$$\dot{\mathbf{a}} = \hat{\mathbf{f}}^G(\mathbf{a}) \quad \text{avec} \quad \hat{\mathbf{f}}^G(\mathbf{a}) = \mathcal{K} + \mathcal{L}\mathbf{a} + (\mathcal{Q} \otimes \mathbf{a})\mathbf{a} \quad (5.25)$$

pour reprendre la notation de l'équation (1.3). Chaque composante  $\{\hat{f}_i^G\}_{i=1}^q$  du vecteur  $\hat{\mathbf{f}}^G(\mathbf{a})$  est donc un polynôme de degré deux portant sur les amplitudes modales, de sorte que  $\hat{f}_i^G(\mathbf{a}) = \mathcal{K}_i + \sum_{j=1}^q \mathcal{L}_{ij}a_j + \sum_{j,k=1}^q \mathcal{Q}_{jik}a_ja_k$ . Les coefficients des polynômes sont calculés à partir des expressions analytiques (5.9) qui proviennent de la projection de Galerkin. Pourvu que les coefficients du modèle d'ordre réduit soient relativement bien calculés, l'intégration en temps de ce système d'équations est possible au moins sur l'intervalle

<sup>1</sup>La raison provient en partie de la troncature effectuée dans la base qui élimine les modes associés aux petites valeurs propres et aux petites structures spatiales qui peuvent participer activement au transfert et à la dissipation d'énergie dans le système.



**Figure 5.1 :** Schématisation du principe de correction du modèle d'ordre réduit POD-Galerkin. On distingue les méthodes de correction a priori (—) des méthodes de correction a posteriori (—) qui reposent sur la définition d'une erreur de reconstruction obtenue à partir d'une première résolution du modèle d'ordre réduit non corrigé.

de temps  $I_e = [t_0; t_0 + T_e]$  et produit des amplitudes modales  $\mathbf{a}(t_m)$  évaluées aux mêmes instants que les clichés. Ces amplitudes sont plus ou moins proches des amplitudes exactes  $\mathbf{a}^{(m)}$  correspondant à la projection des clichés sur les modes POD. Le modèle d'ordre réduit est alors validé soit (i) en comparant les amplitudes  $\mathbf{a}(t_m)$  calculées avec le modèle d'ordre réduit aux amplitudes exactes  $\mathbf{a}^{(m)}$ , soit (ii) en comparant une quantité annexe  $\chi^G = \hat{g}(\mathbf{a}(t_m))$  (variable aérodynamique, coefficient de portance, énergie totale,...) avec la quantité de référence issue du calcul avec le modèle non réduit notée  $\chi^{(m)}$ . Quelle que soit l'approche envisagée, l'objectif est de modifier le modèle d'ordre réduit  $\dot{\mathbf{a}} = \hat{\mathbf{f}}^G(\mathbf{a})$  en un nouveau modèle d'ordre réduit  $\dot{\mathbf{a}} = \hat{\mathbf{f}}^c(\mathbf{a})$  où les termes du polynôme sont corrigés par rapport à ceux de la projection de Galerkin initialement effectuée avec les vecteurs de la base  $\Phi$ .

### 5.3.2 Techniques de correction a priori des modèles d'ordre réduit POD-Galerkin

#### 5.3.2.1 Correction des vecteurs de la base POD

Pour pallier le problème de stabilité du modèle d'ordre réduit, certains auteurs ont évoqué la nécessité de compléter la base des modes POD par de nouvelles fonctions préalablement orthonormalisées, ou de modifier légèrement la façon de calculer les modes POD. Une nouvelle base POD  $\tilde{\Phi}$  est donc construite et donne par conséquent naissance au polynôme corrigé  $\hat{\mathbf{f}}^c$  après projection de Galerkin.

Noack *et al.* [171] ont ainsi proposé de compléter la base des modes POD avec un vecteur supplémentaire associé à une correction liée à l'écart entre la solution stationnaire réelle  $\mathbf{q}_s$  de l'écoulement et la partie moyenne des clichés  $\bar{\mathbf{q}}$  impliquée dans la décomposition affine des variables sur la base des modes POD. Rempfer [188] a quant à lui suggéré de compléter la base avec des éléments appartenant au noyau de l'opérateur de corrélation associé à la décomposition orthogonale propre et de procéder à une analyse de Floquet pour déterminer les éventuels modes instables dans la base POD générée. Enfin, il est possible de chercher à compléter la base POD en utilisant les informations contenues dans l'erreur de reconstruction produite par le modèle d'ordre réduit non corrigé mais dans ce cas il ne s'agit plus d'une correction a priori.

Plutôt que de chercher à compléter la base des modes POD, une solution consiste à en améliorer le contenu en construisant les modes POD au moyen d'un produit scalaire différent de celui de l'espace euclidien  $\mathbb{R}^N$  ou de l'espace  $L^2(\Omega)$ . Iollo *et al.* [117] ont ainsi remarqué que l'utilisation du produit scalaire basé sur l'espace de Sobolev  $H^1$  permet d'améliorer nettement la stabilité du modèle d'ordre réduit obtenu.

En effet, le nouveau produit scalaire (4.55) fait intervenir le gradient des fonctions et par conséquent permet de tenir compte des petites structures associées aux modes négligés. L'amélioration n'a toutefois été constatée que pour un écoulement turbulent dans lequel des structures associées aux petites échelles de la turbulence sont susceptibles d'être présentes. Kirby [132] a lui aussi défini un produit scalaire pour un espace de Sobolev impliquant des dérivées d'ordre supérieur et a obtenu une amélioration significative en introduisant les dérivées secondes du champ aérodynamique pour le calcul des modes POD. Dans les deux cas cependant, la méthode nécessite d'ajuster empiriquement des paramètres de pondération du produit scalaire — comme le paramètre  $\mu_{H^1}$  dans l'expression (4.55).

Henri et Yvon [110] ont suggéré de construire les modes POD en utilisant comme clichés la réponse du système non réduit ainsi que la dérivée *en temps* de cette réponse afin que les modes POD contiennent une partie de la dynamique du système. Par ailleurs l'erreur de reconstruction est parfois améliorée en conservant dans la base de projection des modes POD associés à de petites valeurs propres.

À la différence de ces méthodes de correction des vecteurs de la base de projection, les méthodes exposées dans les paragraphes qui suivent reposent toutes sur la correction directe du polynôme  $\hat{f}^G$ .

### 5.3.2.2 Introduction d'amortissement dans le modèle d'ordre réduit

L'une des raisons les plus souvent invoquées pour expliquer l'instabilité des systèmes POD-Galerkin est la troncature réalisée dans la base POD pour obtenir un modèle d'ordre réduit de petite taille. Bien qu'en général la base tronquée des modes POD contienne la quasi-totalité de l'énergie du système, les modes POD négligés contiennent encore des contributions non négligeables, notamment en ce qui concerne le caractère dissipatif du système. Les modes POD négligés exhibent en effet des petites structures spatiales qui jouent un rôle important dans les transferts d'énergie.

Qualitativement, ce processus est très similaire aux transferts d'énergie qui ont lieu dans le contexte de la turbulence entre les structures cohérentes de grandes et de petites échelles. La méthode POD-Galerkin partage donc quelques traits communs avec les méthodes de Simulation des Grandes Échelles (*Large Eddy Simulation*, ou LES en anglais) qui agissent comme un filtre sur les équations de Navier-Stokes et permettent de ne résoudre que les grandes échelles de la turbulence en modélisant l'action des petites échelles à l'aide d'une viscosité turbulente ou d'un modèle de similarité d'échelle par exemple. Un parallèle entre les transferts d'énergie agissant d'une part entre structures turbulentes et d'autre part entre les modes POD d'un écoulement turbulent incompressible a été établi par Couplet *et al.* [52] et a permis de mettre en place une paramétrisation visqueuse des interactions entre les modes POD résolus et non-résolus.

Dans le cas général d'un écoulement non nécessairement turbulent ni incompressible, l'introduction d'une viscosité artificielle dans le modèle d'ordre réduit POD-Galerkin a été une des premières solutions adoptées pour stabiliser le système. Plusieurs manières d'introduire cette viscosité additionnelle ont été exploitées. La plus simple consiste à multiplier la viscosité physique  $\mu$  par un facteur constant  $1 + \eta_a$  où  $\eta_a = \mu_a/\mu$ , ce qui revient finalement à introduire le nombre de Reynolds corrigé

$$\text{Re}^c = \frac{\rho_0 L_0 U_0}{\mu(1 + \eta_a)} \iff \frac{1}{\text{Re}^c} = \frac{1}{\text{Re}} + \frac{1}{\text{Re}^a}, \quad (5.26)$$

défini à partir des grandeurs aérodynamiques d'adimensionnement ou de façon équivalente à partir du nombre de Reynolds réel de l'écoulement  $\text{Re}$  et du nombre de Reynolds ajouté  $\text{Re}^a$  associé à la viscosité artificielle  $\mu_a$ . Une mesure de la quantité de viscosité artificielle introduite est alors donnée par le ratio

$$\eta_a = \frac{\mu_a}{\mu} = \frac{\text{Re}}{\text{Re}^a} = \frac{\text{Re}}{\text{Re}^c} - 1. \quad (5.27)$$

Une variante permettant d'attribuer à chaque mode un certain niveau de dissipation a été employée par Cazemier [41] : la viscosité additionnelle attribuée au  $i$ -ième mode POD était simplement donnée par  $\mu_{a,i} = i \times \text{Cst}$ , où  $\text{Cst}$  est une constante déterminée de façon empirique après une série d'essais. Ce type de correction conduit au polynôme corrigé

$$\boxed{\hat{f}_i^c(\mathbf{a}) = \mathcal{K}_i^C + \sum_{j=1}^q \mathcal{L}_{ij}^C a_j + \sum_{j,k=1}^q \mathcal{Q}_{jik}^C a_j a_k + \frac{1}{\text{Re}^c} \left( \mathcal{K}_i^D + \sum_{j=1}^q \mathcal{L}_{ij}^D a_j + \sum_{j,k=1}^q \mathcal{Q}_{jik}^D a_j a_k \right)}, \quad (5.28)$$

où les différents coefficients ont été définis à l'équation (5.9). La correction introduite agit sur l'ensemble des termes diffusifs du modèle d'ordre réduit, et par conséquent sur les contributions constantes  $\mathcal{K}_i^D$ , linéaires  $\mathcal{L}_{ij}^D$  et quadratiques  $\mathcal{Q}_{jik}^D$ . En réécrivant le facteur  $1/\text{Re}^c = 1/\text{Re} + 1/\text{Re}^a$  où  $\text{Re}^a$  est le nombre de Reynolds ajouté, il est possible de mettre le polynôme corrigé sous la forme  $\hat{f}_i^c(\mathbf{a}) = \hat{f}_i^G(\mathbf{a}) + \hat{f}_i^{G,D}(\mathbf{a})/\text{Re}^a$  dans laquelle  $\hat{f}_i^{G,D}$  correspond à la partie diffusif du polynôme  $\hat{f}_i^G$ .

La détermination de la viscosité artificielle introduite pour stabiliser le modèle d'ordre réduit avec la technique précédente est empirique et par conséquent, Cazemier *et al.* [42] ont proposé une autre méthode pour tenir compte de l'effet des modes négligés au moyen d'un terme d'amortissement linéaire. L'idée est de déterminer ce terme en imposant que la moyenne du taux de variation de l'énergie dans la direction de chaque mode POD soit conservée. Par définition, le taux de variation de l'énergie dans la direction du mode POD  $i$  est donné par

$$\frac{da_i^2}{dt} = 2 a_i \frac{da_i}{dt}, \quad (5.29)$$

ou encore

$$2 a_i \frac{da_i}{dt} = 2 a_i \mathcal{K}_i + 2 \sum_{j=1}^q \mathcal{L}_{ij} a_i a_j + 2 \sum_{j,k=1}^q \mathcal{Q}_{jik} a_i a_j a_k, \quad (5.30)$$

en multipliant l'équation (5.8) gouvernant le modèle d'ordre réduit par le facteur  $2a_i$ . La moyenne du taux de variation de l'énergie dans la direction de chaque mode POD est alors conservée si l'action de l'opérateur de moyenne  $E[\cdot]$  sur l'équation précédente donne bien un résultat identiquement nul qui correspond à

$$\sum_{j,k=1}^q \mathcal{Q}_{jik} E[a_i a_j a_k] + \lambda_i \mathcal{L}_{ii} = 0, \quad (5.31)$$

en utilisant les propriétés (2.28) et  $E[a_i^{(m)}] = 0$  pour simplifier le résultat<sup>2</sup>. Pour que cette relation soit vérifiée par le modèle d'ordre réduit, on introduit alors un terme linéaire supplémentaire  $\mathcal{D}_i$  dans le polynôme corrigé qui devient

$$\boxed{\hat{f}_i^c(\mathbf{a}) = \mathcal{K}_i + \sum_{j=1}^q \mathcal{L}_{ij} a_j + \sum_{j,k=1}^q \mathcal{Q}_{jik} a_j a_k + \mathcal{D}_i a_i}. \quad (5.32)$$

Le coefficient d'amortissement introduit est alors défini par

$$\mathcal{D}_i = -\frac{1}{\lambda_i} \sum_{j,k=1}^q \mathcal{Q}_{jik} E[a_i a_j a_k] - \mathcal{L}_{ii}. \quad (5.33)$$

Le calcul du terme  $E[a_i a_j a_k]$  s'effectue à partir des amplitudes modales  $\mathbf{a}^{(m)}$  issues de la projection des clichés sur les modes POD. Cazemier *et al.* ont aussi suggéré de ne conserver que les valeurs négatives de  $\mathcal{D}_i$  de manière à introduire exclusivement de l'amortissement dans le modèle d'ordre réduit. À la différence de la correction (5.28), ce modèle agit uniquement sur les coefficients linéaires et la correction utilise des contributions provenant à la fois des termes diffusifs et convectifs.

D'autres manières d'introduire une viscosité dans le modèle d'ordre réduit ont été adoptées par le passé. En particulier, Aubry *et al.* [18] ont utilisé un modèle de fermeture de Heisenberg pour tenir compte des modes non résolus. Cette approche est très similaire à ce qui est fait dans le cadre de la LES avec un modèle de type Smagorinsky. Sirisup et Karniadakis [216] ont quant à eux préféré introduire une viscosité spectrale évanescence pour stabiliser le modèle d'ordre réduit à long terme.

Les méthodes de correction basées sur l'introduction d'une viscosité artificielle permettent de minimiser l'effet des modes tronqués en rajoutant la dissipation qui aurait dû être prise en compte par les modes négligés. Cependant, en dehors de la méthode basée sur la conservation de la moyenne du taux d'énergie dans la direction de chaque mode développée par Cazemier *et al.* [42], la plupart des techniques nécessitent de déterminer de façon empirique un ou plusieurs paramètres dépendant entre autres du nombre de modes conservés dans la base réduite ou encore du type d'écoulement. Par conséquent il est préférable d'avoir recours à des méthodes de stabilisation plus robustes.

<sup>2</sup>La seconde propriété  $E[a_i^{(m)}] = 0$  n'est vraie que pour une réponse périodique évaluée sur un intervalle de temps adéquat.

### 5.3.3 Techniques de correction a posteriori des modèles d'ordre réduit POD-Galerkin

#### 5.3.3.1 Optimisation des coefficients du modèle d'ordre réduit

Une manière moins empirique de procéder à la correction du modèle d'ordre réduit consiste à calibrer certains coefficients des opérateurs  $\mathcal{K}$ ,  $\mathcal{L}$  et  $\mathcal{Q}$  en utilisant l'historique des amplitudes modales intégrées avec le modèle d'ordre réduit. Les coefficients du modèle réduit sélectionnés deviennent alors les paramètres  $\mathbf{w}$  d'un problème d'optimisation dans lequel une certaine erreur de reconstruction doit être minimisée. La résolution du problème d'optimisation fournit alors un ensemble de nouveaux coefficients qui définissent le polynôme corrigé  $\hat{\mathbf{f}}^c$  de sorte que les amplitudes  $\mathbf{a}(t_m)$  et  $\mathbf{a}^{(m)}$  coïncident au moins sur l'intervalle de temps  $I_\epsilon$ .

La méthode a été employée à l'origine par Galletti *et al.* [82] qui cherchaient à optimiser un terme correctif linéaire  $C_{ij}a_j(t)$  dans le modèle d'ordre réduit afin de tenir compte du terme de pression qui est généralement négligé pour les écoulements incompressibles. Plus récemment, Buffoni *et al.* [35] et Galletti *et al.* [81] ont mis en œuvre une méthode de calibrage similaire permettant d'optimiser à la fois les coefficients constants et linéaires du modèle d'ordre réduit. Couplet *et al.* [51] ont ensuite généralisé la méthode en proposant un calibrage de tout ou partie des coefficients du modèle d'ordre réduit. L'avantage de cette approche est que les coefficients linéaires, mais aussi constants et quadratiques peuvent être corrigés. Bourguet *et al.* [34] ont quant à eux comparé divers types d'erreur qui définissent la fonctionnelle à minimiser. De leur côté, Favier *et al.* [73] ont confronté plusieurs méthodes de calibrage à une méthode de stabilisation analogue à celle employée par Rempfer [186] et dans laquelle le nombre de Reynolds est modifié pour chaque mode POD. Enfin, Cordier *et al.* [48] ont présenté les diverses formes de calibrage au moyen d'un même formalisme et ont établi les liens existants avec la méthode d'identification des coefficients. Une comparaison de ces méthodes est par ailleurs effectuée.

Le problème d'optimisation est formulé de manière très générale en notant  $\mathbf{w}$  le vecteur de dimension  $p_c < N'_c = q(q+1)(q+2)/2$  contenant les coefficients du polynôme  $\hat{\mathbf{f}}^G$  qui doivent être calibrés. La fonctionnelle à minimiser porte sur une erreur  $\epsilon$  qui peut être exprimée soit en fonction de la différence entre les amplitudes modales  $\mathbf{a}(t_m)$  et  $\mathbf{a}^{(m)}$ , soit en fonction de la différence entre des quantités physiques  $\chi(t_m)$  et  $\chi^{(m)}$ . Dans les deux cas, la fonctionnelle dépend d'une manière plus ou moins complexe du polynôme  $\hat{\mathbf{f}}^G$  et par conséquent des paramètres  $\mathbf{w}$  à optimiser. La fonctionnelle à minimiser est donc notée  $\mathcal{J}(\hat{\mathbf{f}}^G(\mathbf{w}))$  et le polynôme corrigé recherché  $\hat{\mathbf{f}}^c$  est défini à partir des coefficients calibrés solutions du problème d'optimisation

$$\boxed{\hat{\mathbf{f}}^c = \hat{\mathbf{f}}^G(\mathbf{w}_{\text{opt}}) \quad \text{avec} \quad \mathbf{w}_{\text{opt}} = \arg \min_{\mathbf{w} \in \mathbb{R}^{p_c}} \mathcal{J}(\hat{\mathbf{f}}^G(\mathbf{w}))}. \quad (5.34)$$

En pratique, la résolution du problème d'optimisation est facilitée lorsque la fonctionnelle s'écrit comme une somme de carrés puisqu'un certain nombre de méthodes ont été spécifiquement développées pour des fonctions quadratiques [77]. Idéalement, la fonctionnelle s'écrit donc

$$\mathcal{J}(\mathbf{r}) = \sum_{l=1}^{n_\epsilon} r_l^2 \quad \text{où} \quad \mathbf{r} = \mathbf{r}(\hat{\mathbf{f}}^G(\mathbf{w})) \quad (5.35)$$

est simplement une application vectorielle qui associe au polynôme  $\hat{\mathbf{f}}^G(\mathbf{w})$  à calibrer un vecteur d'erreur de taille  $n_\epsilon$ . Le lien entre les paramètres à calibrer et la fonctionnelle à minimiser se résume par le jeu de relations

$$\mathcal{J} : \mathbb{R}^{p_c} \xrightarrow{\hat{\mathbf{f}}^G} \mathbb{R}^q \xrightarrow{\mathbf{r}} \mathbb{R}^{n_\epsilon} \xrightarrow{\mathcal{J}} \mathbb{R} \\ \mathbf{w} \mapsto \hat{\mathbf{f}}^G(\mathbf{w}) \mapsto \mathbf{r}(\hat{\mathbf{f}}^G(\mathbf{w})) \mapsto \mathcal{J}(\mathbf{r}(\hat{\mathbf{f}}^G(\mathbf{w}))). \quad (5.36)$$

Lors du processus d'optimisation, les  $p_c$  paramètres du vecteur  $\mathbf{w}$  définissent le polynôme  $\hat{\mathbf{f}}^G(\mathbf{w})$  en cours de calibrage. Les équations du modèle d'ordre réduit définies par ce polynôme sont alors employées pour calculer les amplitudes  $\mathbf{a}(t_m)$  — ou une quantité  $\chi(t_m)$  déduite de ces amplitudes — et établir ensuite le vecteur d'erreur  $\mathbf{r} \in \mathbb{R}^{n_\epsilon}$  dont la dimension dépend de la définition de l'erreur  $\epsilon$  choisie. Si l'erreur porte sur les amplitudes modales ( $\epsilon = \mathbf{a}(t_m) - \mathbf{a}^{(m)}$  par exemple), alors il est possible de l'évaluer à chaque instant  $t_m$  correspondant aux clichés et pour chaque mode  $i$ , ce qui conduit à  $n_\epsilon = Mq$  composantes pour le vecteur d'erreur. Si au contraire la variable optimisée est une quantité scalaire évaluée à chaque instant

(par exemple le coefficient de portance ou l'énergie totale du système), alors on ne dispose que de  $n_\epsilon = M$  composantes pour l'erreur. De manière générale, on considère que l'erreur est évaluée à chaque instant  $t_m$  pour  $m = 1, \dots, M$  et que la quantité contrôlée est caractérisée par  $N_{\text{dels}}$  degrés de liberté. La fonctionnelle portant sur l'erreur générique  $\epsilon$  est alors définie par

$$\mathcal{J}(\widehat{\mathbf{f}}^G(\mathbf{w})) = E \left[ \|\epsilon\|_2^2 \right] = \sum_{m=1}^M \alpha_m \|\epsilon(t_m)\|_2^2 = \sum_{m=1}^M \alpha_m \sum_{i=1}^{N_{\text{dels}}} |\epsilon_i(t_m)|^2 \quad (5.37)$$

afin de minimiser en moyenne la norme de l'erreur. Les composantes du vecteur d'erreur  $\mathbf{r}$  de dimension  $n_\epsilon = M N_{\text{dels}}$  introduit à l'équation (5.35) sont donc

$$r_l = r_{m+(i-1)M} = \sqrt{\alpha_m} |\epsilon_i(t_m)|, \quad \forall (m, i) \in \llbracket 1; M \rrbracket \times \llbracket 1; N_{\text{dels}} \rrbracket. \quad (5.38)$$

Il reste maintenant à détailler l'expression de l'erreur générique  $\epsilon$  pour définir complètement le problème d'optimisation. L'approche la plus courante consiste à définir l'erreur comme l'écart entre les (dérivées des) amplitudes modales  $\mathbf{a}(t_m)$  calculées par le modèle d'ordre réduit et celles de référence  $\mathbf{a}^{(m)}$  obtenues par projection des clichés. Ces choix ont notamment été effectués par Bourguet *et al.* [34], Couplet [50], Favier *et al.* [73]. Pour reprendre la terminologie de Bourguet *et al.*, la minimisation de l'erreur  $\epsilon$  basée sur l'écart entre les amplitudes correspond à la *méthode de calibrage de Floquet*, alors que l'erreur basée sur l'écart entre les dérivées des amplitudes porte le nom de *méthode de calibrage de Poincaré*. Dans ce cas, le vecteur d'erreur  $\mathbf{r}$  est de dimension  $n_\epsilon = M q$  puisque l'erreur  $\epsilon_i(t_m)$  est évaluée à chaque instant et pour chaque mode. Les fonctionnelles pour ces deux types d'erreur sont donc respectivement données par

$$\mathcal{J}^F(\widehat{\mathbf{f}}^G(\mathbf{w})) = E \left[ \left\| \mathbf{a}(t_m) - \mathbf{a}^{(m)} \right\|_2^2 \right] \quad \text{avec} \quad \epsilon = \mathbf{a}(t_m) - \mathbf{a}^{(m)} \quad \text{et} \quad (5.39a)$$

$$\mathcal{J}^P(\widehat{\mathbf{f}}^G(\mathbf{w})) = E \left[ \left\| \dot{\mathbf{a}}(t_m) - \dot{\mathbf{a}}^{(m)} \right\|_2^2 \right] \quad \text{avec} \quad \epsilon = \dot{\mathbf{a}}(t_m) - \dot{\mathbf{a}}^{(m)}. \quad (5.39b)$$

Les définitions précédentes conduisent à une formulation de la fonctionnelle fortement non-linéaire en fonction des paramètres  $\mathbf{w}$  à calibrer. En effet, la fonctionnelle dépend des paramètres par le biais des amplitudes  $\mathbf{a}(t_m)$  — ou de leurs dérivées — qui proviennent de l'intégration en temps du modèle d'ordre réduit non-linéaire caractérisé par  $\widehat{\mathbf{f}}^G(\mathbf{w})$ . Par conséquent Couplet [50] a proposé de linéariser le problème en remplaçant les amplitudes  $\mathbf{a}(t_m)$  provenant de l'intégration du modèle d'ordre réduit par  $\mathbf{a}(t_m) \approx \mathbf{a}^{(m)}(0) - \int_{t_0}^{t_0+t_m} \widehat{\mathbf{f}}^G(\mathbf{a}^{(m)}(\tau)) d\tau$ , ce qui revient non plus à calibrer directement les amplitudes mais le polynôme. De la même façon, les dérivées des amplitudes intervenant dans la formulation de Poincaré sont remplacées par  $\dot{\mathbf{a}}(t_m) \approx \widehat{\mathbf{f}}^G(\mathbf{a}^{(m)})$ .

Une autre classe d'erreur est définie pour calibrer une grandeur spécifique  $\chi$  qui peut être aussi bien scalaire (coefficient de portance, énergie totale,...) que vectorielle (champ aérodynamique en un certain nombre de degrés de liberté,...). Lorsque la quantité  $\chi$  est scalaire, l'erreur est évaluée à chaque instant et la dimension du vecteur d'erreur est donc  $n_\epsilon = M$ . La fonctionnelle à minimiser porte alors sur l'erreur  $\epsilon = \chi(t_m) - \chi^{(m)}$  et s'exprime donc en fonction de  $\chi(t_m) = \widehat{g}(\mathbf{a}(t_m))$  suivant la relation

$$\mathcal{J}^\chi(\widehat{\mathbf{f}}^G(\mathbf{w})) = E \left[ \left\| \chi(t_m) - \chi^{(m)} \right\|_2^2 \right] = E \left[ \left\| \widehat{g}(\mathbf{a}(t_m)) - \chi^{(m)} \right\|_2^2 \right]. \quad (5.40)$$

Le calibrage d'une grandeur dérivée du champ aérodynamique devient rapidement très coûteux en temps de calcul, notamment lorsqu'il s'agit d'une grandeur globale. En effet, si l'on souhaite par exemple calibrer le modèle d'ordre réduit de manière à ce que l'énergie totale  $\mathcal{E}_T = \int_{\Omega} \rho e_T d\Omega$  du système soit conservée, il est nécessaire de reconstruire à chaque instant et pour chaque itération de l'algorithme d'optimisation le champ aérodynamique sur l'ensemble du domaine de calcul. Ce problème est moins contraignant pour le calcul du coefficient de portance puisqu'il ne fait intervenir que le champ sur le bord du profil où le nombre de degrés de liberté reste relativement faible.

Le calibrage du modèle d'ordre réduit basé sur la conservation de l'énergie totale semble cependant une idée intéressante puisqu'elle repose sur un principe physique plutôt que sur un simple recalage de variables telles que les amplitudes modales. Pour remédier au problème du coût de calcul, on s'inspire



de la définition de l'énergie cinétique pour un système incompressible. Ainsi, les clichés employés pour un écoulement incompressible correspondent au champ de vitesse  $\mathbf{q} = [u, v, w]^T$  et l'énergie cinétique  $\mathcal{E}_C = \int_{\Omega} 1/2 \rho_0 e_C \, d\Omega$  se déduit donc immédiatement des clichés au moyen du produit scalaire défini sur l'espace  $L^2$  :  $\mathcal{E}_C = 1/2 \rho_0 \langle \mathbf{q}, \mathbf{q} \rangle_{L^2}$ . L'introduction de la décomposition POD  $\mathbf{q}(t) = \sum_{j=1}^q a_j(t) \varphi^{(j)}$  dans l'écriture précédente conduit alors à

$$\mathcal{E}_C \approx \frac{1}{2} \rho_0 \sum_{i=1}^q a_i(t) a_i(t) \quad (5.41)$$

puisque les modes POD sont orthonormaux pour le produit scalaire  $L^2$  par construction. Si on considère la moyenne de l'énergie cinétique, alors celle-ci s'exprime simplement en fonction des valeurs propres  $\lambda_i$  associées aux modes POD puisque la relation (2.28) conduit à  $E[\mathcal{E}_C^{(m)}] = 1/2 \rho_0 \sum_{i=1}^q \lambda_i$ .

Dans le cas d'un écoulement compressible, le lien entre les valeurs propres associées aux modes POD et l'énergie totale de l'écoulement n'est pas évident. En effet, l'énergie totale de l'écoulement ne peut s'exprimer comme une fonction quadratique des clichés comme on l'a vu au paragraphe 4.4.2. Bien que les valeurs propres n'aient plus de lien direct avec l'énergie totale du système, on peut toujours envisager de calibrer le modèle d'ordre réduit à partir de la relation (2.28) qui est vérifiée pour les amplitudes de référence  $\mathbf{a}^{(m)}$  mais pas nécessairement par celles provenant de l'intégration en temps du modèle d'ordre réduit. La nouvelle fonctionnelle

$$\mathcal{J}^{\mathcal{E}_T^*} = \sum_{i=1}^q \left( \lambda_i - \sum_{m=1}^M \alpha_m a_i(t_m) a_i(t_m) \right) \quad (5.42)$$

est donc introduite dans le but d'ajuster le contenu énergétique de chaque mode. Dans ce cas, la méthode revient à calibrer le modèle d'ordre réduit pour la moyenne d'une énergie  $\mathcal{E}_T^*$  définie comme une fonction quadratique des clichés par la relation  $\mathcal{E}_T^* = \langle \mathbf{q}, \mathbf{q} \rangle = \int_{\Omega} (\vartheta^2 + \mathbf{u}^T \mathbf{u} + p^2) \, d\Omega$  telle que  $E[\mathcal{E}_T^*] \approx \sum_{i=1}^q \lambda_i$ . Bien que l'énergie  $\mathcal{E}_T^*$  diffère de l'énergie totale réelle de l'écoulement, on pourra vérifier numériquement que calibrer le modèle d'ordre réduit d'un écoulement compressible en ajustant numériquement le contenu énergétique de chaque mode au moyen de la fonctionnelle (5.42) revient à calibrer l'énergie totale (réelle) de l'écoulement au moyen de l'erreur  $\epsilon = \mathcal{E}_T(t_m) - \mathcal{E}_T^{(m)}$  calculée à partir du champ aérodynamique reconstruit à chaque instant avec la relation (4.50).

Les méthodes basées sur le calibrage de certains des coefficients du modèle d'ordre réduit sont particulièrement efficaces pour corriger le polynôme issu de la projection de Galerkin et conduisent à une solution coïncidant presque exactement avec celle du modèle non réduit sur l'intervalle de temps  $I_e$ . Cependant, le choix des coefficients qui doivent être optimisés n'est pas évident et la résolution du problème d'optimisation devient fastidieuse lorsque le nombre  $p_c$  de coefficients à calibrer devient important. En outre, le problème devient rapidement sur-déterminé lorsque le nombre  $p_c$  de coefficients dépasse la dimension  $n_e$  du vecteur d'erreur sur lequel la définition de la fonctionnelle repose. Enfin, il serait souhaitable lors de la formulation du problème d'optimisation de disposer d'un critère pour fixer des bornes aux intervalles dans lesquels les paramètres varient.

### 5.3.3.2 Stabilisation intrinsèque

Une autre méthode baptisée « stabilisation intrinsèque » par Kalb et Deane [124] consiste aussi à utiliser les informations contenues dans les amplitudes temporelles  $\mathbf{a}^{(m)}$  associées aux clichés pour corriger les coefficients du modèle d'ordre réduit. Le terme correctif est cette fois calculé explicitement à partir de l'erreur entre les amplitudes  $\mathbf{a}^{(m)}$  et celles produites par le modèle d'ordre réduit non corrigé lorsqu'il est initialisé à chaque instant correspondant aux clichés avec les amplitudes  $\mathbf{a}^{(m)}$ . Les amplitudes  $\mathbf{a}(t_m)$  ainsi produites traduisent alors les défaillances liées à l'intégration du polynôme  $\hat{\mathbf{f}}^G$  sur l'intervalle d'échantillonnage.

En effet, en supposant qu'il existe un polynôme de référence  $\hat{\mathbf{f}}^R$  fournissant les amplitudes  $\mathbf{a}^{(m)}$ , l'intégration en temps du modèle d'ordre réduit avec un schéma d'Euler progressif fournit les amplitudes à l'instant  $t_{m+1}$  :  $\mathbf{a}^{(m+1)} = \mathbf{a}^{(m)} + \delta t_c \hat{\mathbf{f}}^R(\mathbf{a}^{(m)})$ , où  $\delta t_c$  est le pas de temps d'échantillonnage. Si le polynôme  $\hat{\mathbf{f}}^G$  calculé au moyen des coefficients définis au début de ce chapitre correspondait au polynôme de référence  $\hat{\mathbf{f}}^R$ , les amplitudes calculées devraient coïncider lorsque le modèle d'ordre réduit défini avec  $\hat{\mathbf{f}}^G$  est intégré en temps. Puisque ce n'est pas le cas, on définit les amplitudes  $\mathbf{a}(t_m)$  en initialisant le modèle d'ordre réduit à chaque instant correspondant aux clichés avec les amplitudes de référence  $\mathbf{a}^{(m)}$  et en intégrant en temps

avec le schéma d'Euler progressif, ce qui conduit à :  $\mathbf{a}(t_{m+1}) = \mathbf{a}^{(m)} + \delta t_c \widehat{\mathbf{f}}^G(\mathbf{a}^{(m)})$ . L'erreur est donc définie à chaque instant  $t_{m+1}$  comme la différence entre les amplitudes exactes et celles calculées avec le polynôme non corrigé  $\widehat{\mathbf{f}}^G$ , ce qui s'écrit

$$\varepsilon(t_{m+1}) = \mathbf{a}^{(m+1)} - \mathbf{a}(t_{m+1}) \approx \delta t_c \widehat{\mathbf{f}}^R(\mathbf{a}^{(m)}) - \delta t_c \widehat{\mathbf{f}}^G(\mathbf{a}^{(m)}). \quad (5.43)$$

Les amplitudes  $\mathbf{a}(t_{m+1})$  sont en pratique calculées avec le même schéma d'intégration en temps que celui adopté pour le modèle d'ordre réduit, et par conséquent le schéma d'Euler progressif est dans le cas présent remplacé par le schéma d'Adams ou un schéma de type BDF suivant la raideur du système. Ensuite, l'erreur est normalisée par le pas de temps d'échantillonnage  $\delta t_c$  de sorte que l'expression (5.43) n'en dépende plus :  $\varepsilon(t_{m+1}) \approx \widehat{\mathbf{f}}^R(\mathbf{a}^{(m)}) - \widehat{\mathbf{f}}^G(\mathbf{a}^{(m)})$ .

L'erreur définie précédemment est alors employée pour améliorer la réponse du modèle d'ordre réduit. En effet, si l'erreur est introduite au second membre du modèle d'ordre réduit pour corriger les défauts du polynôme  $\widehat{\mathbf{f}}^G$ , on obtient la relation suivante :  $\mathbf{a}(t_{m+1}) = \mathbf{a}^{(m)} + \delta t_c \widehat{\mathbf{f}}^G(\mathbf{a}^{(m)}) + \varepsilon(t_{m+1}) \approx \mathbf{a}^{(m+1)}$ . Cependant, pour l'implémentation du terme correctif dans le modèle d'ordre réduit, il est nécessaire d'exprimer cette erreur comme une fonction *temporelle* des amplitudes recherchées puisque pour l'instant la définition (5.43) ne fournit qu'une approximation discrète aux instants  $t_m$  où les clichés ont été prélevés. Pour remédier à ce problème, il suffit de se rappeler que les amplitudes modales  $\mathbf{a}_j = [a_j^{(1)}, \dots, a_j^{(M)}]^\top$  pour  $j = 1, \dots, q$  forment une base orthogonale. L'idée consiste alors à évaluer l'erreur (5.43) aux  $M$  instants associés aux clichés et à projeter chaque composante  $\varepsilon_i = [\varepsilon_i(t_1) \dots \varepsilon_i(t_M)]^\top$  du terme correctif sur cette base. Si on note  $\mathcal{P}_A$  la projection orthogonale sur l'espace engendré par la base des amplitudes modales, alors la projection de  $\varepsilon_i$  est donnée par

$$\mathcal{P}_A \varepsilon_i(t_m) = \sum_{j=1}^q \frac{\langle \varepsilon_i, \mathbf{a}_j^{(m)} \rangle}{\langle \mathbf{a}_j^{(m)}, \mathbf{a}_j^{(m)} \rangle} \mathbf{a}_j^{(m)}. \quad (5.44)$$

La projection est de préférence effectuée sur les erreurs centrées  $\varepsilon_i - \bar{\varepsilon}_i$  où l'opérateur de moyenne employé est identique à celui adopté dans le processus de calcul des modes POD. Finalement, si les erreurs produites par le polynôme  $\widehat{\mathbf{f}}^G$  ne sont pas trop importantes, il est raisonnable d'assimiler la projection des erreurs à l'erreur elle-même, de sorte que  $\varepsilon_i \approx \mathcal{P}_A \varepsilon_i$ . En outre, la projection nécessite une certaine normalisation ; Kalb et Deane [124] ont proposé d'utiliser la norme des amplitudes modales  $\|\mathbf{a}^{(m)}\|_{\mathbb{R}^M}$  définie au moyen du produit scalaire euclidien, bien que cela puisse conduire à des corrections très importantes pour les modes d'indice élevé puisque le produit scalaire est proportionnel aux valeurs propres associées aux modes POD<sup>3</sup>. Le terme correctif d'erreur est finalement approché par l'expression

$$\varepsilon_i(t_m) \approx \bar{\varepsilon}_i + \sum_{j=1}^q \frac{\langle \varepsilon_i - \bar{\varepsilon}_i, \mathbf{a}_j^{(m)} \rangle}{\langle \mathbf{a}_j^{(m)}, \mathbf{a}_j^{(m)} \rangle} \mathbf{a}_j^{(m)} = \mathcal{K}_i^c + \sum_{j=1}^q \mathcal{L}_{ij}^c \mathbf{a}_j^{(m)} \quad (5.45)$$

qui fournit une approximation des fluctuations de l'erreur sur la base des amplitudes exactes. Par extension, le terme correctif pourra être exprimé en fonction des amplitudes  $\mathbf{a}(t)$  recherchées. Ainsi, le polynôme corrigé de la méthode de stabilisation intrinsèque s'écrit

$$\widehat{\mathbf{f}}_i^c(\mathbf{a}) = (\mathcal{K}_i + \mathcal{K}_i^c) + \sum_{j=1}^q (\mathcal{L}_{ij} + \mathcal{L}_{ij}^c) a_j + \sum_{j,k=1}^q \mathcal{Q}_{jik} a_j a_k. \quad (5.46)$$

La méthode de stabilisation intrinsèque semble intéressante puisqu'elle se base sur une correction explicitement définie à partir de l'erreur  $\varepsilon(t_{m+1}) = \mathbf{a}^{(m)} - \mathbf{a}(t_m)$ , sans recourir à la résolution d'un problème d'optimisation comme dans les approches précédentes.

<sup>3</sup>En effet, le produit scalaire  $\langle \mathbf{a}_j^{(m)}, \mathbf{a}_j^{(m)} \rangle$  correspond à l'expression  $E[a_j^{(m)} a_j^{(m)}]$  avec des poids de pondération  $\alpha_m = 1$  pour l'opérateur  $E[\cdot]$ . De la relation (2.28), on en déduit que le produit scalaire vaut ici  $\langle \mathbf{a}_j^{(m)}, \mathbf{a}_j^{(m)} \rangle = M \lambda_j$ , où  $\lambda_j$  est la valeur propre correspondant au mode POD  $\varphi^{(j)}$ .

### 5.3.4 Autres techniques de correction des modèles d'ordre réduit POD-Galerkin

D'autres méthodes de stabilisation ont vu le jour, notamment pour tenter de corriger les problèmes liés au non-respect de certaines conditions aux limites. Noack *et al.* [172] ont ainsi attribué l'instabilité de leur modèle d'ordre réduit POD-Galerkin au fait que le terme de pression généralement négligé pour les écoulements incompressibles entraîne un manque de dissipation. Ils ont donc proposé de corriger l'absence du terme de pression à l'aide d'un modèle quadratique dont les coefficients sont calculés comme les solutions d'une équation de Poisson. Galletti *et al.* [81, 82] et Buffoni *et al.* [35] ont pour leur part modélisé les effets du terme de pression négligé dans le cas incompressible par l'ajout d'un opérateur linéaire dont les coefficients sont calibrés. Plutôt que de corriger le terme de pression, Sirisup et Karniadakis [215] ont préféré introduire au second membre du modèle d'ordre réduit le terme de pénalité

$$\mathcal{G}_i(\mathbf{a}) = -\tau \int_{\Omega} \boldsymbol{\varphi}^{(i)} \cdot (\mathbf{q} - \mathbf{q}_{\infty}) \mathbb{I}_{\partial\Omega}(\mathbf{x}) \, d\Omega, \quad (5.47)$$

où  $\mathbb{I}_{\partial\Omega}$  est la fonction caractéristique du bord telle que  $\mathbb{I}_{\partial\Omega} = 1$  si  $\mathbf{x} \in \partial\Omega$  et  $\mathbb{I}_{\partial\Omega} = 0$  sinon. Les variables aérodynamiques  $\mathbf{q}$  sont donc remplacées par leurs valeurs  $\mathbf{q}_{\infty}$  sur le bord  $\partial\Omega$ . En remplaçant  $\mathbf{q}$  par sa décomposition sur la base des modes POD, le terme de pénalité s'écrit

$$\mathcal{G}_i(\mathbf{a}) = \tau \left( \sum_{j=1}^q a_j \int_{\partial\Omega} \boldsymbol{\varphi}^{(i)} \cdot \boldsymbol{\varphi}^{(j)} \, d\partial\Omega - \int_{\partial\Omega} \boldsymbol{\varphi}^{(i)} \cdot \mathbf{q}_{\infty} \, d\partial\Omega \right). \quad (5.48)$$

La méthode est bien adaptée lorsque les valeurs des variables aérodynamiques sont connues sur le bord du domaine, ce qui est surtout le cas lorsqu'un contrôle est appliqué. Toutefois, la stabilisation du modèle d'ordre réduit dépend fortement du paramètre  $\tau$  qui doit être ajusté correctement.

Vigo [240] a pour sa part tenté d'introduire un terme de stabilisation cubique dans le modèle d'ordre réduit en s'appuyant sur une méthode SUPG (*Streamline Upwind Petrov Galerkin*) déjà existante dans le code de calcul employé pour générer les clichés. Le terme de correction nécessite toutefois de nombreuses approximations pour utiliser les décompositions POD des variables aérodynamiques et est en outre très coûteux à calculer. Enfin, Allery [5] a introduit une correction dans son modèle d'ordre réduit afin de tenir compte des erreurs éventuelles des simulations LES. Pour cela, la correction apportée est déterminée de façon à ce que le champ des vitesses corrigées satisfasse exactement les équations de Navier-Stokes.

Bien que les méthodes de correction présentées soient dans certains cas assez spécifiques, l'éventail des techniques disponibles est assez vaste. Gloerfelt [87] a comparé un grand nombre de méthodes de stabilisation afin de déterminer celle qui sera la plus efficace pour traiter un problème d'oscillations auto-entretenues d'ondes sonores et est arrivé à la conclusion que les méthodes de calibrage basées sur la résolution d'un problème d'optimisation sont certainement les plus robustes. En effet, la plupart des autres méthodes dépendent d'un ou de plusieurs paramètres qui doivent être déterminés au cas par cas après une série d'essais afin de stabiliser correctement le modèle d'ordre réduit. La détermination de ces paramètres reste assez empirique et dépend notamment de l'écoulement modélisé, du nombre de modes tronqués dans la base POD, ... Par ailleurs le calibrage ne constitue pas non plus la solution idéale — bien qu'il soit souvent plus performant. En effet, le problème d'optimisation doit être correctement défini et la convergence vers l'optimum devient délicate lorsque le nombre de paramètres calibrés augmente. Lorsque ces techniques sont insuffisantes, il est toujours possible de se diriger vers la méthode d'identification du modèle d'ordre réduit qui a été présentée au paragraphe 2.4.4. Toutefois, il faut bien comprendre que la correction introduite dans le modèle d'ordre réduit est la manifestation de certains défauts de modélisation qu'il est nécessaire sinon indispensable d'identifier pour pouvoir correctement améliorer la réponse du modèle d'ordre réduit.

Dans ce qui suit, le modèle d'ordre réduit est mis à l'épreuve pour reproduire le champ aérodynamique autour d'un profil NACA0012 en forte incidence. Plusieurs méthodes de correction du modèle d'ordre réduit seront évaluées afin de déterminer celle qui se prête le mieux au problème étudié.

## 5.4 Calcul des modes POD de l'écoulement autour d'un profil NACA0012 fixe en forte incidence

### 5.4.1 Simulation numérique de l'écoulement autour du profil NACA0012 pour la génération des clichés

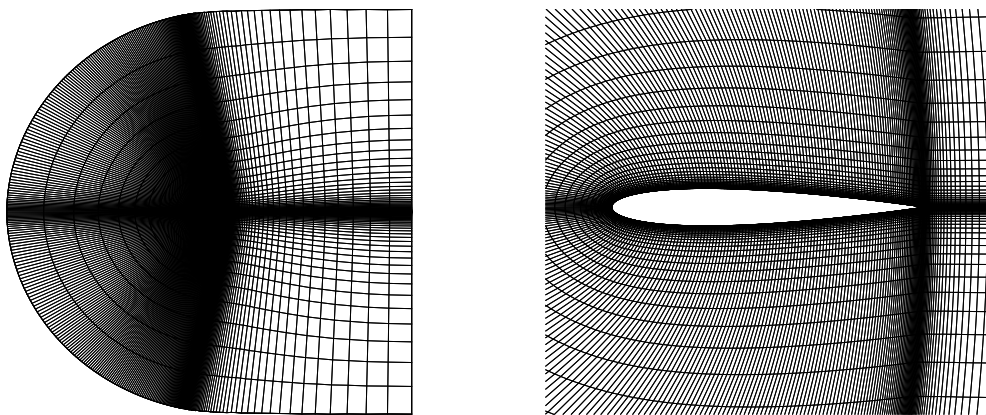
Ce premier exemple est mis en œuvre uniquement dans le but de servir à valider la formulation du modèle d'ordre réduit développé pour un fluide compressible — non turbulent — étudié dans un domaine de frontières fixes. De ce fait, les paramètres de l'écoulement sont déterminés de manière à ce que les hypothèses émises lors de la construction du modèle d'ordre réduit soit vérifiées, plutôt que par souci de réalisme physique.

La première étape de la construction du modèle d'ordre réduit consiste à effectuer une simulation numérique préalable du système afin d'extraire un certain nombre de clichés pour calculer les modes POD. L'instationnarité du système est ici d'origine purement aérodynamique puisqu'elle résulte de détachements de vortex de type Von Kármán qui se forment dans le sillage du profil symétrique NACA0012 pour certaines valeurs de l'angle d'incidence  $\alpha$ , du nombre de Mach  $Ma$  et du nombre de Reynolds  $Re$ .

Pulliam [184] a en particulier étudié l'influence du nombre de Reynolds sur l'intervalle [800; 3 000] pour  $\alpha = 20^\circ$  et  $Ma = 0,2$ , et a mis en évidence une bifurcation qui caractérise la transition entre un détachement simple de vortex et un détachement chaotique. Shan *et al.* [208] ont pour leur part analysé au moyen de simulations directes les détails de la séparation puis du détachement de la couche limite, de l'émission de vortex, de la transition vers la turbulence puis du ré-attachement de la couche limite lorsque le profil est incliné d'un angle d'incidence  $\alpha = 4^\circ$ , pour un nombre de Reynolds  $Re = 10^5$  et un nombre de Mach  $Ma = 0,2$ . D'autres travaux concernent la modélisation du bruit rayonné par l'écoulement autour du profil : Marsden *et al.* [160] ont ainsi simulé le bruit rayonné autour du profil pour un angle d'incidence nul, avec  $Re = 60\,000$  et  $Ma = 0,33$ ; Le Garrec et Gloerfelt [141] ont quant à eux étudié le bruit rayonné pour des angles d'incidence compris entre 0 et  $10^\circ$ , avec  $Re = 5\,000$  et  $Ma = 0,2$ .

Par ailleurs, la simulation de l'écoulement autour d'un profil NACA0012 a servi à plusieurs reprises de cas test pour valider des modèles d'ordre réduit. Favier *et al.* [73] l'ont notamment utilisé dans le cas d'un fluide incompressible afin de valider plusieurs méthodes de calibrage de leur modèle d'ordre réduit. Bourguet *et al.* [34] ont repris le modèle d'ordre réduit développé par Iollo *et al.* [117] et ont validé leur modèle d'ordre réduit en régime transonique ( $Ma = 0,80$  et  $0,85$  pour  $Re = 5\,000$  et  $10\,000$  en incidence nulle) alors que Iollo *et al.* avaient utilisé cet exemple en régime subsonique ( $Ma = 0,2$  et  $Re = 2\,100$ , avec  $\alpha = 20^\circ$ ).

Dans le cas présent, on se place dans des conditions similaires à celles de Iollo *et al.* [117], en inclinant le profil d'un angle d'incidence  $\alpha = 20^\circ$  et en choisissant des paramètres de l'écoulement tels que  $Ma = 0,2$  et  $Re = 2\,000$ . De ce fait, l'écoulement peut donc être considéré comme compressible et laminaire, ce qui rentre dans le cadre des hypothèses qui ont été émises pour élaborer le modèle d'ordre réduit.



**Figure 5.2 :** Maillage du domaine de calcul pour la simulation numérique de l'écoulement autour du profil NACA0012 incliné d'un angle  $\alpha = 20^\circ$ .

Le domaine de calcul représenté sur la figure 5.2 est discrétisé en  $N = 19\,100$  cellules Volumes Finis à l'aide d'un maillage bidimensionnel en C, raffiné localement au niveau de la paroi du profil et vers les bords d'attaque et de fuite. Le maillage a été généré au moyen d'une méthode d'interpolation transfinie [69].

L'écoulement supposé compressible, laminaire et visqueux est modélisé au moyen des équations de Navier-Stokes. Les équations sont résolues pour les variables conservatives avec une méthode multigrille à deux niveaux. L'intégration temporelle des équations est effectuée au moyen du schéma d'Euler rétrograde et les flux sont discrétisés selon l'approche de Jameson [120] (voir l'annexe E.3 pour les détails relatifs à la discrétisation). On peut distinguer deux types de frontières du domaine  $\Omega$  : la première, notée  $\Gamma$  correspond à l'interface fluide-structure où la condition limite (4.45) s'applique avec une vitesse de la structure nulle puisque le profil est fixe ; la seconde frontière notée simplement  $\partial\Omega_\infty$  représente une limite non physique du domaine  $\Omega$  supposé infini. Sur cette portion de la frontière, des conditions aux limites dites de non-réflexion sont imposées pour éviter la réflexion des grandeurs aérodynamiques. Les valeurs du champ sur la frontière perméable sont soit imposées par les valeurs du champ dans  $\Omega$ , soit par les valeurs pré-supposées à l'infini en fonction des relations caractéristiques.

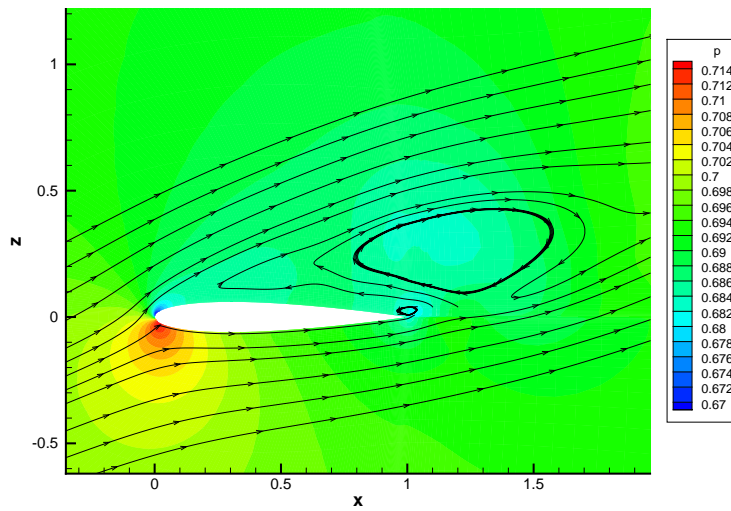


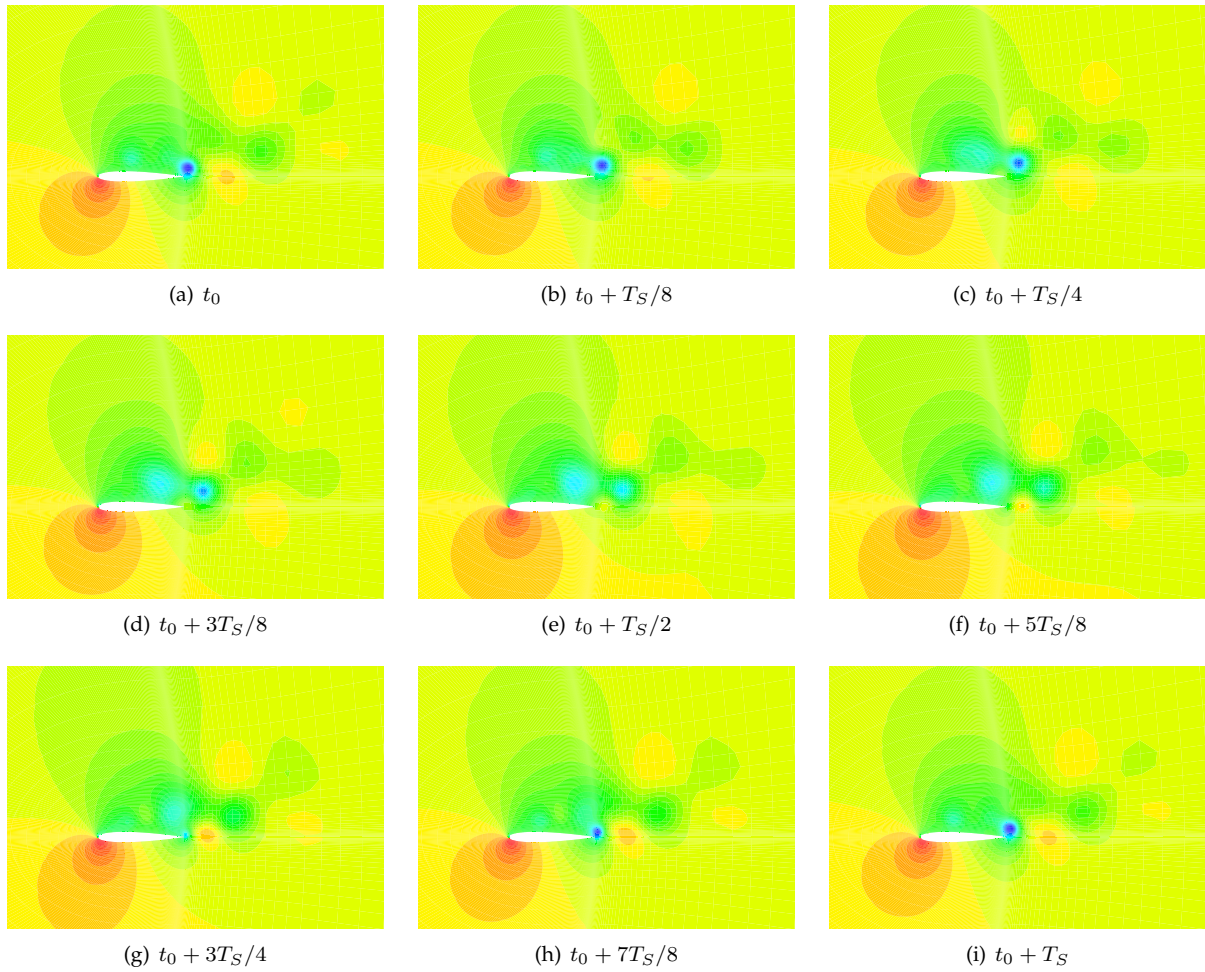
Figure 5.3 : Champ de pression stationnaire et lignes de courant autour du profil NACA0012 en forte incidence.

Une première simulation stationnaire est effectuée afin de mettre en évidence l'existence d'une zone de recirculation sur l'extrados vers le bord de fuite. Celle-ci est visible sur la figure 5.3 où quelques lignes de courant de l'écoulement ont été superposées au champ de pression stationnaire. L'existence d'une telle structure laisse supposer un possible détachement de cette zone de recirculation, conduisant ainsi à la création d'un phénomène instationnaire caractérisé par des allées de Von Kármán.

On procède donc ensuite à des simulations instationnaires afin de vérifier l'existence du détachement de vortex dans le sillage du profil, qui sera le facteur déclencheur de l'instationnarité de l'écoulement. La simulation est initiée à partir de la solution stationnaire calculée précédemment. Après une période transitoire, les différentes grandeurs aérodynamiques se mettent à osciller de façon périodique. La figure 5.4 présente par exemple le champ de pression relevé à différents instants sur une période complète de détachement tourbillonnaire. Les deux vortex dans le sillage proche à  $t_0$  sont convectés progressivement en aval, tandis que de nouveaux vortex commencent à se former sur l'extrados du profil. Ceux-ci sont expulsés vers  $t_0 + T_S/2$  puis convectés à leur tour dans le sillage jusqu'à  $t_0 + T_S$  où ils occupent exactement la même place que les vortex précédemment émis.

La validation du modèle d'ordre réduit effectuée par la suite reposera sur la bonne reproduction du phénomène instationnaire. Ce dernier peut être caractérisé notamment par les courbes d'évolution des coefficients aérodynamiques de portance  $C_L$  et de traînée  $C_D$  définis respectivement par les relations

$$C_D = \frac{F_D}{0.5 \rho_0 U_\infty L_0} \quad \text{et} \quad C_L = \frac{F_L}{0.5 \rho_0 U_\infty L_0}, \quad (5.49)$$



**Figure 5.4 :** Clichés du champ de pression sur une période de détachement tourbillonnaire. Le détachement de vortex dans le sillage s'effectue de façon périodique à la période de Strouhal  $T_S$  et donne ainsi naissance à une allée de Von Kármán.

où  $F_L$  et  $F_D$  sont les efforts de portance et de traînée définis comme la projection des efforts aérodynamiques le long d'un axe perpendiculaire à la direction de l'écoulement pour  $F_L$  et d'un axe parallèle à l'écoulement pour  $F_D$ . Ces deux composantes sont extraites directement d'*elsA* au moyen des flux convectifs et diffusifs dans les deux directions souhaitées<sup>4</sup>. Les courbes d'évolution des coefficients de traînée et de portance sont présentées sur la figure 5.5(a) en fonction du temps adimensionnel  $t^* = t/t_0$  où  $t$  désigne le temps physique et  $t_0 = L_0/U_0$  est le temps utilisé pour l'adimensionnement avec  $U_0 = \sqrt{\gamma R \theta_0}$  —  $\gamma = 1,4$  étant le rapport des chaleurs spécifiques,  $R = 287,04 \text{ J kg}^{-1} \text{ K}^{-1}$  étant la constante des gaz parfaits de l'air et  $\theta_0$  est la température de référence en Kelvin. Après une période transitoire s'étalant sur environ cinq périodes, les variables aérodynamiques présentent une allure sinusoïdale entretenue comme en témoignent les historiques des coefficients aérodynamiques. Les transformées de Fourier de la partie établie du signal représentées sur la figure 5.5(b) confirment le caractère parfaitement sinusoïdal du phénomène d'émission de vortex, et permettent de déterminer la période de Strouhal  $T_S^* = 9,6 \times 10^2$  associée au détachement, ce qui est cohérent avec la valeur trouvée par Iollo *et al.* [117] pour des paramètres d'écoulement proches.

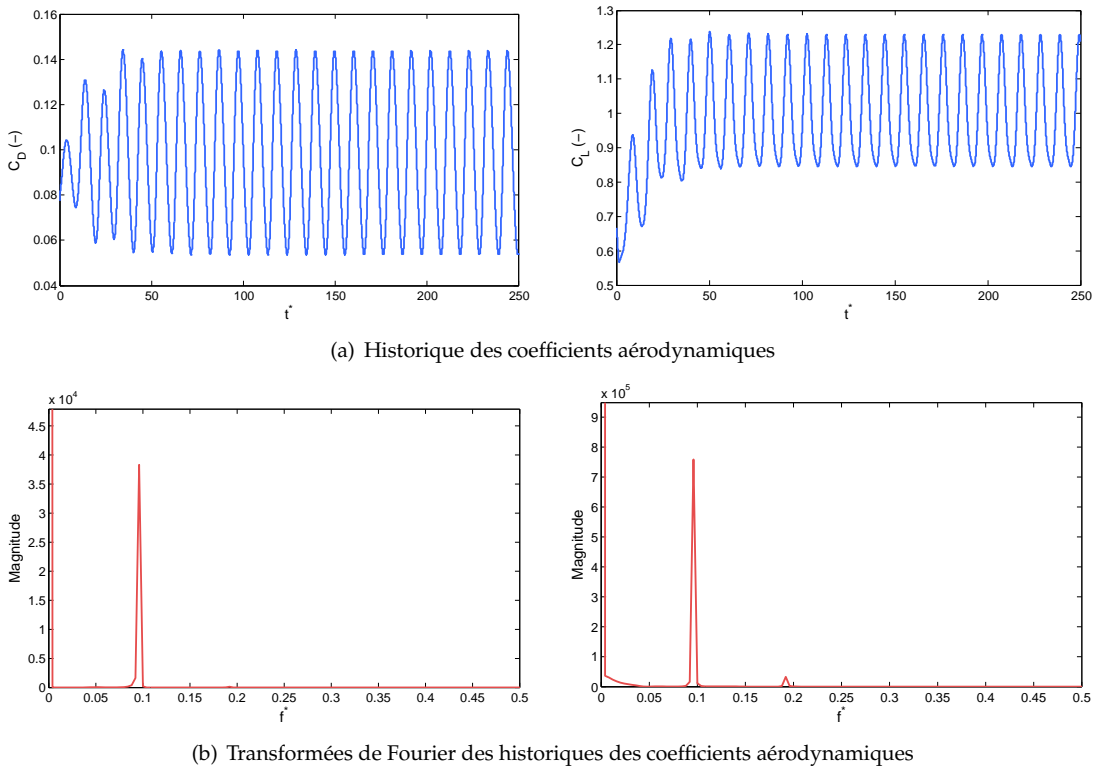


Figure 5.5 : Historiques et transformées de Fourier des coefficients de traînée et de portance.

### 5.4.2 Calcul des modes POD de l'écoulement autour du profil NACA0012

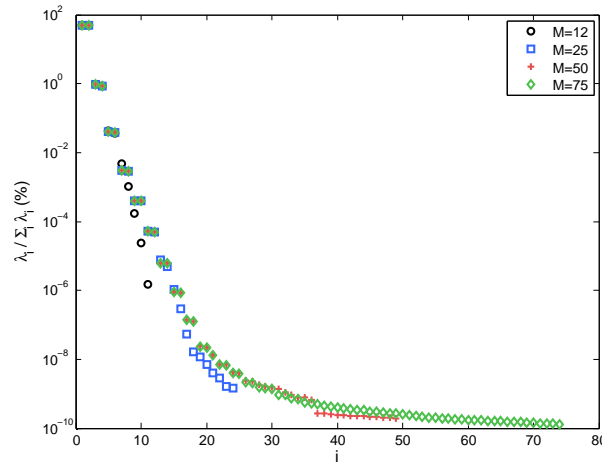
Un ensemble de  $M$  clichés adimensionnels relatifs aux champs  $\mathbf{q}^{(m)} = [\vartheta, u, v, p]^T$ , pour  $m = 1, \dots, M$  est échantillonné régulièrement sur l'intervalle de temps  $I_e$  à partir de la simulation précédente. Les clichés sont sélectionnés sur une durée couvrant un peu plus d'une période puisque l'analyse d'un système périodique dans l'annexe C.1 avait permis de mettre en évidence qu'il n'était pas nécessaire d'allonger l'intervalle d'échantillonnage.

La construction du modèle d'ordre réduit est basée sur une approche POD multivariées si bien que

<sup>4</sup>Dans le cas présent, le profil est horizontal et l'angle d'incidence est introduit au niveau de la condition limite de non-réflexion. Par conséquent les flux issus d'*elsA* fournis dans le repère du profil et non de l'écoulement correspondent aux coefficients  $C_x$  et  $C_z$ . Pour retrouver les valeurs des coefficients de traînée et de portance il sera donc nécessaire de pratiquer un changement de repère, correspondant à une rotation de l'angle d'incidence  $\alpha$ .

les modes POD sont calculés à partir des vecteurs propres du problème (2.81) et de la relation (2.82). Le produit scalaire adopté est celui défini par l'équation (4.52) avec  $\mathbf{P}_{\text{adim}} = \mathbf{I}_{n_v}$ , puisque les clichés extraits sont déjà sous forme adimensionnelle. Les modes POD obtenus seront donc eux aussi adimensionnels, ce qui est cohérent pour effectuer ensuite les produits scalaires qui interviendront au cours de la construction du modèle d'ordre réduit.

La figure 5.6 représente le spectre des valeurs propres de l'opérateur de corrélation spatial formé pour la méthode des clichés proposée par Sirovich [218]. Le graphique indique le pourcentage  $\tilde{\eta}_i$  d'énergie captée par chaque valeur propre qui a été défini à l'équation (3.44). Les premières valeurs propres captent la quasi-totalité de l'énergie contenue dans le système : ainsi pour  $M = 50$ , les  $q = 6$  premières valeurs propres captent déjà  $\eta_q = 99,99\%$  de l'énergie totale et les  $q = 10$  premières valeurs propres conduisent à une erreur de troncature  $\epsilon_q = \sum_{i=q+1}^r \lambda_i$  de l'ordre de  $10^{-4}$ . Le nombre de clichés choisis sur une période d'oscillation du système semble avoir assez peu d'influence sur les valeurs propres, bien que quelques différences soient visibles sur le spectre, notamment pour de faibles valeurs de  $M$ . Les valeurs propres apparaissent par paires, au moins pour  $i \in \llbracket 1; 20 \rrbracket$  et lorsque le nombre de clichés utilisé est suffisamment grand. Ce regroupement par paires peut être attribué à la symétrie de translation des ondes progressives caractérisant l'écoulement [17] et est habituellement observé, que ce soit dans le sillage d'un profil portant comme celui traité ici [34, 73, 117], ou encore dans le sillage d'un cylindre de section circulaire [216] ou carrée [50].



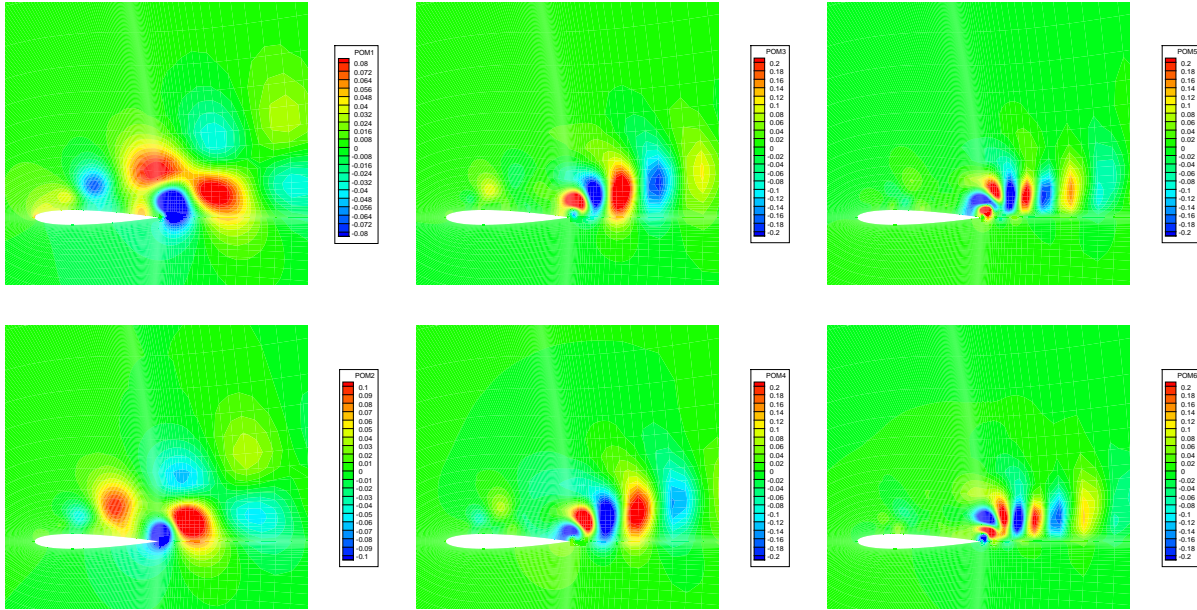
**Figure 5.6 :** Spectre des valeurs propres associées aux modes POD calculés avec la méthode des clichés pour quelques valeurs du nombre  $M$  de clichés.

À la vue de ce spectre, on peut supposer qu'un modèle d'ordre réduit construit à l'aide d'une base POD tronquée à six modes seulement devrait déjà permettre de reconstruire correctement la solution. Au-delà de vingt modes POD, il semble peu utile de rajouter des modes supplémentaires dans la base puisque la pente du spectre devient moins importante, ce qui traduit le fait que chaque nouveau mode ajouté n'apporte que peu d'information complémentaire.

La figure 5.7 donne l'aspect des six premiers modes POD de pression calculés par la méthode des clichés. Il s'agit en réalité d'un abus de langage puisque les « modes POD » représentés correspondent seulement à la partie  $\varphi_p^{(j)}$  du vecteur  $\varphi^{(j)} = [\varphi_\theta^{(j)} \ \varphi_u^{(j)} \ \varphi_p^{(j)}]^T$ . Le regroupement par paires des valeurs propres déjà observé auparavant se retrouve dans l'allure des modes POD : les structures spatiales des modes sont similaires pour les paires (1-2), (3-4), (5-6),... et présentent uniquement un déphasage spatial. Le premier couple de modes POD met en évidence une structure relative à l'émission simultanée de deux vortex contrarotatifs convectés ensuite en aval sur deux axes distincts comme dans le cas des allées de Von Kármán, alors que les couples de modes suivants font plutôt apparaître des structures de vortex émis successivement. De plus, il apparaît que la taille des structures spatiales mises en évidence décroît lorsque l'indice des modes POD associés à des valeurs propres de plus en plus faibles augmente. Ceci signifie que les modes POD d'ordre élevé rendent accessibles des détails de plus en plus fins qui permettront par conséquent de reconstruire de plus en plus précisément la solution. Ce phénomène est très similaire à ce qui se produit avec des dévelop-



pements en séries trigonométriques où les longueurs d'ondes des fonctions de la série sont de plus en plus faibles.



**Figure 5.7 :** Allure des six premiers modes POD de pression  $\varphi_p^{(j)}$  calculés à partir de  $M = 50$  clichés solutions de l'écoulement autour du profil NACA0012 en forte incidence.

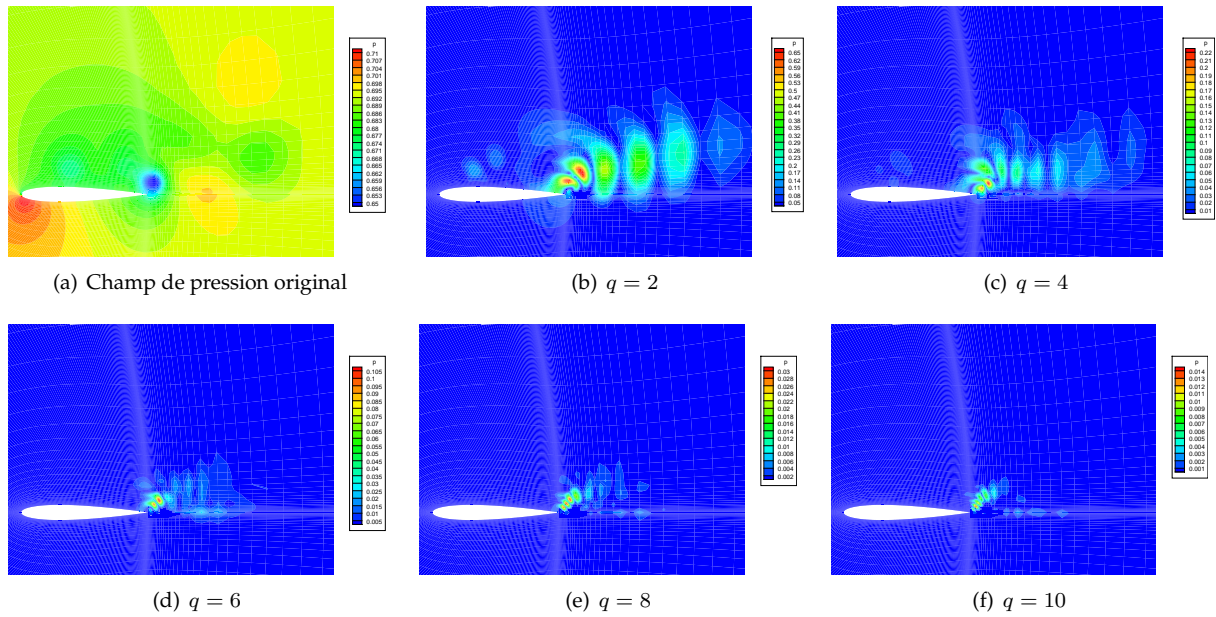
Les clichés initialement utilisés pour calculer les modes POD sont dans un premier temps reconstruits en évaluant les amplitudes modales telles qu'elles satisfassent la décomposition (2.83). Pour cela, il suffit d'utiliser l'orthonormalité des modes POD  $\varphi^{(j)}$  au sens du produit scalaire de  $(L^2(\Omega))^{n_v}$  de sorte que les amplitudes modales associées aux clichés de départ  $\mathbf{q}^{(j)}$  soient données par

$$a_j^{(m)} = \langle \mathbf{q}^{(m)} - \bar{\mathbf{q}}, \varphi^{(j)} \rangle = \int_{\Omega} (\mathbf{q}^{(m)} - \bar{\mathbf{q}}) \cdot \varphi^{(j)} \, d\Omega, \quad \forall (m, j) \in \llbracket 1; M \rrbracket \times \llbracket 1; q \rrbracket. \quad (5.50)$$

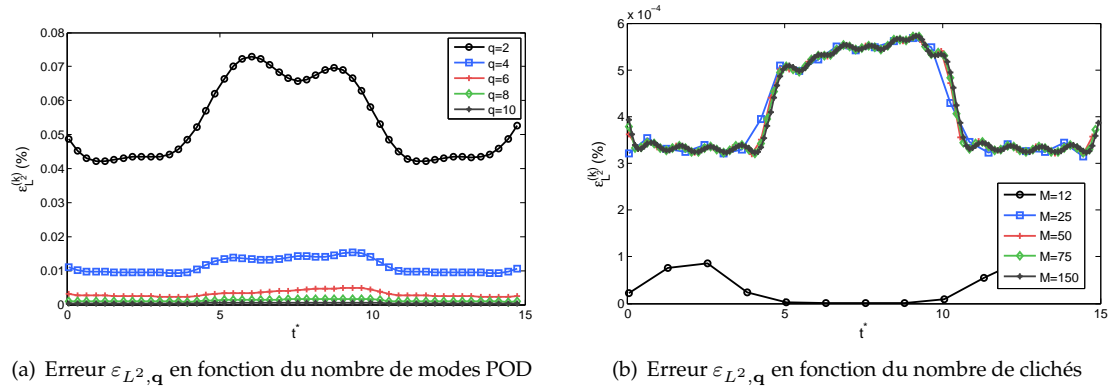
Une fois les amplitudes modales calculées, les clichés sont reconstruits comme combinaison linéaire des modes POD au moyen de la relation (2.83). La figure 5.8(a) représente le champ de pression  $p$  correspondant au premier cliché  $\mathbf{q}^{(1)}$  de la base de clichés initiale. Le champ de pression est alors reconstruit en utilisant une base POD dont on augmente au fur et à mesure le nombre de vecteurs  $q$ . Les figures 5.8(b) à (f) représentent l'erreur relative de reconstruction  $100 |p - p_r| / |p_r|$ , où  $p_r$  est le champ de pression du cliché original et  $p$  est la pression reconstruite à l'aide des amplitudes (5.50) introduites dans la décomposition POD en conservant de deux à dix modes. L'erreur de reconstruction reste très faible, même lorsque peu de modes sont pris en compte : ainsi, avec seulement  $q = 2$  modes, l'erreur relative maximale dans le domaine est seulement de l'ordre 0,65%. En augmentant le nombre de modes dans la base, l'erreur diminue jusqu'à être progressivement noyée dans les erreurs numériques lorsque l'ensemble des modes POD est conservé. La figure 5.8 fournit en outre des informations sur la répartition spatiale des erreurs de reconstruction à un instant donné. En particulier, on observe que lorsque seuls les deux premiers modes sont conservés, l'erreur fait apparaître des structures correspondant à celles des modes 3 et 4 (voir figure 5.7) ; de la même manière, lorsque les quatre premiers modes sont conservés, on observe que les structures de l'erreur correspondent aux modes 5 et 6 qui sont négligés — plus d'autres composantes relatives aux modes d'indices supérieurs mais qui ne sont pas nettement visibles ici.

L'erreur introduite à l'équation (3.41) est redéfinie d'une manière plus générale pour tenir compte de l'erreur commise dans tout le domaine et pour toutes les variables au moyen de l'expression

$$\varepsilon_{L^2, \mathbf{q}} = \frac{\|\mathbf{q}^{(m)} - \mathbf{q}_r^{(m)}\|_{L^2}}{\|\mathbf{q}^{(m)}\|_{L^2}}. \quad (5.51)$$

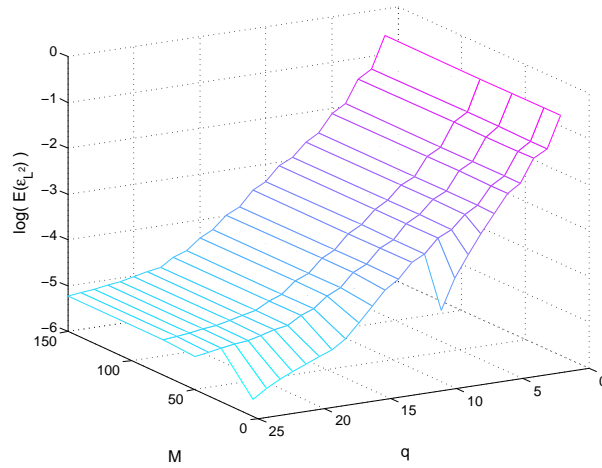


**Figure 5.8 :** Erreur de reconstruction relative entre le champ de pression initial  $p$  représenté en (a) et le champ  $p_r$  reconstruit à l'aide de la base POD constituée de  $q = 2, 4, 6, 8$  et  $10$  modes POD (figures (b) à (f)).



**Figure 5.9 :** Erreur  $\varepsilon_{L^2,q}$  commise lors de la reconstruction des clichés de départ (a) lorsque la base des clichés est constituée de  $M = 50$  clichés et que le nombre de modes POD varie de  $q = 2$  à  $10$ , et (b) lorsque le nombre de clichés varie de  $M = 12$  à  $150$  et que la base des modes POD est constituée de  $q = 10$  modes.

L'erreur  $\varepsilon_{L^2,q}$  est alors estimée pour chaque cliché  $m = 1, \dots, M$  de la base initiale. Le résultat est présenté sur la figure 5.9(a) pour une base de clichés initialement composée de  $M = 50$  clichés et une base POD dont on augmente le nombre de modes de  $q = 2$  à  $10$ . L'ordonnée donne l'erreur  $\varepsilon_{L^2,q}$  et les abscisses indiquent le temps adimensionnel  $t^*$  auquel les clichés ont été pris. L'erreur  $\varepsilon_{L^2,q}$  est déjà très bonne lorsqu'un petit nombre de modes POD est conservé dans la base et on observe une convergence assez rapide de l'erreur vers des valeurs très faibles lorsque la taille de la base POD augmente : l'erreur  $\overline{\varepsilon_{L^2,q}}$  calculée comme la moyenne sur l'ensemble des clichés passe de  $0,05\%$  à  $4 \times 10^{-4}\%$  lorsque  $q$  varie de  $2$  à  $10$ . Par ailleurs, la figure 5.9(b) représente la même erreur  $\varepsilon_{L^2,q}$  lorsque le nombre de clichés dans la base initiale varie de  $M = 12$  à  $150$  et que la dimension de la base POD est maintenue à une valeur constante  $q = 10$ . On observe une certaine convergence de l'erreur lorsque le nombre clichés augmente, mais le cas  $M = 12$  produit une erreur nettement plus faible ce qui peut sembler incohérent. En réalité, ceci s'explique par le fait que la taille de la base POD utilisée est maintenue constante et que dans ce cas le nombre de



**Figure 5.10 :** Convergence de l'erreur infinie de reconstruction  $\varepsilon_\infty$  en fonction du nombre  $M$  de clichés contenus dans la base POD initiale et du nombre  $q$  de modes POD retenus dans la base réduite.

modes POD utilisés correspond pratiquement à la base complète — puisque  $r = 11$  lorsque  $M = 12$  — et par conséquent l'erreur de troncature est extrêmement faible : dans ce cas, le pourcentage d'énergie perdue  $1 - \eta_q$  est de l'ordre de  $\mathcal{O}(10^{-6})$  alors que pour les autres valeurs de  $M$  il est de l'ordre de  $\mathcal{O}(10^{-4})$ .

Finalement, le graphique de la figure 5.10 résume la convergence de l'erreur moyenne  $\bar{\varepsilon}_{L^2, q}$  lorsque le nombre de clichés et de modes POD contenus dans la base de projection POD augmente. Le nombre de clichés employés a peu d'influence sur la convergence de l'erreur. On note cependant que pour de faibles valeurs de  $M$ , l'erreur décroît plus rapidement pour la même raison que celle invoquée ci-dessus, puisque la troncature effectuée dans la base est moins importante.

#### Remarque 5.1 : Calcul des modes POD pour des clichés de grande dimension

Les clichés comportent ici un nombre important  $N_v = N n_v$  de degrés de liberté qui rend le calcul des modes POD relativement coûteux, même avec la méthode des clichés. La matrice des clichés ne peut en effet pas être stockée intégralement en mémoire et le calcul de l'opérateur de corrélation s'effectue donc de manière itérative. Puisque les modes POD peuvent s'obtenir de façon équivalente comme les solutions d'une décomposition en valeurs singulières (voir l'annexe B.3 p.222), il est possible de profiter de certaines méthodes approchées pour le calcul de la SVD. L'idée sous-jacente est d'effectuer initialement la SVD d'une matrice des clichés tronquée qui est améliorée de manière itérative au moyen des clichés absents de la matrice tronquée de départ. Le principe de la méthode est expliqué en détail par Chahlaoui et al. [43] et l'algorithme implémenté ici est celui proposé par Mastronardi et al. [162]. Le coût de calcul des modes POD est alors nettement réduit au moyen de cette technique mais les erreurs de reconstruction deviennent d'autant plus élevées que la matrice tronquée des clichés employée pour initialiser l'algorithme est de faible dimension. Une technique alternative qui n'a pas été évaluée ici consiste à calculer les modes POD au moyen d'une méthode de Lanczos comme l'a proposé Fahl [70].

## 5.5 Réponse du modèle d'ordre réduit de l'écoulement autour du profil NACA0012

Dans le paragraphe précédent, le calcul des modes POD de l'écoulement avait permis de constater que les clichés pouvaient être reconstruits de façon très satisfaisante au moyen des amplitudes  $\mathbf{a}^{(m)}$  issues de la projection des clichés sur les modes POD. En effet, les erreurs de reconstruction obtenues indiquent qu'un petit nombre de modes suffit à reproduire le phénomène de détachement de vortex dans le sillage du profil.

L'objectif est à présent de calculer les amplitudes modales  $\mathbf{a}(t)$  à l'aide du modèle d'ordre réduit et de comparer les erreurs commises par rapport aux amplitudes de référence  $\mathbf{a}^{(m)}$ . Le modèle d'ordre réduit est défini par le système d'équations (5.8) et les expressions des coefficients données à l'équation (5.9). Dans un premier temps, la solution produite par le modèle d'ordre réduit non corrigé tel qu'il a été défini au paragraphe 5.2.1 sera étudiée et comparée à la solution de référence. Les deux paragraphes suivants résumeront un certain nombre de tentatives plus ou moins fructueuses pour améliorer la réponse.

Auparavant, signalons qu'une comparaison des contributions convectives et diffusives d'une part, et surfaciques et volumiques d'autre part est fournie dans l'annexe F.1. Celle-ci révèle que la partie diffusive joue surtout un rôle au travers de la partie moyenne  $\bar{q}$  qui intervient dans la définition des coefficients constants et linéaires. Par ailleurs, les contributions surfaciques issues de l'intégration par parties des termes diffusifs sont peu significatives et pourraient éventuellement être négligées. Cependant, l'ensemble des contributions est pris en compte afin de minimiser les erreurs dans l'évaluation des coefficients qui pourraient conduire à une déstabilisation de la réponse du modèle d'ordre réduit. Enfin, le nombre de Reynolds est calculé à partir de la définition (5.1) en évaluant la viscosité avec la loi de Sutherland :  $\mu = \mu_{273} (\theta/\theta_{273})^{3/2} (\theta_{273} + S)/(\theta + S)$  où  $\mu_{273}$  est une viscosité de référence à 273° Kelvin,  $\theta_{273} = 273^\circ$  Kelvin et  $S = 110,4$  est la constante de Sutherland. La valeur du nombre de Prandtl est quant à elle prise constante et égale à  $Pr = 0,72$ .

### 5.5.1 Réponse du modèle d'ordre réduit non corrigé

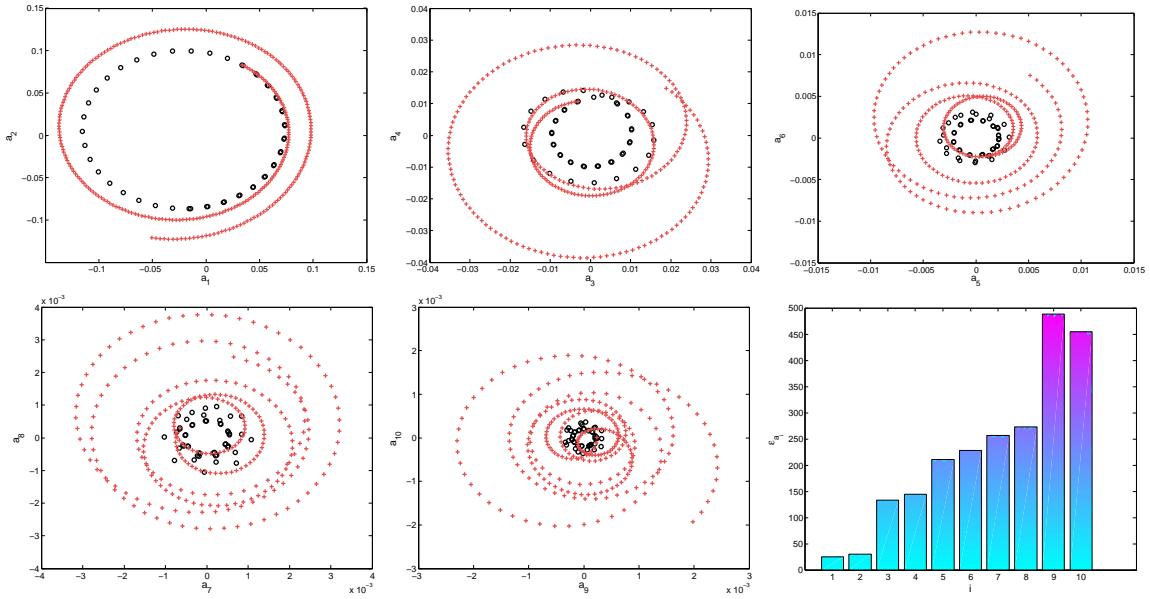
Le modèle d'ordre réduit POD-Galerkin de l'écoulement autour du profil NACA0012 est à présent mis à l'épreuve pour reproduire le champ aérodynamique au moyen des amplitudes modales  $\mathbf{a}(t_m)$  associées aux modes POD. La qualité du modèle d'ordre réduit est principalement évaluée au moyen des amplitudes qui doivent être les plus proches possible des amplitudes de référence  $\mathbf{a}^{(m)}$ . À partir d'elles, il devient alors possible de comparer n'importe quelle grandeur aérodynamique telle que le coefficient de portance  $C_L$  défini à l'équation (5.49) par exemple. On définit pour commencer l'erreur générique

$$\bar{\varepsilon}_\chi = 100 \frac{\|\chi_r(t) - \chi(t)\|_{L^2(I)}}{\|\chi_r(t)\|_{L^2(I)}} \quad \text{avec} \quad \|\chi(t)\|_{L^2(I)} = \sqrt{\int_{t_0}^{t_0+t_f} \chi^2(t) dt}, \quad (5.52)$$

qui mesure l'écart relatif moyen en temps entre la solution de référence  $\chi_r(t)$  et celle reconstruite avec le modèle d'ordre réduit notée simplement  $\chi(t)$ . L'erreur définie ci-dessus sera calculée en pratique sur l'intervalle de temps d'échantillonnage  $I_e$ . Elle permettra ainsi d'évaluer l'erreur moyenne commise pour chaque amplitude modale en choisissant  $\chi(t) = a_j(t)$ , ce qui conduira à l'erreur notée  $\bar{\varepsilon}_{a_j}$ , ou de la même manière à la déviation moyenne  $\bar{\varepsilon}_{C_L}$  observée sur le coefficient de portance reconstruit par le modèle d'ordre réduit en choisissant  $\chi(t) = C_L(t)$ . En outre, on définit à partir des matrices  $\mathbf{A}_r$  et  $\mathbf{A}$  de termes généraux  $a_i^{(j)}$  et  $a_i(t_j)$  respectivement l'erreur globale

$$\varepsilon_{\mathbf{A}} = 100 \frac{\|\mathbf{A}_r - \mathbf{A}\|_2}{\|\mathbf{A}_r\|_2}. \quad (5.53)$$

Pour commencer, le modèle d'ordre réduit est construit à l'aide des  $q = 10$  premiers modes POD puisque d'après les résultats du paragraphe 5.4.2 la base constituée de ces modes POD capture près de 100% de l'énergie du système. Par ailleurs, le nombre de clichés utilisés pour construire les modes POD semble avoir peu d'influence (voir la figure 5.9(b) p.145) et on optera ici pour une base de clichés constituée de  $M = 50$  vecteurs. Le modèle d'ordre réduit est alors intégré en temps à partir d'une solution initiale calculée à l'aide de la relation (5.24) avec  $\mathbf{q}_0 = \mathbf{q}^{(1)}$ . L'intégration en temps du système réduit fournit les amplitudes modales sur le même intervalle de temps  $I_e$  que celui utilisé pour l'échantillonnage des clichés, de façon à pouvoir comparer la solution de référence à celle calculée au moyen du modèle d'ordre réduit.

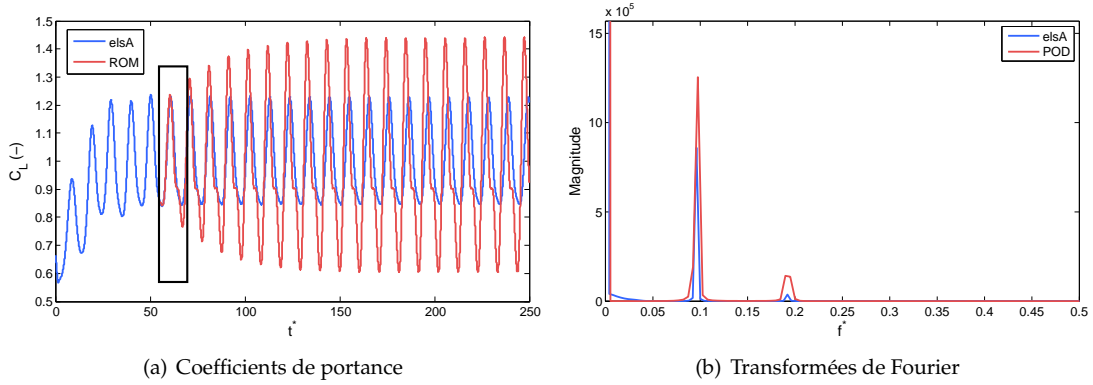


**Figure 5.11 :** Comparaison des portraits de phase des amplitudes modales  $\{a_j(t)\}_{j=1}^{10}$  reconstruites par le modèle d'ordre réduit non corrigé (+) avec celles issues de la projection des clichés (o). Le dernier graphique représente de l'erreur moyenne relative  $\overline{\varepsilon_{a_j}}$ .

Le résultat est tracé sous forme de portraits de phase sur la figure 5.11 où les amplitudes modales  $\{a_j\}_{j=1}^{10}$  associées aux modes POD sont comparées aux amplitudes  $a_j^{(m)}$  de référence. Le modèle d'ordre réduit non corrigé produit des amplitudes qui semblent diverger au fil du temps. L'amplitude des  $a_j(t)$  calculées par le modèle d'ordre réduit est en général trop élevée et les cycles limites sont mal reproduits, tant au niveau de la forme que de l'étendue qu'ils couvrent. L'écart entre les amplitudes issues de la projection des clichés et celles reconstruites par le modèle d'ordre réduit mesuré pour chaque mode au moyen de l'erreur  $\overline{\varepsilon_{a_j}}$  est représenté sur le dernier graphique de la figure 5.11. Une forte augmentation de l'erreur relative est constatée lorsque l'indice  $j$  des modes augmente : d'environ 25% pour le premier mode, l'erreur relative moyenne sur l'amplitude associée aux derniers modes POD passe à près de 500%! Toutefois, plus l'indice  $j$  est grand, et plus les variations des amplitudes sont faibles : le grand axe de l'ellipse du premier portrait de phase a une amplitude de l'ordre de  $\mathcal{O}(10^{-1})$ , alors que sur le dernier portrait de phase, les variations maximales de  $a_{10}$  sont de l'ordre de  $\mathcal{O}(10^{-3})$ . Par conséquent, le fait que l'écart entre les amplitudes de référence et celles calculées par le modèle d'ordre réduit augmente avec l'indice  $j$  est en quelque sorte compensé par la diminution des variations des amplitudes quand  $j$  augmente. Il s'ensuit que les grandeurs macroscopiques ne sont pas si éloignées de la solution de référence.

Pour illustrer ceci, on évalue l'erreur commise sur le coefficient de portance qui correspond à une grandeur macroscopique  $\chi$  annexe au modèle d'ordre réduit. Le coefficient de portance se déduit indirectement des amplitudes modales puisque les efforts de portance  $F_L(t) \approx \int_{\Gamma} p(\mathbf{x}, t) n_z ds$  entrant dans la définition (5.49) du coefficient de portance font intervenir la pression  $p$  qui est reconstruite au moyen de la décomposition orthogonale propre  $p(\mathbf{x}, t) = \bar{p}(\mathbf{x}) + \sum_{j=1}^q a_j(t) \varphi_p^{(j)}(\mathbf{x})$ , de sorte qu'il existe bien une relation  $C_L(t) = \hat{g}(\mathbf{a}(t))$ . Le modèle d'ordre réduit est intégré sur une durée plus longue que celle ayant servi pour l'échantillonnage des clichés afin d'observer le comportement à long terme de la réponse. La figure 5.12(a) confronte ainsi le coefficient de portance calculé au moyen du modèle d'ordre réduit à celui de référence sur une durée correspondant à une quinzaine de détachements de vortex. Le rectangle noir correspond à l'intervalle de temps d'échantillonnage. Au cours de cette période, la solution reconstruite avec le modèle d'ordre réduit est en relativement bon accord avec la solution de référence puisque l'erreur calculée pour le modèle d'ordre réduit caractérisé par  $q = 10$  modes POD et  $M = 50$  clichés ne s'élève qu'à  $\overline{\varepsilon_{C_L}} = 3,8\%$ . Toutefois, lorsque l'intégration en temps du modèle d'ordre réduit est poursuivie en dehors de la fenêtre d'échantillonnage, une amplification anormale de l'amplitude est observée. Bien que les oscillations ne di-

vergent pas, l'amplitude crête-à-crête<sup>5</sup>  $\Delta_{cc,POD} = 0,83$  est pratiquement deux fois plus importante que celle de référence  $\Delta_{cc,r} = 0,38$ . Malgré cette imperfection concernant les amplitudes d'oscillation, la fréquence est correctement évaluée : les transformées de Fourier des signaux de portance données sur la figure 5.12(b) permettent d'identifier le pic principal à  $f_{S,r}^* = 0,0960$  tandis que le modèle d'ordre réduit produit un pic à  $f_{S,POD}^* = 0,0975$ , ce qui correspond à des périodes de Strouhal respectivement égales à  $T_{S,r} = 2,95 \times 10^{-2}$  s et  $T_{S,POD} = 2,91 \times 10^{-2}$  s, soit une erreur ne dépassant pas 1,5%.



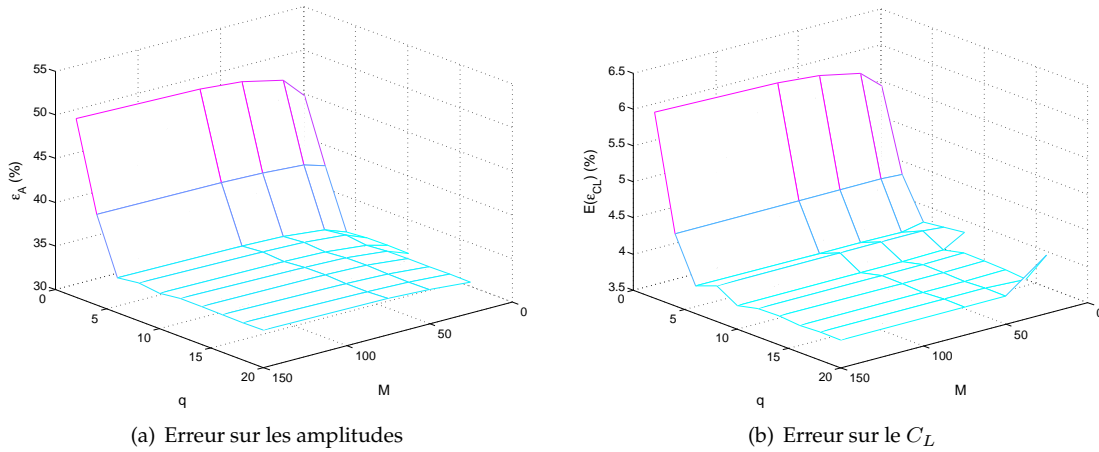
**Figure 5.12 :** Comparaison des coefficients de portance calculés avec elsA et le modèle d'ordre réduit non corrigé sur une longue durée (a) et comparaison des transformées de Fourier associées (b).

Avant de conclure sur la qualité du modèle d'ordre réduit, il est important de contrôler si l'augmentation du nombre de modes POD  $q$  contenus dans la base de projection et/ou du nombre  $M$  de clichés employés pour calculer les modes POD permet(tent) d'améliorer la réponse. Plusieurs critères de convergence peuvent être adoptés et on retiendra ici l'erreur globale  $\varepsilon_A$  définie à l'équation (5.53) pour les amplitudes modales, ainsi que l'erreur  $\overline{\varepsilon_{C_L}}$  définie à partir de l'équation (5.52) pour le coefficient de portance. Le même type d'analyse de convergence de l'erreur commise sur les clichés reconstruits en fonction de  $M$  et  $q$  avait été effectué au paragraphe 5.4.2 en utilisant les amplitudes modales issues de la projection des clichés  $\mathbf{a}^{(m)}$  alors que maintenant les amplitudes  $\mathbf{a}(t_m)$  produites par le modèle d'ordre réduit sont employées. Le résultat est tracé sur les graphiques de la figure 5.13 : quel que soit le type d'erreur considéré, une décroissance de l'erreur de reconstruction est toujours observée lorsque le nombre  $q$  de modes POD conservés dans la base de projection augmente mais la réponse du modèle d'ordre réduit converge très rapidement vers une valeur asymptotique de l'erreur, ce qui coïncide avec les résultats obtenus au paragraphe 3.3 pour le système dynamique linéaire. Ce phénomène est encore plus prononcé pour l'erreur  $\varepsilon_A$  qui semble constante dès que  $q \geq 6$ . Par ailleurs, le nombre  $M$  de clichés employés au départ semble avoir très peu d'influence sur l'erreur de reconstruction. Enfin il faut signaler que l'erreur relative moyenne commise sur le coefficient de portance reste faible — de l'ordre de  $\overline{\varepsilon_{C_L}} = 3,9\%$  — malgré une erreur de reconstruction globale assez élevée pour les amplitudes — de l'ordre de  $\varepsilon_A = 33\%$ .

## 5.5.2 Introduction de termes correctifs dans le modèle d'ordre réduit

La réponse du modèle d'ordre réduit non corrigé révèle bien le manque de stabilité symptomatique de l'approche POD-Galerkin. Par conséquent le modèle d'ordre réduit est corrigé pour commencer au moyen des techniques proposées par Cazemier *et al.* [42] et Kalb et Deane [124].

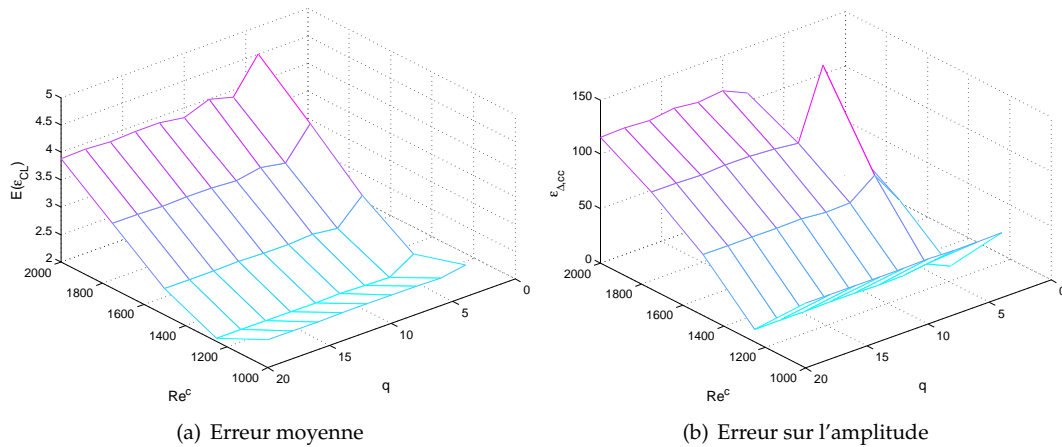
<sup>5</sup>L'amplitude crête-à-crête est calculée pour un signal périodique établi comme la différence entre la valeur maximale et la valeur minimale sur une période d'oscillation. Ici, les trois dernières valeurs des extrema sont relevées lorsque les oscillations sont établies. Ces valeurs permettent de calculer un maximum et un minimum moyens qui serviront à évaluer l'amplitude crête-à-crête du coefficient de portance  $\Delta_{cc} = \overline{C_{L,max}} - \overline{C_{L,min}}$ . À partir de cette quantité, on définit l'erreur relative  $\varepsilon_{\Delta_{cc}} = 100|\Delta_{cc,r} - \Delta_{cc,POD}|/|\Delta_{cc,r}|$ .



**Figure 5.13 :** Convergence en fonction du nombre  $M$  de clichés et du nombre  $q$  de modes POD de l'erreur globale de reconstruction  $\varepsilon_A$  pour les amplitudes modales (a) et de l'erreur  $\overline{\varepsilon_{C_L}}$  pour le coefficient de portance (b).

### 5.5.2.1 Correction du nombre de Reynolds

La réponse du modèle d'ordre réduit peut être améliorée de manière significative en introduisant de la viscosité artificielle dans le système réduit. Il suffit donc de multiplier, comme au paragraphe 5.3.2.2, la viscosité réelle de l'écoulement par un facteur  $1 + \eta_a$  qui donne lieu au nombre de Reynolds corrigé  $Re^c$  défini à l'équation (5.26). La valeur de ce dernier est donc diminuée progressivement par rapport au nombre de Reynolds réel afin d'augmenter la dissipation introduite. La réponse du modèle d'ordre réduit est alors évaluée pour différentes valeurs du nombre de Reynolds corrigé  $Re^c$  et du nombre  $q$  de modes POD retenus dans le modèle d'ordre réduit. La qualité de la réponse reproduite est évaluée d'une part au moyen de l'erreur  $\overline{\varepsilon_{C_L}}$  qui décrit la précision de la réponse sur la durée d'échantillonnage, et d'autre part à l'aide de l'erreur  $\varepsilon_{\Delta_{cc}}$  sur l'amplitude crête-à-crête qui fournit une indication sur l'erreur à long terme. Le résultat est présenté sur la figure 5.14.



**Figure 5.14 :** Erreurs de reconstruction  $\overline{\varepsilon_{C_L}}$  et  $\varepsilon_{\Delta_{cc}}$  du modèle d'ordre réduit corrigé à l'aide d'une viscosité artificielle conduisant au nombre de Reynolds corrigé  $Re^c$ .

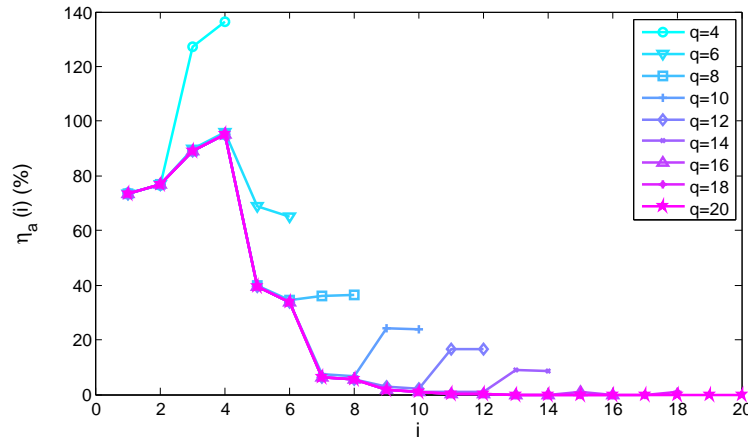
La valeur optimale du nombre de Reynolds corrigé pour un nombre de modes POD  $q$  donné est recherchée en procédant par dichotomie, bien que cette méthode requière un grand nombre d'essais. On observe en effet qu'en dessous de  $Re^c = 1000$ , l'erreur de reconstruction du modèle d'ordre réduit a tendance à augmenter à nouveau puisque les oscillations du système passent d'un régime périodique entretenu à un

régime amorti, ce qui est le signe d'une trop forte dissipation introduite dans le modèle d'ordre réduit. L'amélioration la plus importante est obtenue pour  $Re^c = 1250$  en conservant  $q = 10$  modes POD, ce qui correspond à un taux de dissipation artificiel  $\eta_a = 60\%$ . L'erreur sur les amplitudes d'oscillation tombe alors à  $\varepsilon_{\Delta_{cc}} = 11,1\%$ , tandis que l'erreur moyenne  $\overline{\varepsilon_{C_L}}$  sur la durée d'échantillonnage n'atteint que 2%.

### 5.5.2.2 Correction de Cazemier *et al.*

Afin d'éviter la recherche du nombre de Reynolds adéquat pour corriger le système dynamique, Cazemier *et al.* [42] proposent d'introduire une correction linéaire dans le modèle d'ordre réduit. Les coefficients de la correction sont calculés de façon à ce que l'énergie associée aux modes POD soit conservée sur une période d'oscillation du système dynamique, selon le principe exposé au paragraphe 5.3.2.2. Les coefficients  $\mathcal{D}_i$  de la correction peuvent être interprétés comme une contribution dissipative introduite pour chaque mode POD. En effet, la partie linéaire du modèle d'ordre réduit peut se ré-écrire  $\sum_{j=1}^q (\mathcal{L}_{ij} + \delta_{ij} \mathcal{D}_i) a_j$  et la correction peut être regroupée avec la partie diffusive des coefficients, de sorte que l'ensemble puisse être vu comme une modification du nombre de Reynolds pour chaque mode POD. Les termes diagonaux de  $\mathcal{L}_{ii}$  sont ainsi modifiés en  $\mathcal{L}_{ii}^C + (1/Re \mathcal{L}_{ii}^D + \mathcal{D}_i)$ . L'introduction de la correction linéaire revient alors à calculer la partie diffusive linéaire  $\mathcal{L}_{ii}^D$  avec un nombre de Reynolds corrigé  $Re_i^c$ , de sorte que  $1/Re \mathcal{L}_{ii}^D + \mathcal{D}_i = 1/Re_i^c \mathcal{L}_{ii}^D$ . L'effet du terme  $\mathcal{D}_i$  peut donc être exprimé en fonction d'un nombre de Reynolds ajouté  $Re^a$  en utilisant la relation (5.26) qui conduit à

$$Re_i^a = \frac{\mathcal{L}_{ii}^D}{\mathcal{D}_i}. \quad (5.54)$$



**Figure 5.15** : Pourcentage de viscosité artificielle  $\eta_a$  introduit pour chaque mode POD  $\varphi^{(i)}$  avec la correction proposée par Cazemier *et al.* Le résultat est tracé pour différentes dimensions  $q$  de la base de projection.

La viscosité artificielle  $\mu_a$  introduite de cette manière est alors quantifiée au moyen du ratio  $\eta_a = \mu_a/\mu = Re/Re^a$  calculé pour chaque mode POD. Le résultat est tracé sur la figure 5.15 pour différentes tailles de la base réduite. Les premiers modes POD ( $q \leq 4$ ) sont très fortement amortis puisque la viscosité ajoutée représente à chaque fois plus de 70% de la viscosité réelle. Par ailleurs, on observe que lorsque la taille de la base POD augmente, la dissipation totale introduite diminue : ainsi, lorsqu'on passe de  $q = 8$  à  $q = 10$  par exemple, la dissipation artificielle des 6 premiers modes reste identique mais elle chute fortement pour les modes 7 et 8 qui passent d'environ 35% lorsque la base POD est de taille  $q = 8$  à environ 7% pour  $q = 10$ . Enfin la dissipation associée aux modes 9 et 10 est de l'ordre de 24%, c'est-à-dire supérieure à celle des modes précédents. On peut avancer l'hypothèse que le surplus de dissipation associé aux deux derniers modes est lié à la troncature effectuée dans la base POD puisque ce phénomène se reproduit à chaque fois que la taille de la base POD est augmenté. On retrouve par ailleurs le fait que les modes se comportent par paire, ce qui est lié à leur structure et au spectre des valeurs propres mis en évidence. Enfin, lorsque la taille de la base POD devient assez grande, la viscosité artificielle associée aux modes d'indices élevés devient négligeable : ainsi, pour  $q \geq 16$ , les coefficients  $\mathcal{D}_i$  calculés deviennent positifs au-delà du 12<sup>e</sup>



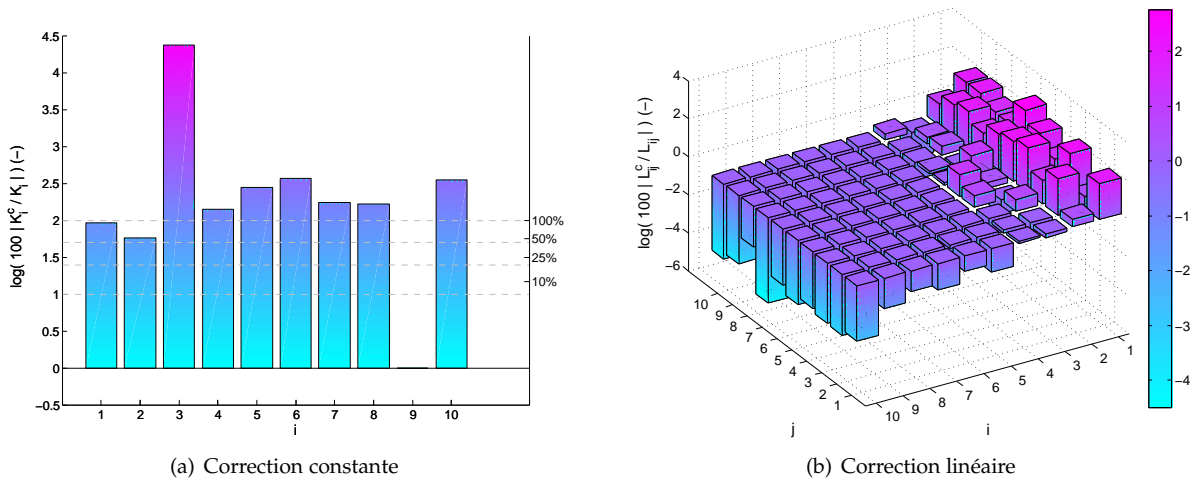
mode et sont donc réinitialisés à zéro comme l'ont suggéré *Cazemier et al.* afin que la correction apportée soit purement dissipative. La correction proposée par *Cazemier et al.* est donc simple à mettre en œuvre et possède un comportement assez intuitif puisque la dissipation totale introduite dans le modèle d'ordre réduit diminue lorsque la taille de la base de projection augmente. Ceci est en accord avec le fait que le manque de dissipation du modèle d'ordre réduit peut être attribué en partie à la troncature effectuée dans la base POD. En conclusion, la correction agit efficacement et les amplitudes sont correctement atténuées. L'erreur moyenne  $\overline{\varepsilon_{C_L}}$  sur la durée d'échantillonnage reste globalement inférieure à 2% dès que  $q > 4$  et l'écart  $\varepsilon_{\Delta_{cc}}$  sur l'amplitude crête-à-crête est de l'ordre de 10%.

### 5.5.2.3 Correction de Kalb et Deane

Une autre technique évaluée ici pour corriger le modèle d'ordre réduit est due à Kalb et Deane [124] qui avaient proposé de tirer parti des informations liées à l'erreur entre les amplitudes modales issues de la projection des clichés et celles provenant d'un premier calcul avec le modèle d'ordre réduit non corrigé. Suivant le principe exposé au paragraphe 5.3.3.2, des coefficients de correction constants  $\mathcal{K}_i^c$  et linéaires  $\mathcal{L}_{ij}^c$  sont calculés en projetant cette erreur sur la base formée des fonctions propres temporelles. Les corrections apportées par cette méthode sont plus complètes puisque l'ensemble des termes constants et linéaires est affecté, alors que l'approche de *Cazemier et al.* ne modifiait que les termes linéaires diagonaux de l'opérateur  $\mathcal{L}$ . Les corrections introduites sont évaluées au moyen des rapports

$$\Pi^{\mathcal{K}} = \log \left| 100 \frac{\mathcal{K}_i^c}{\mathcal{K}} \right| \quad \text{et} \quad \Pi^{\mathcal{L}} = \log \left| 100 \frac{\mathcal{L}_{ij}^c}{\mathcal{L}} \right| \quad (5.55)$$

qui quantifient la correction par rapport aux coefficients du modèle d'ordre réduit non corrigé.

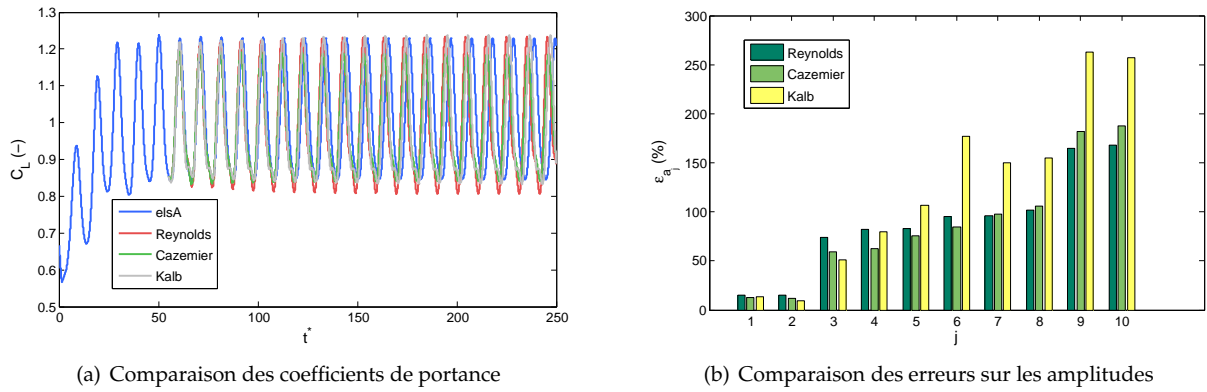


**Figure 5.16 :** Contribution relative des termes correctifs constants et linéaires de la méthode de *Kalb et Deane* par rapport aux coefficients initiaux.

Les corrections apportées sont représentées sur la figure 5.16 au moyen des rapports définis ci-dessus. Il s'avère que pour un certain nombre de coefficients, la correction modifie de manière très importante le coefficient initial : ainsi, le coefficient constant relatif au 3<sup>e</sup> mode est multiplié par plus de 200 et, en dehors du 2<sup>e</sup> et du 9<sup>e</sup> mode, les corrections représentent toujours plus de 100% du coefficient constant non corrigé. Par ailleurs, les corrections linéaires se sont pas uniformément réparties : elles agissent principalement pour  $i = 1, 2$  et sur l'ensemble des colonnes  $j = 1, \dots, q$ , alors que les corrections sur les autres coefficients linéaires sont quasiment négligeables.

La correction proposée par *Kalb et Deane* permet d'obtenir une réponse corrigée satisfaisante puisque

l'erreur de reconstruction  $\overline{\varepsilon_{C_L}}$  sur la période d'échantillonnage tombe à 2,34% et celle sur l'amplitude crête-à-crête est de l'ordre de  $\varepsilon_{\Delta_{cc}} \approx 5\%$ . Toutefois, cette réponse aussi précise a été obtenue à l'aide d'une normalisation adéquate du terme correctif qui diffère de celle proposée par Kalb et Deane. En effet, le facteur de normalisation  $1/\langle a_j^{(m)}, a_j^{(m)} \rangle$  intervenant dans l'expression de la correction linéaire (5.45) est inversement proportionnel aux valeurs propres  $\lambda_j$  associées aux modes POD  $\varphi^{(j)}$ . Puisque les valeurs propres décroissent très rapidement lorsque l'indice  $j$  des modes POD augmente, les corrections linéaires deviennent très importantes et déstabilisent le modèle d'ordre réduit plus qu'elles ne le corrigent. Par conséquent, l'erreur a ici été normalisée par un facteur proportionnel à  $\delta t_c/M$ . Une constante de proportionnalité égale à 10 a été introduite pour atteindre une stabilisation satisfaisante du modèle d'ordre réduit. Ainsi, la méthode proposée par Kalb et Deane n'est pas totalement satisfaisante puisqu'elle nécessite encore la détermination empirique d'une constante afin de corriger correctement le modèle d'ordre réduit.



**Figure 5.17 :** Comparaisons (a) du coefficient de portance et (b) des erreurs sur les amplitudes modales calculés tous les deux par le modèle d'ordre réduit stabilisé par l'introduction de termes correctifs issus (i) d'un Reynolds basé sur une viscosité artificielle, (ii) de la correction de Cazemier et al. [42], et (iii) de la correction de Kalb et Deane [124].

Les graphiques de la figure 5.17 confrontent la réponse du modèle d'ordre réduit corrigé par les trois méthodes présentées dans ce qui précède. Dans tous les cas, la réponse a été obtenue avec une base POD constituée de  $q = 10$  modes POD. Sur le graphique 5.17(a), le coefficient de portance reconstruit au moyen des réponses corrigées par les trois méthodes coïncide pratiquement avec la courbe de référence, même si un déphasage très léger persiste. Le graphique 5.17(b) présente l'erreur de reconstruction  $\overline{\varepsilon_{a_j}}$  pour l'amplitude associée à chaque mode POD obtenue avec les différentes corrections. Globalement, l'erreur relative augmente avec l'indice du mode POD et atteint des valeurs toujours très importantes pour les derniers modes ; l'erreur a néanmoins été fortement réduite puisqu'elle ne culmine « plus qu'à » 250% environ alors qu'elle atteignait pratiquement le double sans correction (voir le dernier graphique de la figure 5.11). Les trois corrections proposées ici conduisent toutes à une réponse pratiquement équivalente d'un point de vue macroscopique. Toutefois, la correction proposée par Cazemier et al. semble être la plus efficace des trois méthodes comparées ici puisqu'elle permet d'amortir correctement le système réduit, sans avoir à déterminer de façon empirique une constante caractéristique de la correction. En outre, cette méthode se comporte idéalement puisque la dissipation introduite diminue lorsque la taille du modèle d'ordre réduit augmente. La méthode basée sur la modification du nombre de Reynolds peut certes fournir une réponse satisfaisante comme celle obtenue ici avec  $Re^c = 1250$ , mais la détermination du nombre de Reynolds adéquat nécessite une série (éventuellement) importante d'essais. Enfin la méthode de stabilisation de Kalb et Deane fournit aussi une réponse très satisfaisante, mais elle nécessite de normaliser correctement l'erreur employée pour la correction. Or aucune expression adéquate n'a été établie pour le facteur de normalisation qui est identifié de manière empirique après une série d'essais.

Dans le paragraphe qui suit, des méthodes de calibrage des coefficients basées sur la résolution d'un problème d'optimisation sont donc mises en œuvre afin de corriger le modèle d'ordre réduit de manière plus efficace.

### 5.5.3 Calibrage des paramètres du modèle d'ordre réduit

Les techniques de correction précédentes requièrent pour la plupart la détermination d'un ou plusieurs paramètres dont l'évaluation s'effectue empiriquement. Un moyen plus rigoureux de déterminer la correction à apporter consiste donc à résoudre un problème d'optimisation dans lequel une erreur de reconstruction  $\epsilon$  caractéristique de la réponse du modèle d'ordre réduit est minimisée. Le problème d'optimisation présenté au paragraphe 5.3.3.1 nécessite d'effectuer un certain nombre de choix concernant :

- l'algorithme d'optimisation mis en œuvre ;
- la formulation linéaire ou non-linéaire de l'erreur à minimiser ;
- la quantité dont on cherche à minimiser l'erreur ;
- les coefficients du modèle d'ordre réduit à calibrer ;
- l'intervalle dans lequel les coefficients à calibrer peuvent varier.

L'objectif n'étant pas ici de reprogrammer un algorithme d'optimisation, le choix est limité en fonction de ceux proposés dans Scilab. Parmi les méthodes disponibles pour les systèmes non-linéaires, l'algorithme de Levenberg-Marquardt [77, 143, 159] semble être un bon candidat puisqu'il permet de traiter des problèmes d'optimisation pour des fonctionnelles définies en particulier par les équations (5.35) à (5.38) pour une erreur générique. Toutefois, la version de cet algorithme proposée dans Scilab est limitée d'une part par l'impossibilité de tenir compte des contraintes éventuellement associées au problème d'optimisation, et d'autre part par le fait que l'algorithme ne traite pas les problèmes sur-déterminés — dans lesquels le nombre de paramètres  $p_c$  à calibrer doit être inférieur ou égal à la dimension  $n_\epsilon$  du vecteur d'erreur  $\mathbf{r}$ . Lorsqu'il est nécessaire de prendre en compte des contraintes ou de traiter un problème sur-déterminé, on aura recours à la méthode Quasi-Newton modifiée détaillée par Bonnans [33].

En ce qui concerne le choix entre une formulation linéaire ou non-linéaire de l'erreur qui doit être minimisée, Couplet [50] proposait — pour simplifier la formulation du problème — de linéariser l'erreur afin qu'elle ne dépende que des amplitudes modales de référence ou de leurs dérivées. La comparaison effectuée par Favier *et al.* [73] entre ces deux types de formulations a montré que la formulation non-linéaire du problème d'optimisation conduisait à de meilleurs résultats, ce que confirment les tentatives effectuées ici. Par conséquent, la formulation non-linéaire sera de préférence adoptée pour l'ensemble des fonctionnelles mises en œuvre dans ce qui suit. Les paramètres calibrés correspondent tout d'abord au nombre de Reynolds qui agit sur les termes diffusifs puis d'une manière plus générale à certains coefficients du modèle d'ordre réduit.

#### 5.5.3.1 Calibrage du nombre de Reynolds

En ce qui concerne les trois derniers points de l'énumération précédente, des éléments de réponse sont fournis dans la suite de ce paragraphe. Dans un premier temps, on s'interroge sur la meilleure définition de la quantité  $\epsilon$  à minimiser. Le problème d'optimisation est donc formulé pour cinq quantités différentes en utilisant (i) la fonctionnelle (5.39a) pour calibrer les amplitudes  $\mathbf{a}(t_m)$  (calibrage de Floquet), (ii) la fonctionnelle (5.39b) pour calibrer les dérivées des amplitudes modales  $\dot{\mathbf{a}}(t_m)$  (calibrage de Poincaré), (iii) la fonctionnelle (5.40) avec  $\chi = C_L$  lorsqu'on optimise le coefficient de portance, ou (iv) avec  $\chi = \mathcal{E}_T$  lorsqu'on s'intéresse plutôt à l'énergie totale, et enfin (v) la fonctionnelle (5.42) pour calibrer l'énergie  $\mathcal{E}_T^*$  captée par chaque mode POD. Le problème d'optimisation correspondant à chacune des erreurs minimisées repose sur une formulation non-linéaire de la fonctionnelle définie pour  $p_c = 1$  paramètre, celui-ci étant le nombre de Reynolds optimal  $Re^c$  commun à l'ensemble des modes POD. Afin d'introduire exclusivement de la dissipation dans le modèle d'ordre réduit ( $\eta_a \geq 0$ ), il est nécessaire d'imposer que le nombre de Reynolds corrigé  $Re^c$  soit inférieur ou égal au nombre de Reynolds réel  $Re$  de l'écoulement. En outre, il est indispensable que le nombre de Reynolds reste positif, de sorte que le problème d'optimisation (5.34) soit en outre soumis à la contrainte que  $\mathbf{w} = Re^c \in ]0; Re]$ . Pour tenir compte de cette contrainte additionnelle, il est nécessaire de travailler avec l'algorithme Quasi-Newton. Afin de juger de l'efficacité des différentes formulations, le problème d'optimisation est résolu dans chaque cas pour différentes tailles de la base de projection. La qualité de la reconstruction est alors évaluée au moyen des erreurs  $\overline{\epsilon_{C_L}}$  et  $\epsilon_{\Delta_{cc}}$  relatives au coefficient de portance, ainsi que des erreurs  $\epsilon_{\mathbf{A}}$  et  $\epsilon_{\dot{\mathbf{A}}}$  concernant les amplitudes modales et leurs dérivées.

## 5.5. Réponse du modèle d'ordre réduit de l'écoulement autour du profil NACA0012

$q$	$\epsilon$	$Re^c$	$\overline{\epsilon_{C_L}}$	$\epsilon_{\Delta_{cc}}$	$\epsilon_A$	$\epsilon_{\dot{A}}$
4	$\mathbf{a}^e - \mathbf{a}^g$	1 099, 16	2, 32	11, 67	20, 13	18, 83
	$\dot{\mathbf{a}}^e - \dot{\mathbf{a}}^g$	1 082, 63	2, 35	12, 62	20, 17	18, 83
	$C_L^g - C_L^e$	1 191, 20	2, 25	6, 06	20, 64	19, 76
	$\mathcal{E}_T^g - \mathcal{E}_T^e$	1 152, 38	2, 26	12, 33	20, 29	19, 19
	$\mathcal{E}_T^{*,g} - \mathcal{E}_T^{*,e}$	1 157, 35	2, 26	12, 61	20, 32	19, 25
6	$\mathbf{a}^e - \mathbf{a}^g$	1 125, 01	2, 12	10, 22	18, 66	17, 20
	$\dot{\mathbf{a}}^e - \dot{\mathbf{a}}^g$	1 117, 78	2, 13	9, 98	18, 68	17, 21
	$C_L^g - C_L^e$	1 219, 68	2, 03	7, 87	19, 10	17, 70
	$\mathcal{E}_T^g - \mathcal{E}_T^e$	1 133, 40	2, 10	11, 05	18, 66	17, 20
	$\mathcal{E}_T^{*,g} - \mathcal{E}_T^{*,e}$	1 134, 00	2, 10	11, 40	18, 66	17, 20
10	$\mathbf{a}^e - \mathbf{a}^g$	1 126, 26	2, 14	10, 99	18, 77	17, 36
	$\dot{\mathbf{a}}^e - \dot{\mathbf{a}}^g$	1 120, 23	2, 15	10, 48	18, 78	17, 37
	$C_L^g - C_L^e$	1 224, 09	2, 05	5, 42	19, 24	17, 94
	$\mathcal{E}_T^g - \mathcal{E}_T^e$	1 138, 06	2, 12	11, 03	18, 77	17, 36
	$\mathcal{E}_T^{*,g} - \mathcal{E}_T^{*,e}$	1 138, 95	2, 12	10, 90	18, 77	17, 36
14	$\mathbf{a}^e - \mathbf{a}^g$	1 127, 70	2, 13	10, 68	18, 77	17, 36
	$\dot{\mathbf{a}}^e - \dot{\mathbf{a}}^g$	1 119, 31	2, 15	10, 46	18, 78	17, 37
	$C_L^g - C_L^e$	1 221, 73	2, 05	5, 49	19, 23	17, 93
	$\mathcal{E}_T^g - \mathcal{E}_T^e$	1 137, 29	2, 12	10, 93	18, 76	17, 36
	$\mathcal{E}_T^{*,g} - \mathcal{E}_T^{*,e}$	1 138, 17	2, 12	10, 81	18, 76	17, 36
20	$\mathbf{a}^e - \mathbf{a}^g$	1 125, 05	2, 14	10, 83	18, 77	17, 35
	$\dot{\mathbf{a}}^e - \dot{\mathbf{a}}^g$	1 119, 23	2, 15	10, 43	18, 78	17, 36
	$C_L^g - C_L^e$	1 221, 67	2, 05	5, 48	19, 22	17, 91
	$\mathcal{E}_T^g - \mathcal{E}_T^e$	×	×	×	×	×
	$\mathcal{E}_T^{*,g} - \mathcal{E}_T^{*,e}$	1 138, 14	2, 12	10, 77	18, 76	17, 35

**Tableau 5.1** : Erreurs produites par le modèle d'ordre réduit dont le nombre de Reynolds a été calibré au moyen de diverses formulations de l'erreur  $\epsilon$  et pour différentes tailles  $q$  de la base de projection.

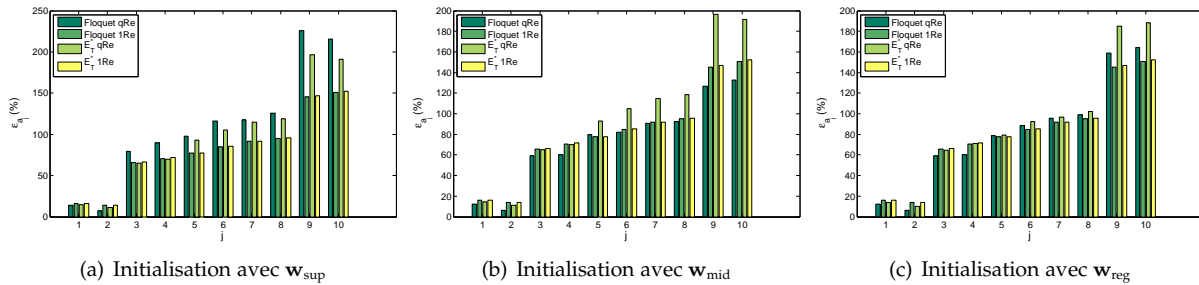
Les erreurs regroupées dans le tableau 5.1 indiquent que le nombre de Reynolds corrigé varie légèrement en fonction des quantités dont on minimise l'erreur : le calibrage basé sur le coefficient de portance est la méthode la moins dissipative, suivie des méthodes de calibrage de l'énergie totale puis de celles basées sur les amplitudes modales et leurs dérivées. Dans la plupart des cas, la dissipation introduite a légèrement tendance à décroître lorsque  $q$  augmente ce qui est cohérent avec l'hypothèse qu'une partie du manque de dissipation du modèle d'ordre réduit est due à la troncature effectuée dans la base POD. En dehors de la méthode de calibrage basée sur le coefficient de portance, les autres méthodes sont trop dissipatives sur le long terme et la réponse du modèle d'ordre réduit a tendance à être amortie au lieu d'être entretenue. On constate aussi sans surprise que l'erreur  $\overline{\epsilon_{C_L}}$  est toujours minimale pour le calibrage du coefficient de portance puisque dans ce cas le modèle d'ordre réduit est calibré pour minimiser l'erreur sur cette quantité. En contrepartie, l'erreur  $\epsilon_A$  sur les amplitudes modales est la plus importante avec cette méthode. Par conséquent, le calibrage devient inadapté dès lors qu'une autre quantité que le coefficient de portance devra être reconstruite avec les amplitudes issues de l'intégration du modèle d'ordre réduit.

Pour la suite de l'étude, on retiendra seulement la formulation de Floquet basée sur la minimisation de l'erreur sur les amplitudes, ainsi que la formulation reposant sur la minimisation du contenu énergétique de chaque mode POD via l'énergie  $\mathcal{E}_T^*$ . En effet, la formulation de Floquet vise à minimiser l'erreur sur les amplitudes modales, ce qui lui procure la capacité de reproduire correctement n'importe quelle quantité aérodynamique au moyen de la décomposition POD, dès lors que les amplitudes sont bien évaluées. Par ailleurs, la formulation basée sur la minimisation de l'erreur sur l'énergie présente l'avantage de faire appel à un critère physique qui permet de s'assurer de la conservation de l'énergie du système initial par le modèle d'ordre réduit. Les valeurs des erreurs fournies dans le tableau 5.1 indiquent que le calibrage portant sur l'énergie totale (réelle)  $\mathcal{E}_T$  de l'écoulement fournit pratiquement les mêmes résultats que lorsque le calibrage est effectué au moyen de l'erreur  $\mathcal{E}_T^*$ , ce qui justifie le fait que la formulation basée sur  $\mathcal{E}_T^*$  repose bien sur un critère énergétique. Cette dernière formulation est bien plus avantageuse que celle portant directement sur l'énergie totale réelle puisqu'elle ne nécessite pas de reconstruire à chaque fois le champ global afin d'évaluer l'énergie dans le domaine.

Le calibrage du nombre de Reynolds seul ne semble pas suffire puisque dans la plupart des cas le taux de dissipation artificiel déterminé par le processus d'optimisation est légèrement trop élevé. On cherche donc

à améliorer le processus de calibrage en attribuant à chaque mode POD  $\varphi^{(i)}$  un certain taux de dissipation  $\eta_{a,i}$  via un nombre de Reynolds  $Re_i^c$  comme dans la méthode de Cazemier *et al.* [42]. On impose par ailleurs la contrainte supplémentaire que le nombre de Reynolds associé à chaque mode ne peut varier que dans l'intervalle  $]0; Re]$  et on se limitera pour la suite à  $q = 10$  modes POD au maximum puisque l'introduction de modes supplémentaires dans le modèle d'ordre réduit n'améliore pas de manière significative l'erreur de reconstruction.

Le problème d'optimisation résolu ici devient beaucoup plus compliqué puisque le nombre de paramètres à optimiser passe de 1 à  $q$ . Par conséquent, les fonctionnelles à minimiser sont susceptibles de contenir des minima locaux vers lesquels l'algorithme d'optimisation risque de converger. Plusieurs résolutions du problème d'optimisation doivent donc être menées afin de vérifier que l'optimum global est bien atteint. À cet effet, le problème d'optimisation sera initialisé avec un jeu de paramètres  $w_0$  correspondant soit à la borne inférieure<sup>6</sup>  $w_{\text{inf}} \rightarrow 0$  de l'intervalle  $]0; Re]$  considéré, soit à la borne supérieure  $w_{\text{sup}} = Re$ , soit à la valeur moyenne  $w_{\text{mid}} = Re/2$ , ou encore en utilisant une régression linéaire en fonction de l'indice  $i$  du mode en partant de la valeur moyenne  $Re/2$  pour le premier mode :  $w_{i,\text{reg}} = Re/2 + (Re/2 - Re)/(q - 1) \times (i - 1)$ . Dans le dernier cas, l'initialisation consiste à imposer au départ un taux de dissipation de  $\eta_{a,1} = 100\%$  pour le premier mode et de  $\eta_{a,q} = 0\%$  pour le dernier mode.



**Figure 5.18 :** Comparaison des erreurs sur les amplitudes modales commises avec la méthode de calibrage de Floquet et celle basée sur l'énergie en optimisant le taux de dissipation pour chaque mode POD.

Les graphiques de la figure 5.18 détaillent ainsi pour  $q = 10$  et les différents types d'initialisation du problème d'optimisation, les erreurs sur les amplitudes modales obtenues avec la formulation de Floquet et la formulation énergétique. En outre, les erreurs obtenues dans le tableau 5.1 avec le calibrage d'un nombre de Reynolds commun à tous les modes sont rappelées à titre de comparaison<sup>7</sup>. L'amélioration constatée concerne surtout les premières amplitudes modales qui contribuent majoritairement à la reconstruction des grandeurs physiques au moyen de la décomposition orthogonale propre. L'initialisation effectuée au moyen de  $w_{\text{mid}}$  ou  $w_{\text{reg}}$  conduit à des résultats pratiquement identiques. En revanche, si l'algorithme est initialisé avec  $w_{\text{sup}}$ , les erreurs deviennent beaucoup plus importantes et dépassent même largement celles obtenues avec le calibrage 1Re. Enfin, l'initialisation  $w_{\text{inf}}$  avec la borne inférieure est peu adéquate car elle revient à introduire une dissipation quasiment infinie sur chacun des modes. L'algorithme ne parvient alors pas à converger vers un minimum autre que local et une majorité des nombres de Reynolds corrigés restent bloqués à la borne inférieure.

Les corrections apportées au modèle d'ordre réduit par le biais des calibrages 1Re et qRe ont permis d'améliorer encore la réponse du modèle d'ordre réduit par rapport aux corrections du paragraphe précédent. En particulier l'erreur sur les amplitudes modales qui était de  $\varepsilon_A = 19,53\%$  en utilisant une correction « à la main » avec  $Re^c = 1250$ , puis de  $\varepsilon_A = 18,39\%$  avec la correction de Kalb et Deane et enfin de  $\varepsilon_A = 14,50\%$  avec la correction de Cazemier *et al.*, a ainsi pu être réduite encore de quelques pour cents — près de 6% par rapport au premier cas, et environ 1% par rapport au dernier. Les corrections apportées jusqu'ici n'ont toutefois porté que sur les termes dissipatifs et la réponse peut encore être améliorée en agissant directement sur les coefficients du modèle d'ordre réduit.

<sup>6</sup>En pratique on prendra pour la borne inférieure un nombre de Reynolds égal à 1 puisque la valeur  $Re^c = 0$  conduirait à un taux de dissipation ajouté infini d'après la définition (5.27).

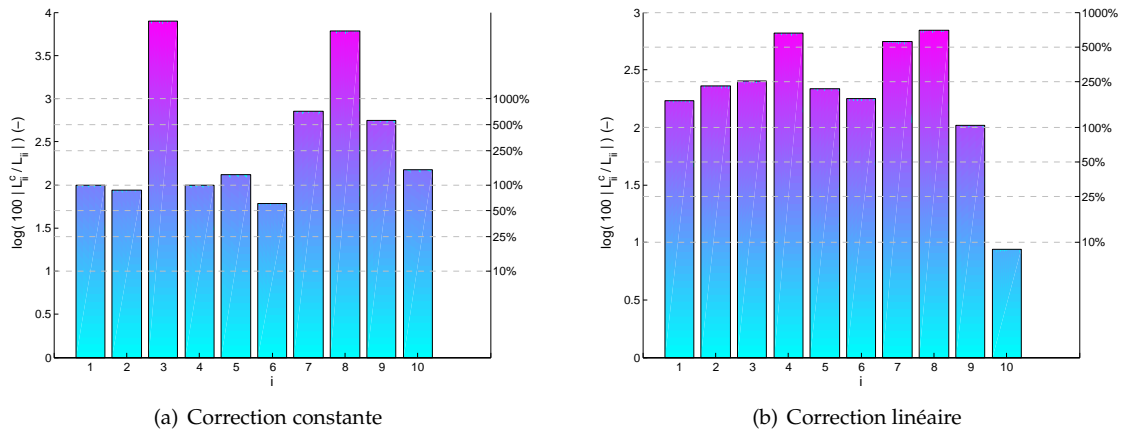
<sup>7</sup>On note 1Re le calibrage d'un seul nombre de Reynolds  $Re^c$  pour tous les modes POD par opposition au calibrage qRe dans lequel un nombre de Reynolds  $Re_i^c$  est déterminé pour chaque mode POD  $\varphi^{(i)}$ .

## 5.5.3.2 Calibrage des coefficients du modèle d'ordre réduit

Les paramètres du problème d'optimisation correspondent à présent à un certain nombre de coefficients du modèle d'ordre réduit. Le nombre de paramètres à corriger peut ainsi être augmenté en gardant à l'esprit que l'algorithme Quasi-Newton devra nécessairement être adopté si le problème devient sur-déterminé. Dans le cas contraire, on lui préférera l'algorithme de Levenberg-Marquardt puisque rien ne dicte ici les limites dans lesquelles les coefficients à calibrer doivent rester.

On envisage donc tout d'abord d'optimiser les termes constants  $\mathcal{K}$  et ceux de la diagonale du terme linéaire  $\mathcal{L}$ , ce qui représente un compromis entre les méthodes de correction de [Cazemier et al.](#) et [Kalb et Deane](#) puisque dans le premier cas seule la diagonale du terme linéaire était modifiée alors que dans le second cas, les termes constants et l'ensemble des termes linéaires étaient corrigés. On notera kDL ce calibrage qui fait intervenir  $p_c = 2q$  paramètres. Ensuite, on cherchera à calibrer comme dans la méthode de [Kalb et Deane](#) les  $q$  termes constants et l'ensemble des  $q^2$  termes linéaires si bien que ce calibrage sera désigné par KL.

La méthode de calibrage basée sur la conservation de l'énergie fournit des résultats de très mauvaise qualité par rapport à ceux résultant de la formulation de Floquet. Une explication probable est que le problème devient sur-déterminé puisque le vecteur d'erreur  $\mathbf{r}$  que la fonctionnelle  $\mathcal{J}^{\varepsilon_T^*}$  doit minimiser est seulement de taille  $n_\varepsilon = q$  alors que le nombre de paramètres à calibrer passe à  $p_c = 2q$  ou  $p_c = q + q^2$ . Par conséquent, on se concentrera pour les exemples qui suivent sur la formulation de Floquet, dans laquelle la dimension du vecteur d'erreur  $\mathbf{r}$  est  $n_\varepsilon = qM$ , ce qui permet de conserver un problème sous-déterminé tant que le nombre de clichés  $M$  reste assez grand ( $M \geq 2$  pour kDL et  $M \geq 1 + q$  pour KL).

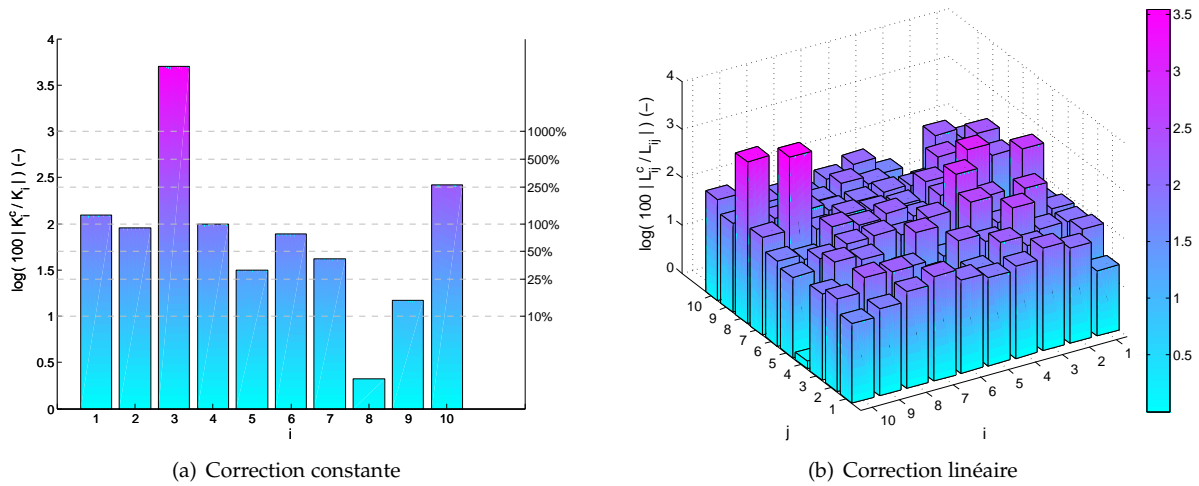


**Figure 5.19** : Contribution relative des termes constants et linéaires diagonaux calibrés avec la formulation de Floquet kDL. On représente le rapport entre les valeurs de la correction et celles des coefficients initiaux.

Le calibrage kDL conduit pour  $q = 10$  avec la formulation de Floquet aux erreurs suivantes :  $\overline{\varepsilon_{C_L}} = 1,67\%$ ,  $\varepsilon_{\Delta_{cc}} = 11,47\%$ ,  $\varepsilon_{\mathbf{A}} = 9,86\%$  et  $\varepsilon_{\dot{\mathbf{A}}} = 9,63\%$ . Cette formulation permet donc d'améliorer encore sensiblement l'erreur sur les amplitudes modales, certainement en raison du fait que la contribution convective des coefficients est aussi modifiée. Cependant, aucune contrainte n'a été imposée ici sur les paramètres, si bien que l'algorithme de Levenberg-Marquardt employé pour la résolution est libre de faire varier les paramètres dans  $\mathbb{R}^{p_c}$ . Les graphiques de la figure 5.19 représentent ainsi le rapport relatif entre les coefficients du modèle d'ordre réduit corrigé et ceux du modèle d'ordre réduit non corrigé. Les graphiques mettent alors en évidence que l'amélioration de la solution passe par une modification relativement importante des coefficients du modèle d'ordre réduit.

On peut espérer réduire encore l'erreur avec le calibrage KL puisque le nombre de paramètres corrigés s'élève cette fois à  $p_c = q + q^2$ . Avec  $q = 10$  modes POD, le modèle d'ordre réduit calibré au moyen de la formulation de Floquet conduit aux erreurs suivantes :  $\overline{\varepsilon_{C_L}} = 1,52\%$ ,  $\varepsilon_{\Delta_{cc}} = 4,80\%$ ,  $\varepsilon_{\mathbf{A}} = 0,03\%$  et  $\varepsilon_{\dot{\mathbf{A}}} = 0,16\%$ . Les amplitudes modales ainsi que leurs dérivées obtenues avec le modèle d'ordre réduit sont donc excellentes et les grandeurs reconstruites au moyen de la décomposition POD sont très satisfaisantes. Toutefois, la question de la validité du modèle d'ordre réduit corrigé de cette manière se pose. En effet, les

coefficients sont modifiés de manière très importante, comme l'indiquent les graphiques de la figure 5.20. Un certain nombre de coefficients constants voient leur contribution nettement réduite et les coefficients linéaires — en particulier hors de la diagonale — subissent de fortes variations. Par conséquent le modèle d'ordre réduit corrigé s'éloigne de plus en plus du modèle d'ordre réduit initialement obtenu par projection de Galerkin. À cela s'ajoute le problème que la fonctionnelle dépend à présent de  $q + q^2$  paramètres et l'existence de minima locaux vers lesquels l'algorithme risque de se précipiter est très probable. Le problème est en outre accentué par le fait qu'aucune contrainte n'a été imposée : si les coefficients sont par exemple contraints à rester dans un intervalle de  $\pm 25\%$  autour de la valeur initiale du coefficient, l'amélioration obtenue est peu satisfaisante et la réponse du modèle d'ordre réduit n'est pas assez dissipative.



**Figure 5.20 :** Contribution relative des termes constants et linéaires calibrés avec la formulation de Floquet KL par rapport aux coefficients initiaux.

Les méthodes de calibrage mises en œuvre par Bourguet *et al.* [34], Couplet [50], Favier *et al.* [73] doivent donc être manipulées avec précaution puisque les modifications apportées aux coefficients deviennent très importantes. La capacité des algorithmes d'optimisation à déterminer un optimum correct est rapidement dépassée quand le nombre de paramètres augmente, et ce d'autant plus lorsque aucune borne n'est imposée pour contraindre le problème.

En conclusion de ce paragraphe, il s'avère que l'utilisation de la formulation non-linéaire de Floquet ou celle basée sur l'énergie  $\mathcal{E}_T^*$  fournit de bons résultats pour calibrer un nombre de Reynolds associé à chaque mode POD. Le problème d'optimisation est dans ce cas résolu avec l'algorithme Quasi-Newton avec la contrainte supplémentaire de n'introduire que de la dissipation.

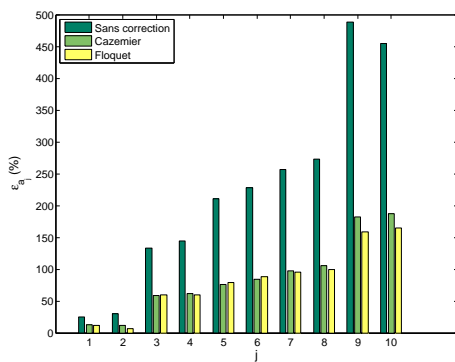
#### 5.5.4 Comparaison des réponses corrigées

Le choix d'une méthode de correction du modèle d'ordre réduit doit se faire en gardant à l'esprit que les équations sur lesquelles la méthode de projection est basée ne correspondent pas strictement à celles résolues pour générer les clichés. Dans le cas présent, l'hypothèse d'une viscosité du fluide constante peut expliquer les différences qui apparaissent dans la réponse du modèle d'ordre réduit non corrigé. Ainsi, la correction du nombre de Reynolds est un moyen cohérent de remédier aux défauts du modèle d'ordre réduit qui peuvent être attribués à l'hypothèse d'une viscosité constante. Par extension, l'optimisation d'un nombre de Reynolds pour chaque mode POD puis la correction des coefficients linéaires diagonaux s'inscrit dans la lignée d'une correction des termes dissipatifs. La méthode de Cazemier *et al.* [42] et celle basée sur le calibrage des nombres de Reynolds avec une formulation de Floquet ont donc permis d'améliorer nettement la solution tout en maîtrisant l'importance de la correction introduite. Les méthodes de calibrage des coefficients du modèle d'ordre réduit impliquant un nombre plus important de paramètres sont par

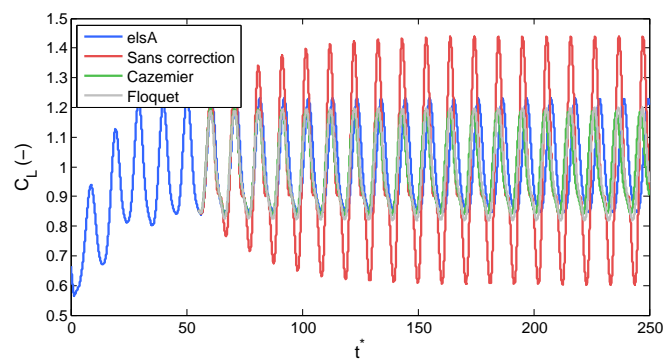
ailleurs capables de produire des réponses dont la qualité est encore meilleure, mais il est plus difficile de maîtriser les corrections apportées, tant au niveau de leur intensité que de leur interprétation.

Pour conclure ce paragraphe, la réponse du modèle d'ordre réduit POD-Galerkin corrigé au moyen de la méthode de [Cazemier et al.](#) et de celle de Floquet pour les paramètres  $\alpha_{Re}$  est comparée à la réponse du modèle d'ordre réduit non corrigé. Le graphique 5.21(a) indique que les deux méthodes produisent des amplitudes modales assez similaires et permettent de réduire considérablement les erreurs commises. Ceci se traduit immédiatement par une amélioration du coefficient de portance dont l'évolution à long terme est tracée sur le graphique de la figure 5.21(b). Après correction, le système réduit est donc en mesure de reproduire correctement l'amortissement du modèle initial.

Les corrections apportées se justifient par l'existence d'erreurs dont les origines ont été évoquées au paragraphe 5.3.1 p.130. La principale source d'erreur est certainement liée à la modélisation de la viscosité qui a été supposée constante alors que la réponse de référence issue du modèle non réduit tient compte de la loi de Sutherland et introduit une dissipation artificielle pour dissiper correctement la réponse. La correction apportée ici peut donc raisonnablement être interprétée comme un opérateur de dissipation artificielle.



(a) Erreur sur les amplitudes



(b) Coefficient de portance

**Figure 5.21 :** Comparaison de réponses du modèle d'ordre réduit corrigé au moyen de la méthode de [Cazemier et al.](#) et du calibrage  $\alpha_{Re}$  du nombre de Reynolds pour chaque mode POD.



## 5.6 Synthèse du chapitre 5

Dans ce chapitre, les différents éléments nécessaires à la construction du modèle d'ordre réduit d'un écoulement compressible instationnaire évoluant dans un domaine de frontières fixes ont été présentés. La formulation POD-Galerkin s'est révélée adéquate pour développer un modèle d'ordre réduit non-linéaire basé sur les équations de Navier-Stokes formulées au moyen des variables primitives modifiées. Une décomposition orthogonale propre multivariées a été pratiquée afin de tenir compte des corrélations entre les variables et a conduit à un modèle d'ordre réduit de dimension extrêmement faible. Le modèle d'ordre réduit résultant de la projection sur la base des modes POD est simplement composé d'un ensemble de  $q$  équations différentielles ordinaires quadratiques dont l'intégration en temps s'effectue très rapidement.

Le modèle d'ordre réduit ainsi développé a été mis à l'épreuve pour reproduire l'écoulement instationnaire autour d'un profil NACA0012 fortement incliné. L'objectif a alors consisté à reproduire le phénomène de détachement tourbillonnaire dans le sillage du profil qui se caractérise notamment par une oscillation périodique du coefficient de portance. Le calcul des modes POD a permis de retrouver une organisation des modes par paires qui est caractéristique du phénomène d'émission de vortex dans le sillage d'obstacles.

La réponse du modèle d'ordre réduit dont les coefficients ont été calculés au moyen des expressions analytiques établies au début de ce chapitre a révélé un manque significatif de dissipation. Plusieurs méthodes de correction ont alors été mises en œuvre avec plus ou moins de succès. En particulier, la méthode de correction de [Cazemier \*et al.\*](#) dans laquelle un terme d'amortissement calculé pour conserver l'énergie de chaque mode fournit une stabilisation efficace. Des techniques de correction a posteriori reposant sur la résolution d'un problème d'optimisation ont par ailleurs été employées avec succès. Parmi les variantes évaluées, la formulation de Floquet utilisant l'erreur sur les amplitudes modales ainsi que la formulation énergétique dans laquelle l'énergie captée par chaque mode est calibrée permettent de corriger efficacement le modèle d'ordre réduit. Afin de maîtriser la correction introduite, la méthode adoptée a consisté à optimiser un nombre de Reynolds pour dissiper correctement chaque mode POD, en imposant que la correction soit purement dissipative.

Les réponses obtenues avec le modèle d'ordre réduit corrigé sont donc satisfaisantes puisqu'elles permettent de reproduire avec une bonne précision l'amplitude des forces aérodynamiques. Cette première étape dans le développement d'un modèle d'ordre réduit aérodynamique est capitale et ouvre donc la voie au couplage avec le mouvement d'une structure dans le cadre d'un système aéroélastique.

# Construction et mise en œuvre d'un modèle d'ordre réduit POD-Galerkin non-linéaire et non-autonome pour l'aéroélasticité

## Sommaire

<b>6.1</b>	<b>Introduction</b>	<b>162</b>
<b>6.2</b>	<b>Simplification des équations de Navier-Stokes pour un mouvement de corps rigide</b>	<b>163</b>
6.2.1	Applicabilité de la POD en aéroélasticité	163
6.2.1.1	Obstacles à la définition de la POD sur un domaine mobile	163
6.2.1.2	État de l'art des modèles d'ordre réduit basés sur la POD en aéroélasticité	164
6.2.2	Simplification des équations de Navier-Stokes	167
<b>6.3</b>	<b>Construction du modèle d'ordre réduit POD-Galerkin pour un mouvement de corps rigide</b>	<b>169</b>
6.3.1	Développement d'un modèle d'ordre réduit basé sur la formulation en vitesse absolue	169
6.3.2	Développement d'un modèle d'ordre réduit basé sur la formulation en vitesse relative	175
6.3.3	Construction de modèles d'ordre réduit par identification des coefficients	178
6.3.4	Résumé et comparaison qualitative des différents modèles d'ordre réduit	180
<b>6.4</b>	<b>Calcul des modes POD de l'écoulement autour d'un profil NACA0064 oscillant</b>	<b>182</b>
6.4.1	Simulation numérique de l'écoulement transsonique autour du profil NACA0064	182
6.4.2	Simplification des coefficients du modèle d'ordre réduit pour les équations d'Euler	186
6.4.3	Calcul de la base des modes POD de l'écoulement autour du profil NACA0064	187
<b>6.5</b>	<b>Réponse du modèle d'ordre réduit de l'écoulement autour du profil NACA0064</b>	<b>191</b>
6.5.1	Réponse du modèle d'ordre réduit POD-Galerkin corrigé	191
6.5.1.1	Correction de <i>Cazemier et al.</i> du modèle d'ordre réduit POD-Galerkin	191
6.5.1.2	Calibrage des coefficients du modèle d'ordre réduit POD-Galerkin	193
6.5.2	Réponse du modèle d'ordre réduit POD-Galerkin identifié	198
<b>6.6</b>	<b>Synthèse du chapitre 6</b>	<b>204</b>

## 6.1 Introduction

Le travail effectué jusqu'à présent a permis de construire et de valider le modèle d'ordre réduit d'un écoulement dans un domaine de frontières fixes au moyen de la formulation FEUL-RAUA basée sur une description eulérienne des équations de Navier-Stokes dans le repère absolu. On s'intéresse à présent à la prise en compte d'un mouvement de corps rigide de la structure afin de progresser vers la construction d'un modèle d'ordre réduit d'aérodynamique instationnaire pour l'aéroélasticité. On justifie dans un premier temps le choix d'un mouvement de corps rigide avant de décrire deux formulations du modèle d'ordre réduit obtenues à partir des équations de Navier-Stokes écrites dans le repère entraîné avec les composantes de la vitesse absolue ou relative.

La réponse du modèle d'ordre réduit sera alors évaluée sur l'exemple d'un profil NACA0064 animé d'un mouvement d'oscillation harmonique placé dans un écoulement transsonique. Le fluide modélisé par les équations d'Euler exhibe alors un phénomène d'oscillation des chocs qui apparaissent sur la paroi du profil. Plusieurs techniques de construction et de correction du modèle d'ordre réduit seront alors mises en œuvre pour parvenir à reproduire correctement le phénomène.

## 6.2 Simplification des équations de Navier-Stokes pour un mouvement de corps rigide

Jusqu'à présent, les modes POD ont été calculés à partir d'un ensemble de clichés définis sur un domaine de l'espace  $\Omega$  invariant au cours du temps. Dans le cadre d'un problème aéroélastique cependant, la structure placée dans l'écoulement est en général supposée déformable et peut en outre être animée d'un mouvement de corps rigide. Par conséquent le domaine se déplace et/ou se déforme au cours du temps. Cela nécessite dans un premier temps de reformuler les équations du fluide dans le domaine arbitraire associé à la description ALE si le domaine se déforme et éventuellement dans le repère entraîné au mouvement de corps rigide si il en existe un. Ce travail a été effectué dans le chapitre 4 et fournit une description de l'écoulement dans de telles conditions. Les équations du fluide ont en particulier été écrites avec les variables primitives modifiées afin de conserver une forme quadratique des équations différentielles ordinaires qui soit propice à la projection de Galerkin.

La seconde partie du travail consiste à trouver une formulation du modèle d'ordre réduit qui s'adapte au mouvement et si possible à la déformation de la structure. Cependant, la signification des modes POD est remise en cause lorsque le domaine spatial est variable au cours du temps. L'objectif de cette section est d'éclaircir dans un premier temps la problématique de l'applicabilité de la décomposition orthogonale propre en aéroélasticité. Les différentes solutions adoptées par le passé seront alors recensées et la solution adoptée dans le cadre de cette thèse sera exposée pour terminer.

### 6.2.1 Applicabilité de la POD en aéroélasticité

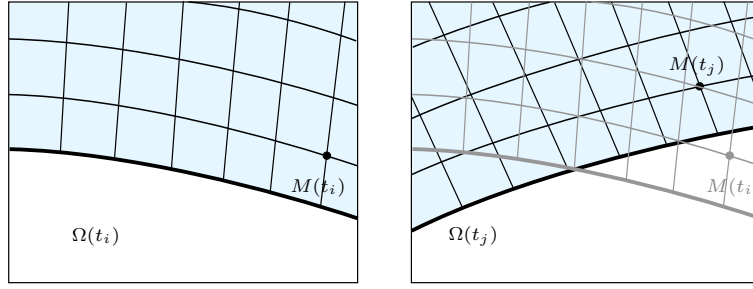
#### 6.2.1.1 Obstacles à la définition de la POD sur un domaine mobile

À l'origine employée comme un outil d'analyse de données, la décomposition orthogonale propre fournit une base optimale pour représenter l'ensemble des clichés, indépendamment de toute notion de variation du domaine de définition des données. En mécanique, les clichés  $\mathbf{q}^{(m)}$  solutions pour  $m = 1, \dots, M$  du système dynamique aux instants  $t_m$  doivent donc pouvoir être décomposés en une somme finie de modes POD spatiaux  $\varphi^{(j)}(\mathbf{x})$  pondérés par des amplitudes modales  $a_j^{(m)}$ . Par extension on suppose que la solution à tout instant  $t$  doit pouvoir se décomposer en  $\mathbf{q}(\mathbf{x}, t) = \bar{\mathbf{q}}(\mathbf{x}) + \sum_{j=1}^q a_j(t) \varphi^{(j)}(\mathbf{x})$  pour  $\mathbf{x} \in \Omega$ . Le principe repose donc sur une séparation espace-temps des solutions du système à l'aide des modes POD qui sont des structures spatiales *indépendantes du temps* capturant de façon optimale les caractéristiques du système.

Or, lorsque le domaine spatial varie au cours du temps, cette séparation des variables est remise en cause. En effet, à chaque instant  $t_m$ , les clichés sont définis sur un domaine  $\Omega(t_m)$  différent pour lequel la position des nœuds du maillage est modifiée, comme le montre le schéma de la figure 6.1. Par conséquent il est difficile de déterminer le domaine spatial de définition des modes POD, et la décomposition des modes POD sur les clichés introduite à l'équation (2.82) n'a plus aucun sens. En effet, lorsque la dépendance des variables en fonction du temps et de l'espace est précisée la relation s'écrit :  $\varphi^{(j)}(\mathbf{x}) = \sum_{m=1}^M c_m^{(j)} (\mathbf{q}(\mathbf{x}, t_m) - \bar{\mathbf{q}}(\mathbf{x}))$ . Ainsi se pose la question de savoir sur quel domaine les modes POD  $\varphi^{(j)}$  doivent être définis puisque chaque cliché  $\mathbf{q}(\mathbf{x}, t_m)$  est défini sur un domaine  $\Omega(t_m)$  différent. Leur domaine de définition devrait couvrir au moins la réunion des  $\Omega(t_m)$  de sorte que chaque cliché puisse être ensuite reconstruit sur son domaine de définition par la décomposition POD. Toutefois, dans l'optique de la construction d'un modèle d'ordre réduit, la réponse n'est pas connue a priori et par conséquent il faut supposer que les solutions restent comprises dans un certain domaine spatial sur lequel les modes doivent être définis.

Quand bien même cela serait possible, il reste à calculer les coefficients  $c_m^{(j)}$  qui pondèrent les clichés dans l'expression des modes POD (2.82). Or ces coefficients sont solutions du problème aux valeurs propres de la méthode des clichés dont le terme général  $R_{ij}^* = \sqrt{\alpha_i \alpha_j} \langle \mathbf{q}^{(i)} - \bar{\mathbf{q}}, \mathbf{q}^{(j)} - \bar{\mathbf{q}} \rangle_H$  requiert le calcul du produit scalaire entre des clichés centrés à deux instants différents. Jusqu'à présent, les clichés étaient supposés être contenus dans l'espace  $H = (L^2(\Omega))^{n_v}$  des fonctions de carré intégrable définies sur un domaine spatial  $\Omega$  auquel on avait adjoint le produit scalaire

$$\langle \mathbf{q}^{(i)}, \mathbf{q}^{(j)} \rangle = \sum_{k=1}^{n_v} \int_{\Omega} q_k(\mathbf{x} \in \Omega(t_i), t_i) q_k(\mathbf{x} \in \Omega(t_j), t_j) d\Omega. \quad (6.1)$$



**Figure 6.1 :** Schématisation du mouvement du domaine fluide au cours du temps. Le domaine de définition  $\Omega(t)$  du fluide (en bleu) change entre les instants  $t_i$  et  $t_j$  et le même nœud  $M$  change de position.

Cependant, les fonctions sont cette fois définies sur des domaines  $\Omega(t_i)$  et  $\Omega(t_j)$  différents si bien que le domaine d'intégration  $\Omega$  n'est pas clairement défini. L'ensemble des incohérences relevées découle de la mauvaise définition de l'espace de travail. En effet, pour formuler correctement la décomposition orthogonale propre dans le cas d'un domaine variable, il serait nécessaire de travailler sur l'espace de Hilbert  $H = (L^2(\Omega(t)))^{n_v}$  et de définir le produit scalaire associé en tenant compte du fait que le domaine de définition change au cours du temps. Il s'agit donc bien d'un problème intrinsèque à la formulation POD employée jusqu'à présent, et non pas seulement de l'utilisation de la méthode des clichés au lieu de la méthode directe<sup>1</sup>. La solution consiste donc à se ramener autant que possible à un domaine de référence fixe sur lequel les clichés et les modes POD sont définis. De cette façon, la même formulation POD que celle utilisée jusqu'à présent peut être conservée. Cependant, la difficulté apparaît lorsqu'il s'agit de formuler les équations gouvernant le système dynamique sur ce domaine de référence.

### 6.2.1.2 État de l'art des modèles d'ordre réduit basés sur la POD en aéroélasticité

Malgré la difficulté de formuler le problème de la décomposition orthogonale propre sur un domaine mobile, quelques travaux ont été menés pour des systèmes couplés fluide-structure. Tout d'abord, la décomposition orthogonale propre a été employée comme outil d'analyse, sans chercher à construire de modèle d'ordre réduit du système dynamique. Ainsi, Zhang *et al.* [252] ont étudié le comportement aérodynamique d'ailes volantes grâce à l'analyse de modes POD calculés à partir de clichés prélevés sur un domaine mobile. Pour cela, les clichés sont interpolés à chaque instant sur un domaine fixe  $\Omega_0$  au moyen d'une méthode d'interpolation basée sur l'inverse de la distance à un point. Un cliché  $q(\mathbf{x}_i, t_m)$  défini au centre de la cellule de contrôle  $\Omega_i$  sur le maillage mobile à un instant  $t_m$  est donc interpolé en un point  $\mathbf{x}_{i,0}$  sur le maillage fixe de référence au moyen de la relation

$$\tilde{q}(\mathbf{x}_{i,0}, t_m) = \sum_{k=1}^K \frac{q(\mathbf{x}_{i,k}, t_m)}{r_k} \bigg/ \sum_{k=1}^K \frac{1}{r_k}, \quad (6.2)$$

où  $q(\mathbf{x}_{i,k}, t_m)$  représente la valeur du champ aérodynamique à l'instant  $t_m$  au  $k$ -ième coin de la cellule  $\Omega_i$ , et  $r_k$  est la distance entre ce même point et le point d'interpolation  $\mathbf{x}_{i,0}$ .

D'une façon plus générale, Utturkar *et al.* [234] se sont intéressés à l'application de la POD en mécanique des fluides sur un domaine avec une frontière mobile au cours du temps, que ce soit pour des problèmes d'interaction fluide-structure comme ici, ou encore pour des problèmes de cavitation ou de changement de phase. Aucune modification à la formulation POD n'a été introduite pour l'exemple de la cavitation, ni pour le problème de changement de phase puisque dans ce cas le fluide est modélisé au moyen d'une fraction de phase liquide. L'exemple en interaction fluide-structure est celui traité dans [252] où les clichés sont interpolés sur un maillage de référence. Dans ces deux études, aucun modèle d'ordre réduit n'a été construit et les modes POD ont simplement été calculés en vue d'analyser les caractéristiques des systèmes dynamiques.

<sup>1</sup>Avec la méthode directe, il faut résoudre une équation intégrale de Fredholm qui est aussi mal définie puisqu'elle fait intervenir le tenseur de corrélation spatiale entre deux points  $M_i$  et  $M_j$  dont les positions changent au cours du temps.

Afin de reproduire les phénomènes physiques impliqués dans les différents exemples, il semble que la base POD doit être composée d'un nombre plus important de modes, en particulier pour les problèmes d'interaction fluide-structure. En outre le nombre de modes dépend de l'intervalle de temps simulé [234].

Très peu de travaux ont conduit à la construction de modèles d'ordre réduit pour l'interaction fluide-structure en raison de la difficulté provoquée par le mouvement du domaine. Pour remédier à ce problème, les premiers travaux étaient basés sur les équations linéarisées du fluide [189]. Dans ce cas, puisque les équations sont formulées pour de petites perturbations autour d'un état d'équilibre, on peut admettre que les effets du mouvement du domaine sont négligeables pour le calcul des modes POD, de sorte que la formulation POD établie dans le cas d'un domaine fixe reste encore valable. Ainsi, Hall *et al.* [104] ont construit un modèle d'ordre réduit POD des équations d'Euler linéarisées qui a ensuite été couplé à un modèle à deux degrés de liberté de la structure pour détecter les limites de stabilité d'un profil isolé en écoulement transsonique et d'une cascade de profils bidimensionnelle. Les équations d'Euler employées reposent sur une formulation eulérienne gouvernant l'évolution des fluctuations du champ aérodynamique sur une grille de calcul invariante au cours du temps. Une autre approche consiste à linéariser les équations d'Euler ou de Navier-Stokes écrites en formulation ALE, auquel cas le terme lié au mouvement de maillage doit aussi être linéarisé. C'est la solution adoptée notamment par Mortchéléwicz [170] et Lieu *et al.* [148]. Dans tous les cas, le modèle d'ordre réduit du domaine fluide peut être aisément couplé aux équations de la structure afin de construire un modèle d'ordre réduit du système couplé aéroélastique qui sert ensuite à obtenir la réponse à différents paramètres tels que l'amplitude ou la fréquence d'une perturbation imposée.

Anttonen *et al.* [12, 13] se sont intéressés de leur côté au calcul de modes POD et à la construction d'un modèle d'ordre réduit POD pour des mouvements de grande amplitude d'une frontière du domaine fluide. Dans un premier temps, l'évaluation de l'erreur de reconstruction des clichés à partir d'une base POD calculée avec les équations formulées sur un domaine fixe en repère entraîné et dans un domaine mobile en repère fixe démontre qu'il est nécessaire de conserver plus de modes POD dans la base de projection avec la formulation en repère fixe et domaine mobile. Dans les deux cas, les modes POD sont calculés avec le produit scalaire de  $\mathbb{R}^{N_v}$ , où  $N_v$  est le nombre de degrés de liberté du système — et ceci malgré le fait que dans le second cas les positions des nœuds changent à chaque instant. L'argument avancé est que la connectivité des nœuds n'est jamais modifiée et par conséquent les modes POD sont calculés dans un espace *indiciel*. Les structures mises en évidence par les modes POD sont alors associées à la connectivité du maillage et non pas aux positions spatiales. Toutefois, cette base POD calculée est inappropriée lorsqu'il s'agit de construire un modèle d'ordre réduit et de reproduire la réponse du système à différents paramètres comme l'amplitude du mouvement par exemple. En effet, en procédant de la sorte les informations sur la déformation de maillage sont totalement absentes et les modes POD ne sont pas localisés. La première solution développée — appelée *Blended POD* dans [12] — qui consiste à construire une base POD à partir de clichés calculés pour différentes fréquences et amplitudes de mouvement n'est donc pas satisfaisante puisque les informations sur la déformation du maillage sont absentes : le modèle d'ordre réduit obtenu est instable à toutes les fréquences et amplitudes. Par conséquent une seconde méthode — appelée *Multi-POD* dans [12] — est développée à partir de plusieurs modèles d'ordre réduit construits pour différentes amplitudes ; la sélection du modèle d'ordre réduit adéquat se fait en fonction d'un critère de déformation du maillage qui permet d'identifier le modèle d'ordre réduit qui a été construit pour l'amplitude de déformation la plus proche de la déformation actuellement calculée par le modèle d'ordre réduit. Pour déterminer le meilleur modèle d'ordre réduit, une base POD du mouvement de maillage est établie pour chaque modèle d'ordre réduit. L'erreur de projection du champ de déformation actuel sur cette base est alors calculée pour les différentes bases POD de maillage et l'erreur de projection minimale indique le modèle d'ordre réduit le plus approprié. La méthode nécessite donc de construire un certain nombre de modèles d'ordre réduit puis de passer de l'un à l'autre en fonction de la déformation de maillage calculée. Un exemple d'application à un profil en mouvement de tangage et de plongeon est traité dans [13] mais les résultats sont peu probants.

Finalement, une formulation élégante du problème couplé fluide-structure pour un écoulement incompressible a été proposée par Liberge et Hamdouni [145, 146]. Le principe consiste à poser le problème de façon monolithique, *i. e.* à employer les mêmes équations pour le fluide et la structure. La formulation généralisée des équations de Navier-Stokes développée par Glowinski *et al.* [88] a donc été reprise par Liberge pour formuler le problème sur le domaine  $\Omega = \Omega_f \cup \Omega_s$ , de sorte le domaine solide puisse être considéré comme un domaine fluide « fictif ». Le problème couplé s'écrit donc au moyen d'une formulation faible globale sur tout le domaine spatial  $\Omega$  qui apparaît comme invariant au cours du temps. L'objectif est alors



- la *Multi-POD* d'Anttonen *et al.* [12] nécessite la construction d'un ensemble de plusieurs modèles d'ordre réduit entre lesquels on alterne en fonction de la déformation du maillage. Or il serait préférable de disposer d'une formulation tenant compte intrinsèquement de la mobilité du domaine, et de réserver les changements — ou mieux les interpolations — de modèles d'ordre réduit aux variations de paramètres tels que le nombre de Mach ou l'angle d'incidence par exemple ;
- la formulation de Liberge [145] basée sur les équations généralisées de Navier-Stokes est attractive puisque le cadre théorique est bien posé et permet de traiter les non-linéarités de l'écoulement. La seule limitation est que la formulation s'applique pour l'instant à un solide rigide et à un écoulement incompressible.

Au vu des descriptions des différentes approches, il semblerait que la meilleure solution soit de travailler sur un domaine spatial fixe, à moins de trouver une définition pertinente du produit scalaire sur un domaine mobile. Le choix d'un domaine invariant n'est simple qu'en apparence, car si il est facile d'interpoler les clichés sur un domaine de référence sur lequel les modes POD seront calculés, la difficulté est reportée lorsqu'il s'agit d'adapter les équations du fluide à ce domaine de référence. En effet, si les clichés peuvent être interpolés sur un domaine de référence au moyen de la relation (6.2) par exemple, alors les modes POD  $\tilde{\varphi}^{(j)}$  sont déterminés sur ce même domaine par l'expression (2.82). Toutefois, on obtient la décomposition POD des *variables interpolées* sur le domaine  $\Omega_0$

$$\tilde{q}(\mathbf{x}_0, t) = \tilde{q}(\mathbf{x}_0) + \sum_{j=1}^q a_j(t) \tilde{\varphi}^{(j)}(\mathbf{x}_0), \quad (6.8)$$

et non une décomposition des variables sur le domaine variable  $\Omega(t)$ . Or, les équations du fluide formulées au moyen de la description ALE par exemple gouvernent les variables aérodynamiques  $q(\mathbf{x}, t)$  et non les variables interpolées  $\tilde{q}(\mathbf{x}_0, t)$ . Ainsi, la décomposition POD (6.8) n'est pas d'un grand secours pour la construction d'un modèle d'ordre réduit puisqu'elle ne peut être employée pour remplacer les variables aérodynamiques dans les équations à projeter.

La solution adoptée dans le cadre de cette thèse va donc consister à supposer un mouvement de corps rigide puis à formuler les équations dans le repère entraîné associé au mouvement. De cette manière, le domaine spatial reste invariant et les équations peuvent être écrites dans le repère entraîné, au prix d'un certain nombre de calculs détaillés dans la section suivante.

## 6.2.2 Simplification des équations de Navier-Stokes dans le cas d'un mouvement de corps rigide de la structure

Les équations locales de Navier-Stokes ont été formulées au paragraphe 4.4.3 à l'aide des variables primitives modifiées pour obtenir une forme quadratique propice à la projection de Galerkin. La description ALE adoptée a ainsi permis de tenir compte de l'effet d'une structure déformable placée dans l'écoulement et d'un éventuel mouvement de corps rigide. Dans ce dernier cas, le choix des composantes de la vitesse absolue ou relative a conduit aux formulations FALÉ-REUE et FALÉ-REVE fournies respectivement aux équations (4.62) et (4.68) avec les variables primitives modifiées.

L'hypothèse d'un mouvement de corps rigide permet alors de simplifier les formulations précédentes et de se ramener finalement à une formulation dans laquelle des termes sources supplémentaires sont introduits pour tenir compte du mouvement du domaine qui est alors indissociable de celui du repère entraîné. L'expression de la vitesse des nœuds du maillage (4.30) est alors simplifiée puisque les positions relatives  $\tilde{\mathbf{x}}_E$  sont invariantes dans le repère entraîné et la vitesse de déformation des points du maillage  $\mathbf{s}_{d,E}$  s'annule. En outre, la vitesse d'entraînement correspond à la vitesse du mouvement de corps rigide, *i. e.* au mouvement imposé à la structure dans le cas d'un couplage faible. Sous l'hypothèse d'un mouvement de corps rigide les vitesses d'entraînement et de déformation sont données par

$$\boxed{\begin{aligned} \mathbf{s}_{d,E} &= \mathbf{0} \\ \mathbf{s}_E &= \mathbf{s}_{e,E} = \mathbf{s}_{0,E} + \boldsymbol{\omega}_E \wedge \tilde{\mathbf{x}}_E \end{aligned}}. \quad (6.9)$$

Les relations précédentes permettent alors de simplifier les expressions des équations de Navier-Stokes formulées dans le repère entraîné. En particulier les termes liés à la vitesse du maillage sont explicitement



connus dès lors que la vitesse de translation  $\mathbf{s}_{0,E}$  et la vitesse de rotation  $\boldsymbol{\omega}_E$  sont connus. Les équations (4.62) formulées avec les composantes de la vitesse absolue exprimée dans le repère entraîné sont alors rendues adimensionnelles selon le même procédé qu'au paragraphe 5.2.1. La vitesse  $\mathbf{s}_E$  est ensuite explicitée avec les relations (6.9) sous l'hypothèse d'un mouvement de corps rigide, ce qui conduit à

$$\left\{ \begin{array}{l} \frac{\partial \vartheta}{\partial t} \Big|_{\mathcal{X}} + \mathbf{u}_E \cdot \nabla_E \vartheta = \vartheta \operatorname{div}_E \mathbf{u}_E + \mathbf{s}_{e,E} \cdot \nabla_E \vartheta \\ \frac{\partial \mathbf{u}_E}{\partial t} \Big|_{\mathcal{X}} + \mathbf{u}_E \cdot \nabla_E \mathbf{u}_E = -\vartheta \nabla_E p + \frac{\vartheta}{\operatorname{Re}} \operatorname{div}_E \boldsymbol{\tau}(\mathbf{u}_E) - \boldsymbol{\omega}_E \wedge \mathbf{u}_E + \mathbf{s}_{e,E} \cdot \nabla_E \mathbf{u}_E \\ \frac{\partial p}{\partial t} \Big|_{\mathcal{X}} + \mathbf{u}_E \cdot \nabla_E p = -\gamma p \operatorname{div}_E \mathbf{u}_E + \frac{\gamma-1}{\operatorname{Re}} \nabla_E \mathbf{u}_E : \boldsymbol{\tau}(\mathbf{u}_E) + \frac{\gamma}{\operatorname{Re} \operatorname{Pr}} \Delta_E(p \vartheta) + \mathbf{s}_{e,E} \cdot \nabla_E p \end{array} \right. \quad (6.10)$$

La formulation en vitesse relative s'obtient de manière analogue à partir de l'équation (4.68) :

$$\left\{ \begin{array}{l} \frac{\partial \vartheta}{\partial t} \Big|_{\mathcal{X}} + \mathbf{v}_E \cdot \nabla_E \vartheta = \vartheta \operatorname{div}_E \mathbf{v}_E \\ \frac{\partial \mathbf{v}_E}{\partial t} \Big|_{\mathcal{X}} + \mathbf{v}_E \cdot \nabla_E \mathbf{v}_E = -\vartheta \nabla_E p + \frac{\vartheta}{\operatorname{Re}} \operatorname{div}_E \boldsymbol{\tau}(\mathbf{v}_E) + \mathbf{f}_{\text{Cor}} + \mathbf{f}_{\text{cen}} \\ \frac{\partial p}{\partial t} \Big|_{\mathcal{X}} + \mathbf{v}_E \cdot \nabla_E p = -\gamma p \operatorname{div}_E \mathbf{v}_E + \frac{\gamma-1}{\operatorname{Re}} \nabla_E \mathbf{v}_E : \boldsymbol{\tau}(\mathbf{v}_E) + \frac{\gamma}{\operatorname{Re} \operatorname{Pr}} \Delta_E(p \vartheta) \end{array} \right. \quad (6.11)$$

Dans ce cas, les équations sont similaires à celles d'une formulation eulérienne, à la différence près que le bilan de quantité de mouvement fait intervenir des termes gyroscopiques. Ceci provient du fait que le domaine reste rigide et peut être considéré comme un domaine invariant puisque les équations sont formulées dans le repère associé au mouvement.

### 6.3 Construction du modèle d'ordre réduit POD-Galerkin pour un mouvement de corps rigide

Le problème du mouvement d'une structure mobile placée dans le fluide a été résolu en faisant l'hypothèse d'un mouvement de corps rigide et en reformulant les équations dans le repère associé à ce mouvement. Il en résulte une modification des équations de Navier-Stokes par l'introduction de termes sources reflétant le mouvement du domaine. Les modes POD peuvent être calculés de façon cohérente puisque le domaine spatial est invariant lorsqu'on se place dans le repère entraîné. Deux formulations du modèle d'ordre réduit sont alors développées ici suivant que les composantes de la vitesse absolue ou relative sont employées.

#### 6.3.1 Développement d'un modèle d'ordre réduit POD-Galerkin basé sur la formulation en vitesse absolue

Ce premier modèle d'ordre réduit est construit par projection de Galerkin des équations continues (6.10) en suivant le même principe que dans le paragraphe 5.2.1. Cependant, le système dynamique devient non-autonome en raison de la présence des termes sources et s'écrit donc

$$\dot{\mathbf{q}} = \mathbf{Q}^C(\mathbf{q}, \mathbf{q}) + \frac{1}{\text{Re}} \mathbf{Q}^D(\mathbf{q}, \mathbf{q}) + \mathbf{T}(\mathbf{q}, \mathbf{s}_{0,E}, \boldsymbol{\omega}_E), \quad (6.12)$$

où les opérateurs introduits sont donnés par

$$\mathbf{Q}^C(\mathbf{q}, \mathbf{q}) = \begin{bmatrix} -\mathbf{u}_E \cdot \nabla_E \vartheta + \vartheta \text{div}_E \mathbf{u}_E \\ -\mathbf{u}_E \cdot \nabla_E \mathbf{u} - \vartheta \nabla_E p \\ -\mathbf{u}_E \cdot \nabla_E p - \gamma p \text{div}_E \mathbf{u}_E \end{bmatrix}, \quad \mathbf{T}(\mathbf{q}, \mathbf{s}_{0,E}, \boldsymbol{\omega}_E) = \begin{bmatrix} \mathbf{s}_{e,E} \cdot \nabla_E \vartheta \\ \mathbf{s}_{e,E} \cdot \nabla_E \mathbf{u}_E - \boldsymbol{\omega}_E \wedge \mathbf{u}_E \\ \mathbf{s}_{e,E} \cdot \nabla_E p \end{bmatrix}, \quad (6.13)$$

$$\mathbf{Q}^D(\mathbf{q}, \mathbf{q}) = \begin{bmatrix} 0 \\ \vartheta \text{div}_E \boldsymbol{\tau}(\mathbf{u}_E) \\ (\gamma - 1) \nabla_E \mathbf{u}_E : \boldsymbol{\tau}(\mathbf{u}_E) + \frac{\gamma}{\text{Pr}} \Delta_E(\vartheta p) \end{bmatrix}.$$

Les opérateurs  $\mathbf{Q}^C$  et  $\mathbf{Q}^D$  ont exactement la même forme que ceux de l'équation (5.6) mais ils s'appliquent cette fois à la vitesse absolue  $\mathbf{u}_E$  exprimée dans le repère entraîné au lieu de la vitesse  $\mathbf{u}_A$  exprimée dans le repère absolu. Enfin, le terme source supplémentaire  $\mathbf{T}$  est linéaire en fonction des variables aérodynamiques  $\mathbf{q}$ , mais il dépend aussi du mouvement imposé qui résulte d'une translation et d'une rotation caractérisées respectivement par  $\mathbf{s}_{0,E}$  et  $\boldsymbol{\omega}_E$ . Le terme source ne dépend que de ces deux degrés de liberté de mouvement — ce qu'on note  $\mathbf{T}(\mathbf{q}, \mathbf{s}_{0,E}, \boldsymbol{\omega}_E)$  — puisque la vitesse du maillage  $\mathbf{s}_{e,E}$  est donnée par l'expression (6.9).

Le modèle d'ordre réduit est construit comme dans le paragraphe 5.2.1 : les variables aérodynamiques  $\mathbf{q} = [\vartheta, \mathbf{u}_E, p]^T$  sont remplacées par leur décomposition POD (2.83) qui est introduite dans le système d'équations (6.12), puis le tout est projeté sur chaque mode POD. Cette opération conduit au système dynamique réduit

$$\dot{a}_i = \mathcal{K}_i + \sum_{j=1}^q \mathcal{L}_{ij} a_j + \sum_{j,k=1}^q \mathcal{Q}_{jik} a_j a_k + \mathcal{K}_i^E(t) + \sum_{j=1}^q \mathcal{L}_{ij}^E(t) a_j, \quad (6.14)$$

dans lequel les amplitudes modales ont été extraites des opérateurs en recourant à la propriété de multilinearité de  $\mathbf{Q}^C$ ,  $\mathbf{Q}^D$  et  $\mathbf{T}$ . La nouvelle expression du modèle d'ordre réduit introduit deux types de termes : d'une part les coefficients *autonomes*  $\mathcal{K}_i$ ,  $\mathcal{L}_{ij}$  et  $\mathcal{Q}_{jik}$ , et d'autre part les coefficients *non-autonomes* (ou d'entraînement)  $\mathcal{K}_i^E(t)$  et  $\mathcal{L}_{ij}^E(t)$  qui dépendent implicitement du temps et rendent donc le système d'équations différentielles ordinaires non-autonome. Les coefficients autonomes sont identiques à ceux du modèle d'ordre réduit construit à partir des équations en formulation eulérienne dans le repère absolu, bien que les modes POD employés à présent aient été obtenus à partir des vitesses absolues exprimées dans le repère entraîné. Leurs expressions sont données par les équations (5.9). La projection du terme source linéaire

induit quant à elle un ensemble de coefficients constants et linéaires supplémentaires en raison de la décomposition affine des variables sur les modes POD. On obtient donc les termes

$$\begin{aligned} \left\langle \mathbf{T}(\mathbf{q}, \mathbf{s}_{0,E}, \boldsymbol{\omega}_E), \boldsymbol{\varphi}^{(i)} \right\rangle &= \left\langle \mathbf{T} \left( \bar{\mathbf{q}} + \sum_{j=1}^q a_j \boldsymbol{\varphi}^{(j)}, \mathbf{s}_{0,E}, \boldsymbol{\omega}_E \right), \boldsymbol{\varphi}^{(i)} \right\rangle \\ &= \left\langle \mathbf{T}(\bar{\mathbf{q}}, \mathbf{s}_{0,E}, \boldsymbol{\omega}_E), \boldsymbol{\varphi}^{(i)} \right\rangle + \sum_{j=1}^q \left\langle \mathbf{T}(\boldsymbol{\varphi}^{(j)}, \mathbf{s}_{0,E}, \boldsymbol{\omega}_E), \boldsymbol{\varphi}^{(i)} \right\rangle a_j, \end{aligned} \quad (6.15)$$

à partir desquels sont introduits les coefficients non-autonomes constants  $\mathcal{K}_i^E$  et linéaires  $\mathcal{L}_{ij}^E$  issus du mouvement du domaine et définis par

$$\mathcal{K}_i^E(t) = \left\langle \mathbf{T}(\bar{\mathbf{q}}, \mathbf{s}_{0,E}, \boldsymbol{\omega}_E), \boldsymbol{\varphi}^{(i)} \right\rangle \quad \text{et} \quad \mathcal{L}_{ij}^E(t) = \left\langle \mathbf{T}(\boldsymbol{\varphi}^{(j)}, \mathbf{s}_{0,E}, \boldsymbol{\omega}_E), \boldsymbol{\varphi}^{(i)} \right\rangle. \quad (6.16)$$

En développant le produit scalaire, les coefficients linéaires s'écrivent

$$\begin{aligned} \mathcal{L}_{ij}^E(t) &= \left\langle T_{\vartheta}(\boldsymbol{\varphi}^{(j)}), \boldsymbol{\varphi}_{\vartheta}^{(i)} \right\rangle + \left\langle \mathbf{T}_{\mathbf{u}_E}(\boldsymbol{\varphi}^{(j)}), \boldsymbol{\varphi}_{\mathbf{u}_E}^{(i)} \right\rangle + \left\langle T_p(\boldsymbol{\varphi}^{(j)}), \boldsymbol{\varphi}_p^{(i)} \right\rangle \\ &= \int_{\Omega} \left( \mathbf{s}_{e,E} \cdot \nabla_E \boldsymbol{\varphi}_{\vartheta}^{(j)} \right) \boldsymbol{\varphi}_{\vartheta}^{(i)} \, d\Omega \\ &\quad + \int_{\Omega} \left( \mathbf{s}_{e,E} \cdot \nabla_E \boldsymbol{\varphi}_{\mathbf{u}_E}^{(j)} - \boldsymbol{\omega}_E \wedge \boldsymbol{\varphi}_{\mathbf{u}_E}^{(j)} \right) \cdot \boldsymbol{\varphi}_{\mathbf{u}_E}^{(i)} \, d\Omega \\ &\quad + \int_{\Omega} \left( \mathbf{s}_{e,E} \cdot \nabla_E \boldsymbol{\varphi}_p^{(j)} \right) \boldsymbol{\varphi}_p^{(i)} \, d\Omega, \end{aligned} \quad (6.17)$$

et les coefficients constants se déduisent simplement de l'expression précédente en remplaçant le mode POD  $\boldsymbol{\varphi}^{(j)}$  par la partie moyenne  $\bar{\mathbf{q}}$ .

La formulation précédente du modèle d'ordre réduit n'est pas adéquate puisque les coefficients non-autonomes  $\mathcal{K}_i^E(t)$  et  $\mathcal{L}_{ij}^E(t)$  dépendent implicitement du temps comme en témoigne l'expression (6.17). Les coefficients non-autonomes doivent être calculés à chaque pas de temps d'intégration du modèle d'ordre réduit puisque les termes instationnaires entrent dans le calcul des intégrales. Par conséquent le calcul de ces coefficients devient rapidement trop coûteux lorsque le nombre de cellules de contrôle dans le domaine de calcul augmente et l'intérêt de la réduction de modèle en pâtit.

Il est donc essentiel de scinder les expressions des termes  $\mathcal{K}_i^E(t)$  et  $\mathcal{L}_{ij}^E(t)$  afin d'obtenir des coefficients scalaires indépendants du temps qui soient pondérés par les degrés de liberté instationnaires du mouvement  $\mathbf{s}_{0,E}$  et  $\boldsymbol{\omega}_E$ . Ainsi, les coefficients peuvent être calculés une fois pour toute avant l'intégration en temps du modèle d'ordre réduit. Le travail est effectué pour les coefficients linéaires, puisque les coefficients constants se déduisent simplement en remplaçant les modes POD  $\boldsymbol{\varphi}^{(j)}$  par la partie moyenne des clichés  $\bar{\mathbf{q}}$ .

La modification des expressions des coefficients non-autonomes repose sur l'analyse de la vitesse de déplacement des nœuds du maillage donnée par l'expression (6.9). En effet, cette vitesse est la somme de la vitesse de translation  $\mathbf{s}_{0,E}$  de l'origine  $M_0$  du repère entraîné et de la vitesse de rotation calculée comme le produit vectoriel de la vitesse angulaire  $\boldsymbol{\omega}_E$  avec la position relative  $\tilde{\mathbf{x}}_E$ . Parmi ces trois termes, seule la position relative  $\tilde{\mathbf{x}}_E$  dépend de la position des points du maillage et ne peut par conséquent être extraite des intégrales impliquées dans l'expression (6.17) des coefficients. En effet, la vitesse de translation  $\mathbf{s}_{0,E}$  ne dépend que de la position de l'origine  $M_0$  du repère entraîné et le vecteur vitesse angulaire ne dépend que de la vitesse de rotation imposée au repère entraîné et qui est indépendante de la position des points du domaine. Si on précise explicitement la dépendance des différentes quantités, la vitesse de déplacement d'un nœud  $M$  du maillage s'écrit

$$\mathbf{s}_{e,E}(M, t) = \mathbf{s}_{0,E}(M_0, t) + \boldsymbol{\omega}_E(t) \wedge \tilde{\mathbf{x}}_E(M). \quad (6.18)$$

Le point capital est que la position relative  $\tilde{\mathbf{x}}_E$  ne dépend que de la position du point  $M$  et pas du temps. Ceci est dû à l'hypothèse d'un mouvement de corps rigide sans laquelle il ne serait pas possible d'effectuer les modifications qui suivent.

L'idée est donc d'extraire des intégrales de l'équation (6.17) les quantités  $s_{0,E}$  et  $\omega_E$  qui dépendent du temps. On considère pour cela un scalaire  $s(M, t) = \boldsymbol{\eta}(t) \cdot \boldsymbol{\mu}(M)$  fonction de la position d'un point  $M$  et du temps, défini comme le produit scalaire d'un vecteur  $\boldsymbol{\eta}(t)$  de fonctions dépendant uniquement du temps et d'un vecteur  $\boldsymbol{\mu}(M)$  de fonctions dépendant uniquement de l'espace. Les termes scalaires sous les intégrales de l'équation (6.17) sont alors interprétés comme un produit scalaire  $s(M, t) = \boldsymbol{\eta}(t) \cdot \boldsymbol{\mu}(M)$ , de sorte que l'intégrale devienne

$$\int_{\Omega} s(M, t) \, d\Omega = \int_{\Omega} \boldsymbol{\eta}(t) \cdot \boldsymbol{\mu}(M) \, d\Omega = \boldsymbol{\eta}(t) \cdot \int_{\Omega} \boldsymbol{\mu}(M) \, d\Omega. \quad (6.19)$$

L'intégrale du vecteur de fonctions  $\boldsymbol{\mu}(M) = [\mu_1(M), \mu_2(M), \mu_3(M)]^T$  est simplement définie comme le vecteur des intégrales de chacune des composantes sur le domaine  $\Omega$ . D'un point de vue pratique, si le domaine  $\Omega$  est discrétisé en un ensemble de  $N$  cellules de contrôle dans lesquelles on connaît les valeurs  $s(e)$ ,  $\eta_i(e)$  et  $\mu_i(e)$  du scalaire  $s$  et des composantes de  $\boldsymbol{\eta}$  et  $\boldsymbol{\mu}$  au centre des cellules, alors l'intégrale du scalaire  $s$  est approchée de manière discrète au moyen de la règle du point milieu par

$$\int_{\Omega} s(M, t) \, d\Omega \approx \sum_{e=1}^N s(e) \delta\Omega_e = \sum_{e=1}^N \sum_{i=1}^3 \eta_i(e) \mu_i(e) \delta\Omega_e. \quad (6.20)$$

Par ailleurs, la discrétisation de la relation (6.19) conduit à

$$\boldsymbol{\eta}(t) \cdot \int_{\Omega} \boldsymbol{\mu}(M) \, d\Omega \approx \sum_{i=1}^3 \eta_i \sum_{e=1}^N \mu_i(e) \delta\Omega_e. \quad (6.21)$$

Les expressions (6.21) et (6.20) sont en réalité identiques puisque le vecteur  $\boldsymbol{\eta}$  est indépendant de la position des points  $M$ ; par conséquent, pour tout élément du domaine discrétisé, la valeur de  $\eta_i(e)$  est identique quelque soit  $e$ . Il suffit donc de modifier les expressions des termes sous les intégrales de l'équation (6.17) afin de faire apparaître un produit scalaire de type  $s(M, t) = \boldsymbol{\eta}(t) \cdot \boldsymbol{\mu}(M)$  à partir duquel les contributions instationnaires puissent être isolées puis extraites de l'intégrale.

Pour y parvenir, on établit d'abord les quatre relations suivantes pour un scalaire  $s$  défini par des produits d'opérateurs intervenant dans l'équation (6.17) et portant sur des scalaires  $a, b$  ou des vecteurs  $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}$ :

$$s = (\boldsymbol{\alpha} \cdot \nabla a) b = \boldsymbol{\alpha} \cdot [(\nabla a) b], \quad (6.22a)$$

$$s = [(\boldsymbol{\alpha} \wedge \boldsymbol{\beta}) \cdot \nabla a] b = \boldsymbol{\alpha} \cdot [\boldsymbol{\beta} \wedge ((\nabla a) b)], \quad (6.22b)$$

$$s = (\boldsymbol{\alpha} \cdot \nabla \boldsymbol{\beta}) \cdot \boldsymbol{\gamma} = \boldsymbol{\alpha} \cdot [\boldsymbol{\gamma}^T (\nabla \boldsymbol{\beta})]^T, \quad (6.22c)$$

$$s = [(\boldsymbol{\alpha} \wedge \boldsymbol{\beta}) \cdot \nabla \boldsymbol{\gamma}] \cdot \boldsymbol{\delta} = \boldsymbol{\alpha} \cdot \left[ \boldsymbol{\beta} \wedge \left( \boldsymbol{\delta}^T (\nabla \boldsymbol{\gamma}) \right)^T \right]. \quad (6.22d)$$

À chaque fois, le vecteur  $\boldsymbol{\alpha}$  représente le vecteur vitesse de translation  $s_{0,E}$  ou de rotation  $\omega_E$ . Ainsi, on parvient à une expression sous forme d'un produit scalaire de  $\boldsymbol{\alpha}$  avec un autre vecteur, ce qui permettra d'extraire de l'intégrale  $\boldsymbol{\alpha}$  avec la relation (6.19).

Les relations (6.22) se démontrent facilement. Pour la première relation (6.22a) il suffit de remarquer que  $(\boldsymbol{\alpha} \cdot \nabla) a = \boldsymbol{\alpha} \cdot (\nabla a)$  puisque les deux relations donnent  $\alpha_i a_{,i}$ . Pour la seconde relation, on a  $[(\boldsymbol{\alpha} \wedge \boldsymbol{\beta}) \cdot \nabla a] b = (\boldsymbol{\alpha} \wedge \boldsymbol{\beta}) \cdot ((\nabla a) b)$  avec le même argument que précédemment. On note alors  $\mathbf{B} = (\nabla a) b$ , de sorte que la deuxième relation s'écrit  $(\boldsymbol{\alpha} \wedge \boldsymbol{\beta}) \cdot \mathbf{B}$  qui est un produit mixte. Puisque le produit mixte est invariant par permutation circulaire, on peut écrire la deuxième relation  $\boldsymbol{\alpha} \cdot (\boldsymbol{\beta} \wedge \mathbf{B})$  ce qui donne bien l'égalité (6.22b). La troisième relation s'écrit en notation indicelle  $\alpha_j \beta_{i,j} \gamma_i$ ; en permutant  $\alpha_j$  et  $\gamma_i$  on peut donc écrire le premier produit  $\boldsymbol{\gamma}^T \nabla \boldsymbol{\beta}$  puis il suffit de prendre la transposée de ce vecteur pour obtenir le produit scalaire avec  $\boldsymbol{\alpha}$ , ce qui fournit bien la relation (6.22c). Enfin pour obtenir la dernière relation, on applique la relation (6.22c) en prenant  $\boldsymbol{\alpha} = (\boldsymbol{\alpha} \wedge \boldsymbol{\beta}) = \mathbf{A}$ ,  $\boldsymbol{\beta} = \boldsymbol{\gamma}$  et  $\boldsymbol{\gamma} = \boldsymbol{\delta}$  ce qui donne :  $(\mathbf{A} \cdot \nabla \boldsymbol{\gamma}) \cdot \boldsymbol{\delta} = \mathbf{A} \cdot [\boldsymbol{\delta}^T (\nabla \boldsymbol{\gamma})^T]$ . On note ensuite  $\mathbf{B} = \boldsymbol{\delta}^T (\nabla \boldsymbol{\gamma})^T$  et on développe  $\mathbf{A}$  ce qui donne l'expression  $(\boldsymbol{\alpha} \wedge \boldsymbol{\beta}) \cdot \mathbf{B}$  qui est un produit mixte. Par permutation circulaire, on a finalement  $(\boldsymbol{\alpha} \wedge \boldsymbol{\beta}) \cdot \mathbf{B} = \boldsymbol{\alpha} \cdot (\boldsymbol{\beta} \wedge \mathbf{B})$  d'où la relation (6.22d).

Les trois contributions de l'équation (6.17) sont d'abord détaillées au moyen de la relation (6.18) pour la vitesse du maillage. La première contribution associée au covolume s'écrit

$$\begin{aligned} \langle T_\vartheta(\varphi^{(j)}), \varphi_\vartheta^{(i)} \rangle &= \int_\Omega (\mathbf{s}_{0,E} \cdot \nabla_E \varphi_\vartheta^{(j)}) \varphi_\vartheta^{(i)} d\Omega + \int_\Omega [(\boldsymbol{\omega}_E \wedge \tilde{\mathbf{x}}_E) \cdot \nabla_E \varphi_\vartheta^{(j)}] \varphi_\vartheta^{(i)} d\Omega \\ &= \mathbf{s}_{0,E} \cdot \int_\Omega \varphi_\vartheta^{(i)} \nabla_E \varphi_\vartheta^{(j)} d\Omega + \boldsymbol{\omega}_E \cdot \int_\Omega [\tilde{\mathbf{x}}_E \wedge (\nabla_E \varphi_\vartheta^{(j)})] \varphi_\vartheta^{(i)} d\Omega. \end{aligned} \quad (6.23)$$

La dernière égalité est obtenue au moyen des relations (6.22a) et (6.22b) établies ci-dessus et de la relation (6.19) pour extraire de l'intégrale les degrés de liberté du mouvement  $\mathbf{s}_{0,E}$  et  $\boldsymbol{\omega}_E$ . Ainsi, les intégrales peuvent être évaluées une fois pour toutes avant l'intégration en temps du modèle d'ordre réduit. De la même manière, la contribution en pression aux coefficients non-autonomes est

$$\begin{aligned} \langle T_p(\varphi^{(j)}), \varphi_p^{(i)} \rangle &= \int_\Omega (\mathbf{s}_{0,E} \cdot \nabla_E \varphi_p^{(j)}) \varphi_p^{(i)} d\Omega + \int_\Omega [(\boldsymbol{\omega}_E \wedge \tilde{\mathbf{x}}_E) \cdot \nabla_E \varphi_p^{(j)}] \varphi_p^{(i)} d\Omega \\ &= \mathbf{s}_{0,E} \cdot \int_\Omega \varphi_p^{(i)} \nabla_E \varphi_p^{(j)} d\Omega + \boldsymbol{\omega}_E \cdot \int_\Omega [\tilde{\mathbf{x}}_E \wedge (\nabla_E \varphi_p^{(j)})] \varphi_p^{(i)} d\Omega. \end{aligned} \quad (6.24)$$

Enfin, la contribution liée à la vitesse absolue exprimée dans le repère entraîné est d'abord transformée au moyen des relations (6.22c), (6.22d) et (6.19). Ensuite, la permutation circulaire du produit mixte  $(\boldsymbol{\omega}_E \wedge \boldsymbol{\varphi}_{\mathbf{u}_E}^{(j)}) \cdot \boldsymbol{\varphi}_{\mathbf{u}_E}^{(i)} = \boldsymbol{\omega}_E \cdot (\boldsymbol{\varphi}_{\mathbf{u}_E}^{(j)} \wedge \boldsymbol{\varphi}_{\mathbf{u}_E}^{(i)})$  conduit à

$$\begin{aligned} \langle T_{\mathbf{u}_E}(\varphi^{(j)}), \boldsymbol{\varphi}_{\mathbf{u}_E}^{(i)} \rangle &= \int_\Omega (\mathbf{s}_{0,E} \cdot \nabla_E \boldsymbol{\varphi}_{\mathbf{u}_E}^{(j)}) \cdot \boldsymbol{\varphi}_{\mathbf{u}_E}^{(i)} d\Omega + \int_\Omega [(\boldsymbol{\omega}_E \wedge \tilde{\mathbf{x}}_E) \cdot \nabla_E \boldsymbol{\varphi}_{\mathbf{u}_E}^{(j)}] \cdot \boldsymbol{\varphi}_{\mathbf{u}_E}^{(i)} d\Omega \\ &\quad - \int_\Omega (\boldsymbol{\omega}_E \wedge \boldsymbol{\varphi}_{\mathbf{u}_E}^{(j)}) \cdot \boldsymbol{\varphi}_{\mathbf{u}_E}^{(i)} d\Omega \\ &= \mathbf{s}_{0,E} \cdot \int_\Omega [(\boldsymbol{\varphi}_{\mathbf{u}_E}^{(i)})^\top \nabla_E \boldsymbol{\varphi}_{\mathbf{u}_E}^{(j)}]^\top d\Omega + \boldsymbol{\omega}_E \cdot \int_\Omega \tilde{\mathbf{x}}_E \wedge [(\boldsymbol{\varphi}_{\mathbf{u}_E}^{(i)})^\top \nabla_E \boldsymbol{\varphi}_{\mathbf{u}_E}^{(j)}]^\top d\Omega \\ &\quad - \boldsymbol{\omega}_E \cdot \int_\Omega \boldsymbol{\varphi}_{\mathbf{u}_E}^{(j)} \wedge \boldsymbol{\varphi}_{\mathbf{u}_E}^{(i)} d\Omega. \end{aligned} \quad (6.25)$$

L'expression générale des coefficients linéaires non-autonomes s'obtient alors comme la somme des contributions (6.23), (6.24) et (6.25). Le regroupement des termes conduit à l'expression

$$\begin{aligned} \mathcal{L}_{ij}^E(t) &= \mathbf{s}_{0,E}(t) \cdot \left[ \int_\Omega \varphi_\vartheta^{(i)} \nabla_E \varphi_\vartheta^{(j)} d\Omega + \int_\Omega [(\boldsymbol{\varphi}_{\mathbf{u}_E}^{(i)})^\top \nabla_E \boldsymbol{\varphi}_{\mathbf{u}_E}^{(j)}]^\top d\Omega + \int_\Omega \varphi_p^{(i)} \nabla_E \varphi_p^{(j)} d\Omega \right] \\ &\quad + \boldsymbol{\omega}_E(t) \cdot \left[ \int_\Omega [\tilde{\mathbf{x}}_E \wedge (\nabla_E \varphi_\vartheta^{(j)})] \varphi_\vartheta^{(i)} d\Omega + \int_\Omega [\tilde{\mathbf{x}}_E \wedge (\nabla_E \varphi_p^{(j)})] \varphi_p^{(i)} d\Omega \right. \\ &\quad \left. + \int_\Omega \tilde{\mathbf{x}}_E \wedge [(\boldsymbol{\varphi}_{\mathbf{u}_E}^{(i)})^\top \nabla_E \boldsymbol{\varphi}_{\mathbf{u}_E}^{(j)}]^\top d\Omega - \int_\Omega \boldsymbol{\varphi}_{\mathbf{u}_E}^{(j)} \wedge \boldsymbol{\varphi}_{\mathbf{u}_E}^{(i)} d\Omega \right] \end{aligned} \quad (6.26)$$

qui correspond à la forme recherchée puisque le coefficient linéaire non-autonome  $\mathcal{L}_{ij}^E(t)$  est la somme de deux produits scalaires entre d'une part des vecteurs de coefficients linéaires d'entraînement en translation  $\mathcal{L}_{ij}^{ET}$  et en rotation  $\mathcal{L}_{ij}^{ER}$  et d'autre part l'excitation temporelle instationnaire caractérisée par le vecteur vitesse de translation  $\mathbf{s}_{0,E}(t)$  et le vecteur vitesse angulaire  $\boldsymbol{\omega}_E(t)$ . On aboutit donc à l'expression synthétique

$$\boxed{\mathcal{L}_{ij}^E(t) = \mathbf{s}_{0,E}(t) \cdot \mathcal{L}_{ij}^{ET} + \boldsymbol{\omega}_E(t) \cdot \mathcal{L}_{ij}^{ER}}. \quad (6.27)$$

Les coefficients constants d'entraînement issus du mouvement de corps rigide se déduisent simplement en remplaçant les modes POD  $\varphi^{(j)}$  par la partie moyenne des clichés  $\bar{\mathbf{q}}$  dans l'expression (6.17), ce qui conduit à l'expression similaire

$$\boxed{\mathcal{K}_i^E(t) = \mathbf{s}_{0,E}(t) \cdot \mathcal{K}_i^{ET} + \boldsymbol{\omega}_E(t) \cdot \mathcal{K}_i^{ER}}. \quad (6.28)$$

La forme détaillée du modèle d'ordre réduit des équations de Navier-Stokes écrites dans le repère entraîné avec les composantes de la vitesse absolue s'écrit donc

$$\begin{aligned}
 \dot{a}_i(t) &= \mathcal{K}_i + \sum_{j=1}^q \mathcal{L}_{ij} a_j(t) + \sum_{j,k=1}^q \mathcal{Q}_{jik} a_j(t) a_k(t) \\
 &+ \mathbf{s}_{0,E}(t) \cdot \mathcal{K}_i^{ET} + \boldsymbol{\omega}_E(t) \cdot \mathcal{K}_i^{ER} \\
 &+ \sum_{j=1}^q \left( \mathbf{s}_{0,E}(t) \cdot \mathcal{L}_{ij}^{ET} + \boldsymbol{\omega}_E(t) \cdot \mathcal{L}_{ij}^{ER} \right) a_j(t)
 \end{aligned} \tag{6.29}$$

en remplaçant les expressions (6.27) et (6.28) dans l'équation (6.14). Pour chaque coefficient constant (resp. linéaire) d'entraînement  $\mathcal{K}_i^E(t)$  (resp.  $\mathcal{L}_{ij}^E(t)$ ), il faut calculer les vecteurs des coefficients constants (resp. linéaires) d'entraînement en translation  $\mathcal{K}_i^{ET}$  et en rotation  $\mathcal{K}_i^{ER}$  (resp.  $\mathcal{L}_{ij}^{ET}$  et  $\mathcal{L}_{ij}^{ER}$ ). Tout comme pour les coefficients autonomes, le calcul est effectué *hors ligne* avant l'intégration en temps du modèle d'ordre réduit. L'évaluation des coefficients est assez coûteuse puisqu'un modèle d'ordre réduit construit avec une base de projection comprenant  $q$  modes POD est défini par  $N_c^a = q + q^2 + q^3$  coefficients autonomes auxquels viennent s'ajouter  $N_c^{na} = 6q + 6q^2$  coefficients non-autonomes supplémentaires associés au mouvement de corps rigide. Une fois ce travail effectué, le coût de calcul pour l'intégration en temps du modèle d'ordre réduit est nettement réduit. En effet, il ne reste plus qu'à effectuer les produits scalaires  $\mathbf{s}_{0,E}(t) \cdot \mathcal{L}_{ij}^{ET}$  et  $\boldsymbol{\omega}_E(t) \cdot \mathcal{L}_{ij}^{ER}$  et leur équivalent pour les termes constants d'entraînement. Cette opération est relativement peu coûteuse puisqu'elle n'implique que des vecteurs de  $\mathbb{R}^3$ .

Le vecteur des coefficients linéaires d'entraînement en translation est défini par la relation

$$\mathcal{L}_{ij}^{ET} = \int_{\Omega} \varphi_{\vartheta}^{(i)} \nabla_E \varphi_{\vartheta}^{(j)} d\Omega + \int_{\Omega} \left[ \left( \varphi_{\mathbf{u}_E}^{(i)} \right)^{\top} \nabla_E \varphi_{\mathbf{u}_E}^{(j)} \right]^{\top} d\Omega + \int_{\Omega} \varphi_p^{(i)} \nabla_E \varphi_p^{(j)} d\Omega, \tag{6.30}$$

détaillée ci-dessous en développant les opérateurs tensoriels dans la base  $\mathcal{B}_E$  supposée cartésienne, ortho-normée et directe :

$$\mathcal{L}_{ij}^{ET} = \begin{bmatrix} \int_{\Omega} \left[ \varphi_{\vartheta}^{(i)} \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial x} + \varphi_p^{(i)} \frac{\partial \varphi_p^{(j)}}{\partial x} + \varphi_u^{(i)} \frac{\partial \varphi_u^{(j)}}{\partial x} + \varphi_v^{(i)} \frac{\partial \varphi_v^{(j)}}{\partial x} + \varphi_w^{(i)} \frac{\partial \varphi_w^{(j)}}{\partial x} \right] d\Omega \\ \int_{\Omega} \left[ \varphi_{\vartheta}^{(i)} \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial y} + \varphi_p^{(i)} \frac{\partial \varphi_p^{(j)}}{\partial y} + \varphi_u^{(i)} \frac{\partial \varphi_u^{(j)}}{\partial y} + \varphi_v^{(i)} \frac{\partial \varphi_v^{(j)}}{\partial y} + \varphi_w^{(i)} \frac{\partial \varphi_w^{(j)}}{\partial y} \right] d\Omega \\ \int_{\Omega} \left[ \varphi_{\vartheta}^{(i)} \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial z} + \varphi_p^{(i)} \frac{\partial \varphi_p^{(j)}}{\partial z} + \varphi_u^{(i)} \frac{\partial \varphi_u^{(j)}}{\partial z} + \varphi_v^{(i)} \frac{\partial \varphi_v^{(j)}}{\partial z} + \varphi_w^{(i)} \frac{\partial \varphi_w^{(j)}}{\partial z} \right] d\Omega \end{bmatrix}. \tag{6.31}$$

Le vecteur des coefficients linéaires d'entraînement en rotation est quant à lui défini par la relation

$$\begin{aligned}
 \mathcal{L}_{ij}^{ER} &= \int_{\Omega} \left[ \tilde{\mathbf{x}}_E \wedge \left( \nabla_E \varphi_{\vartheta}^{(j)} \right) \right] \varphi_{\vartheta}^{(i)} d\Omega + \int_{\Omega} \left[ \tilde{\mathbf{x}}_E \wedge \left( \nabla_E \varphi_p^{(j)} \right) \right] \varphi_p^{(i)} d\Omega \\
 &+ \int_{\Omega} \tilde{\mathbf{x}}_E \wedge \left[ \left( \varphi_{\mathbf{u}_E}^{(i)} \right)^{\top} \nabla_E \varphi_{\mathbf{u}_E}^{(j)} \right]^{\top} d\Omega - \int_{\Omega} \varphi_{\mathbf{u}_E}^{(j)} \wedge \varphi_{\mathbf{u}_E}^{(i)} d\Omega
 \end{aligned}, \tag{6.32}$$

dont le détail est aussi obtenu en développant les opérateurs tensoriels dans la base  $\mathcal{B}_E$ . En notant les composantes du vecteur des positions relatives  $\tilde{\mathbf{x}}_E = [x_E, y_E, z_E]^{\top}$  on aboutit après calcul à l'expression

$$\mathcal{L}_{ij}^{ER} = \left[ \begin{array}{l} \int_{\Omega} \left[ \left( y_E \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial z} - z_E \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial y} \right) \varphi_{\vartheta}^{(i)} + \left( y_E \frac{\partial \varphi_p^{(j)}}{\partial z} - z_E \frac{\partial \varphi_p^{(j)}}{\partial y} \right) \varphi_p^{(i)} - \left( \varphi_v^{(j)} \varphi_w^{(i)} - \varphi_w^{(j)} \varphi_v^{(i)} \right) \right. \\ \left. + y_E \left( \varphi_u^{(i)} \frac{\partial \varphi_u^{(j)}}{\partial z} + \varphi_v^{(i)} \frac{\partial \varphi_v^{(j)}}{\partial z} + \varphi_w^{(i)} \frac{\partial \varphi_w^{(j)}}{\partial z} \right) - z_E \left( \varphi_u^{(i)} \frac{\partial \varphi_u^{(j)}}{\partial y} + \varphi_v^{(i)} \frac{\partial \varphi_v^{(j)}}{\partial y} + \varphi_w^{(i)} \frac{\partial \varphi_w^{(j)}}{\partial y} \right) \right] d\Omega \\ \int_{\Omega} \left[ \left( z_E \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial x} - x_E \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial z} \right) \varphi_{\vartheta}^{(i)} + \left( z_E \frac{\partial \varphi_p^{(j)}}{\partial x} - x_E \frac{\partial \varphi_p^{(j)}}{\partial z} \right) \varphi_p^{(i)} - \left( \varphi_w^{(j)} \varphi_u^{(i)} - \varphi_u^{(j)} \varphi_w^{(i)} \right) \right. \\ \left. + z_E \left( \varphi_u^{(i)} \frac{\partial \varphi_u^{(j)}}{\partial x} + \varphi_v^{(i)} \frac{\partial \varphi_v^{(j)}}{\partial x} + \varphi_w^{(i)} \frac{\partial \varphi_w^{(j)}}{\partial x} \right) - x_E \left( \varphi_u^{(i)} \frac{\partial \varphi_u^{(j)}}{\partial z} + \varphi_v^{(i)} \frac{\partial \varphi_v^{(j)}}{\partial z} + \varphi_w^{(i)} \frac{\partial \varphi_w^{(j)}}{\partial z} \right) \right] d\Omega \\ \int_{\Omega} \left[ \left( x_E \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial y} - y_E \frac{\partial \varphi_{\vartheta}^{(j)}}{\partial x} \right) \varphi_{\vartheta}^{(i)} + \left( x_E \frac{\partial \varphi_p^{(j)}}{\partial y} - y_E \frac{\partial \varphi_p^{(j)}}{\partial x} \right) \varphi_p^{(i)} - \left( \varphi_u^{(j)} \varphi_v^{(i)} - \varphi_v^{(j)} \varphi_u^{(i)} \right) \right. \\ \left. + x_E \left( \varphi_u^{(i)} \frac{\partial \varphi_u^{(j)}}{\partial y} + \varphi_v^{(i)} \frac{\partial \varphi_v^{(j)}}{\partial y} + \varphi_w^{(i)} \frac{\partial \varphi_w^{(j)}}{\partial y} \right) - y_E \left( \varphi_u^{(i)} \frac{\partial \varphi_u^{(j)}}{\partial x} + \varphi_v^{(i)} \frac{\partial \varphi_v^{(j)}}{\partial x} + \varphi_w^{(i)} \frac{\partial \varphi_w^{(j)}}{\partial x} \right) \right] d\Omega \end{array} \right] \quad (6.33)$$

Le modèle d'ordre réduit construit par projection de Galerkin des équations de Navier-Stokes formulées dans le repère entraîné avec les composantes de la vitesse absolue pour un mouvement de corps rigide est donc défini par les équations (6.29) gouvernant les amplitudes modales. Les coefficients d'entraînement constants et linéaires  $\mathcal{K}_i^E(t)$  et  $\mathcal{L}_{ij}^E(t)$  ont été décomposés en différentes contributions au moyen des relations (6.28) et (6.27) afin de faire apparaître des coefficients indépendants du temps qui peuvent être calculés une fois pour toute avant l'intégration en temps du modèle d'ordre réduit. Pour cela, des coefficients d'entraînement en translation et en rotation ont été définis aux équations (6.30) et (6.32).

### Remarque 6.1 : Écriture tensorielle des termes non-autonomes.

Le modèle d'ordre réduit établi auparavant dans le cas de frontières fixes avait pu être écrit sous la forme tensorielle suivante :  $\dot{\mathbf{a}} = \underline{\mathcal{K}} + \underline{\mathcal{L}} \otimes \mathbf{a} + (\underline{\mathcal{Q}} \otimes \mathbf{a}) \otimes \mathbf{a}$ , où le nombre de barres sous les différents tenseurs indique leur ordre. Dans cette écriture, l'opérateur  $\otimes$  désigne le produit tensoriel contracté une fois qui correspond au produit matriciel classique lorsqu'il opère entre une matrice et un vecteur. Cette expression doit à présent être complétée par les termes non-autonomes  $\mathcal{K}_i^E(t)$  et  $\mathcal{L}_{ij}^E(t)$  qui ont été décomposés avec les expressions (6.28) et (6.27).

Pour les coefficients constants d'entraînement, on définit les matrices des coefficients constants d'entraînement en translation  $\underline{\mathcal{K}}^{ET} = [\mathcal{K}_1^{ET}, \dots, \mathcal{K}_q^{ET}]$  et en rotation  $\underline{\mathcal{K}}^{ER} = [\mathcal{K}_1^{ER}, \dots, \mathcal{K}_q^{ER}]$  qui regroupent les différents vecteurs des coefficients constants d'entraînement en translation  $\mathcal{K}_i^{ET}$  et en rotation  $\mathcal{K}_i^{ER}$ . Ces deux matrices sont dans  $\mathcal{M}(3, q)$  et le vecteur des coefficients constants non-autonomes est alors calculé par la relation  $\underline{\mathcal{K}}^E(t) = (\underline{\mathfrak{s}}_{0,E}^T \underline{\mathcal{K}}^{ET} + \underline{\omega}_E^T \underline{\mathcal{K}}^{ER})^T$  qui généralise l'équation (6.28) établie pour chaque composante. Cette relation peut aussi s'écrire avec un produit tensoriel contracté puisque pour tout vecteur  $\mathbf{u}$  et toute matrice  $\mathbf{D}$ ,  $\mathbf{u} \otimes \mathbf{D} = u_i \delta_i \otimes (D_{jk} \delta_j \otimes \delta_k) = u_i D_{jk} \delta_{ij} \delta_k = u_j D_{jk} \delta_k = \mathbf{u}^T \mathbf{D}$ . Par conséquent, le vecteur des coefficients constants non-autonomes s'écrit

$$\underline{\mathcal{K}}^E(t) = \underline{\mathfrak{s}}_{0,E}(t) \otimes \underline{\mathcal{K}}^{ET} + \underline{\omega}_E(t) \otimes \underline{\mathcal{K}}^{ER}. \quad (6.34)$$

On procède alors de façon similaire pour le vecteur des coefficients linéaires non-autonomes. Pour cela, les vecteurs  $\mathcal{L}_{ij}^{ET}$  et  $\mathcal{L}_{ij}^{ER}$  sont regroupés dans des tenseurs d'ordre 3 de taille  $3 \times q \times q$  notés  $\underline{\underline{\mathcal{L}}}^{ET}$  et  $\underline{\underline{\mathcal{L}}}^{ER}$ . En effet, le produit tensoriel contracté entre un vecteur  $\mathbf{u}$  et un tenseur d'ordre 3 noté  $\underline{\underline{\mathbf{L}}}$  donne une matrice de terme général  $u_i L_{ikl}$ . Par conséquent si l'on choisit de placer dans chaque colonne d'indices  $(k, l)$  le vecteur  $\mathcal{L}_{kl}^{ET}$  (et de même avec  $\mathcal{L}_{kl}^{ER}$ ), le terme général de la matrice  $\underline{\underline{\mathcal{L}}}_{kl}^E(t)$  définie par l'équation (6.27) s'obtient bien au moyen de la relation

$$\underline{\underline{\mathcal{L}}}^E(t) = \underline{\underline{\mathfrak{s}}}_{0,E}(t) \otimes \underline{\underline{\mathcal{L}}}^{ET} + \underline{\underline{\omega}}_E(t) \otimes \underline{\underline{\mathcal{L}}}^{ER}. \quad (6.35)$$

Une fois la matrice précédente multipliée par un autre vecteur  $\mathbf{v}$ , on retrouve bien un vecteur dont les composantes sont  $\mathcal{L}_{kl}^E(t) v_l$ , ce qui correspond à l'expression initiale de l'équation (6.29) avec  $\mathbf{v} = \mathbf{a}$ . Finalement, la forme tensorielle

du modèle d'ordre réduit avec les termes non-autonomes issus du mouvement est

$$\begin{aligned} \dot{\mathbf{a}}(t) &= \underline{\mathbf{K}} + \underline{\mathbf{L}} \overline{\mathbf{a}}(t) + \left( \underline{\mathbf{Q}} \overline{\mathbf{a}}(t) \right) \overline{\mathbf{a}}(t) \\ &+ \underline{\mathbf{s}}_{0,E}(t) \overline{\underline{\mathbf{K}}}^{ET} + \underline{\boldsymbol{\omega}}_E(t) \overline{\underline{\mathbf{K}}}^{ER} + \left( \underline{\mathbf{s}}_{0,E}(t) \overline{\underline{\mathbf{L}}}^{ET} + \underline{\boldsymbol{\omega}}_E(t) \overline{\underline{\mathbf{L}}}^{ER} \right) \overline{\mathbf{a}}(t) \end{aligned} \quad (6.36)$$

### 6.3.2 Développement d'un modèle d'ordre réduit POD-Galerkin basé sur la formulation en vitesse relative

Ce second modèle d'ordre réduit est construit par projection de Galerkin des équations continues (6.11) en suivant toujours le même principe que dans le paragraphe 5.2.1. Le système d'équations obtenu a exactement la même forme que celle associée au système d'équations (6.12) relatif à la formulation en vitesse absolue. La seule différence est que les opérateurs  $\mathbf{Q}^C$  et  $\mathbf{Q}^D$  s'appliquent cette fois à la vitesse relative  $\mathbf{v}_E$  au lieu de la vitesse absolue  $\mathbf{u}_E$ , et le vecteur des variables primitives modifiées est cette fois  $\mathbf{q} = [\vartheta, \mathbf{v}_E, p]^T$ . En outre, la définition du terme source change aussi et les différents opérateurs sont donc donnés par

$$\begin{aligned} \mathbf{Q}^C(\mathbf{q}, \mathbf{q}) &= \begin{bmatrix} -\mathbf{v}_E \cdot \nabla_E \vartheta + \vartheta \operatorname{div}_E \mathbf{v}_E \\ -\mathbf{v}_E \cdot \nabla_E \mathbf{v} - \vartheta \nabla_E p \\ -\mathbf{v}_E \cdot \nabla_E p - \gamma p \operatorname{div}_E \mathbf{v}_E \end{bmatrix}, \quad \mathbf{T}(\mathbf{q}, \mathbf{s}_{0,E}, \boldsymbol{\omega}_E) = \begin{bmatrix} 0 \\ \mathbf{f}_{\text{Cor}} + \mathbf{f}_{\text{cen}} \\ 0 \end{bmatrix}, \\ \mathbf{Q}^D(\mathbf{q}, \mathbf{q}) &= \begin{bmatrix} 0 \\ \vartheta \operatorname{div}_E \boldsymbol{\tau}(\mathbf{v}_E) \\ (\gamma - 1) \nabla_E \mathbf{v}_E : \boldsymbol{\tau}(\mathbf{v}_E) + \frac{\gamma}{\text{Pr}} \Delta_E(\vartheta p) \end{bmatrix}. \end{aligned} \quad (6.37)$$

Le terme source contient cette fois les accélérations de Coriolis et centrifuge définies à l'équation (4.34). Le terme source dépend donc encore de la vitesse du maillage  $\mathbf{s}_{e,E}$  et plus particulièrement des deux degrés de liberté du système que sont la vitesse de translation  $\mathbf{s}_{0,E}$  et la vitesse angulaire  $\boldsymbol{\omega}_E$ .

Le modèle d'ordre réduit est donc construit par projection de Galerkin des équations (6.11) sur chaque mode POD après avoir remplacé les variables par leur décomposition sur la base des modes POD. Les coefficients autonomes du modèle d'ordre réduit  $\mathcal{K}_i$ ,  $\mathcal{L}_{ij}$  et  $\mathcal{Q}_{ijk}$  restent identiques. La projection du terme source sur les modes POD se réduit à la projection des termes gyroscopiques sur la contribution du mode POD en vitesse qui s'écrit

$$\left\langle \mathbf{T}(\mathbf{q}, \mathbf{s}_{0,E}, \boldsymbol{\omega}_E), \boldsymbol{\varphi}^{(i)} \right\rangle = \left\langle \mathbf{f}_{\text{Cor}} + \mathbf{f}_{\text{cen}}, \boldsymbol{\varphi}_{\mathbf{v}_E}^{(i)} \right\rangle \quad (6.38)$$

puisque les autres composantes du terme source sont nulles. Les accélérations de Coriolis et centrifuge sont des termes linéaires en fonction des variables aérodynamiques et leur projection sur les modes POD conduit donc à l'apparition de coefficients constants et linéaires dans l'expression du modèle d'ordre réduit. Toutefois, le même problème que celui rencontré dans le paragraphe précédent ressurgit : les coefficients issus de la projection des termes gyroscopiques dépendent du temps puisque les accélérations de Coriolis et centrifuge sont définies à partir de la vitesse des nœuds du maillage. Par conséquent, le même travail que dans le paragraphe précédent doit être mené pour parvenir à déterminer des coefficients d'entraînement indépendants du temps.

**Projection de l'accélération de Coriolis** Le développement de la projection de l'accélération de Coriolis au moyen de l'expression (4.34) fait apparaître un produit mixte invariant par permutation circulaire :

$$\left\langle \mathbf{f}_{\text{Cor}}, \boldsymbol{\varphi}^{(i)} \right\rangle = -2 \int_{\Omega} (\boldsymbol{\omega}_E \wedge \mathbf{v}_E) \cdot \boldsymbol{\varphi}_{\mathbf{v}_E}^{(i)} d\Omega = -2 \int_{\Omega} \boldsymbol{\omega}_E \cdot (\mathbf{v}_E \wedge \boldsymbol{\varphi}_{\mathbf{v}_E}^{(i)}) d\Omega = -2 \boldsymbol{\omega}_E \cdot \int_{\Omega} \mathbf{v}_E \wedge \boldsymbol{\varphi}_{\mathbf{v}_E}^{(i)} d\Omega. \quad (6.39)$$

La dernière égalité est vraie puisque le vecteur vitesse angulaire est indépendant de la position des points du domaine d'intégration, si bien que la relation (6.19) peut être employée. L'introduction de la décomposition orthogonale propre du vecteur vitesse relative mène ensuite à

$$\left\langle \mathbf{f}_{\text{Cor}}, \boldsymbol{\varphi}^{(i)} \right\rangle = -2 \boldsymbol{\omega}_E \cdot \int_{\Omega} \nabla_E \wedge \boldsymbol{\varphi}_{\mathbf{v}_E}^{(i)} d\Omega - 2 \sum_{j=1}^q \left( \boldsymbol{\omega}_E \cdot \int_{\Omega} \boldsymbol{\varphi}_{\mathbf{v}_E}^{(j)} \wedge \boldsymbol{\varphi}_{\mathbf{v}_E}^{(i)} d\Omega \right) a_j. \quad (6.40)$$



Par analogie avec le formalisme adopté dans ce qui précède, la projection est écrite sous la forme

$$\boxed{\langle \mathbf{f}_{\text{Cor}}, \boldsymbol{\varphi}^{(i)} \rangle = -2 \boldsymbol{\omega}_E \cdot \boldsymbol{\mathcal{K}}_i^{\text{Cor}} - 2 \sum_{j=1}^q \left( \boldsymbol{\omega}_E \cdot \boldsymbol{\mathcal{L}}_{ij}^{\text{Cor}} \right) a_j}, \quad (6.41)$$

ce qui permet de définir les vecteurs de coefficients constants  $\boldsymbol{\mathcal{K}}_i^{\text{Cor}}$  et linéaires  $\boldsymbol{\mathcal{L}}_{ij}^{\text{Cor}}$  d'accélération de Coriolis. Ce dernier est donné par

$$\boxed{\boldsymbol{\mathcal{L}}_{ij}^{\text{Cor}} = \int_{\Omega} \boldsymbol{\varphi}_{\mathbf{v}_E}^{(j)} \wedge \boldsymbol{\varphi}_{\mathbf{v}_E}^{(i)} d\Omega} \quad (6.42)$$

et l'expression du vecteur de coefficients constants  $\boldsymbol{\mathcal{K}}_i^{\text{Cor}}$  se déduit de celle de  $\boldsymbol{\mathcal{L}}_{ij}^{\text{Cor}}$  en remplaçant  $\boldsymbol{\varphi}_{\mathbf{v}_E}^{(j)}$  par la partie moyenne des clichés de vitesse relative  $\bar{\mathbf{v}}_E$ . Le détail de l'expression précédente dans la base cartésienne et orthonormée du repère entraîné  $\mathcal{R}_E$  est donné par

$$\boxed{\boldsymbol{\mathcal{L}}_{ij}^{\text{Cor}} = \begin{bmatrix} \int_{\Omega} \varphi_v^{(j)} \varphi_w^{(i)} - \varphi_w^{(j)} \varphi_v^{(i)} d\Omega \\ \int_{\Omega} \varphi_w^{(j)} \varphi_u^{(i)} - \varphi_u^{(j)} \varphi_w^{(i)} d\Omega \\ \int_{\Omega} \varphi_u^{(j)} \varphi_v^{(i)} - \varphi_v^{(j)} \varphi_u^{(i)} d\Omega \end{bmatrix}}. \quad (6.43)$$

**Projection de l'accélération centrifuge** Le même travail est effectué pour l'accélération centrifuge dont l'expression est donnée à l'équation (4.34). À la différence de l'accélération de Coriolis, l'accélération centrifuge est indépendante du champ aérodynamique et résulte seulement du mouvement du repère entraîné. La projection sur les modes POD engendrera donc seulement des coefficients constants dans le modèle d'ordre réduit. L'expression de la projection de l'accélération centrifuge est donc décomposée au moyen des expressions (4.34) et (6.9) en deux parties telles que

$$\begin{aligned} \langle \mathbf{f}_{\text{cen}}, \boldsymbol{\varphi}^{(i)} \rangle &= - \int_{\Omega} \frac{\partial \mathbf{s}_{e,E}}{\partial t} \cdot \boldsymbol{\varphi}_{\mathbf{v}_E}^{(i)} d\Omega - \int_{\Omega} (\boldsymbol{\omega}_E \wedge \mathbf{s}_{e,E}) \cdot \boldsymbol{\varphi}_{\mathbf{v}_E}^{(i)} d\Omega \\ &= - \underbrace{\int_{\Omega} \frac{\partial \mathbf{s}_{0,E}}{\partial t} \cdot \boldsymbol{\varphi}_{\mathbf{v}_E}^{(i)} d\Omega - \int_{\Omega} \left( \frac{\partial \boldsymbol{\omega}_E}{\partial t} \wedge \tilde{\mathbf{x}}_E \right) \cdot \boldsymbol{\varphi}_{\mathbf{v}_E}^{(i)} d\Omega - \int_{\Omega} \left( \boldsymbol{\omega}_E \wedge \frac{\partial \tilde{\mathbf{x}}_E}{\partial t} \right) \cdot \boldsymbol{\varphi}_{\mathbf{v}_E}^{(i)} d\Omega}_{\textcircled{1}} \\ &\quad - \underbrace{\int_{\Omega} (\boldsymbol{\omega}_E \wedge \mathbf{s}_{0,E}) \cdot \boldsymbol{\varphi}_{\mathbf{v}_E}^{(i)} d\Omega - \int_{\Omega} (\boldsymbol{\omega}_E \wedge (\boldsymbol{\omega}_E \wedge \tilde{\mathbf{x}}_E)) \cdot \boldsymbol{\varphi}_{\mathbf{v}_E}^{(i)} d\Omega}_{\textcircled{2}}. \end{aligned} \quad (6.44)$$

Dans le premier terme de la contribution  $\textcircled{1}$ , on peut extraire directement de l'intégrale la dérivée partielle du vecteur vitesse de translation  $\mathbf{s}_{0,E}$  puisque celle-ci ne dépend que de la position de l'origine du repère entraîné. Le second terme de  $\textcircled{1}$  est modifié par permutation circulaire du produit mixte, afin d'obtenir un produit scalaire entre  $\partial \boldsymbol{\omega}_E / \partial t$  et les autres vecteurs. Ainsi le premier vecteur  $\partial \boldsymbol{\omega}_E / \partial t$  sera extrait de la seconde intégrale. Enfin, le troisième terme de  $\textcircled{1}$  est nul puisque la position relative  $\tilde{\mathbf{x}}_E$  est invariante dans le repère entraîné sous l'hypothèse d'un mouvement de corps rigide. Le terme  $\textcircled{1}$  se réécrit donc

$$\textcircled{1} = - \frac{\partial \mathbf{s}_{0,E}}{\partial t} \cdot \int_{\Omega} \boldsymbol{\varphi}_{\mathbf{v}_E}^{(i)} d\Omega - \frac{\partial \boldsymbol{\omega}_E}{\partial t} \cdot \int_{\Omega} \tilde{\mathbf{x}}_E \wedge \boldsymbol{\varphi}_{\mathbf{v}_E}^{(i)} d\Omega. \quad (6.45)$$

Dans le premier terme de  $\textcircled{2}$ , le produit vectoriel  $\boldsymbol{\omega}_E \wedge \mathbf{s}_{0,E}$  peut être entièrement extrait de l'intégrale puisque les deux vecteurs sont indépendants de la position. Le deuxième terme est plus délicat à traiter. On commence par noter  $\mathbf{A} = \boldsymbol{\omega}_E \wedge \tilde{\mathbf{x}}_E$ . Ainsi le double produit vectoriel projeté sur le mode  $\boldsymbol{\varphi}_{\mathbf{v}_E}^{(i)}$  s'interprète comme un produit mixte dont on effectue une permutation circulaire :  $(\boldsymbol{\omega}_E \wedge (\boldsymbol{\omega}_E \wedge \tilde{\mathbf{x}}_E)) \cdot \boldsymbol{\varphi}_{\mathbf{v}_E}^{(i)} = (\boldsymbol{\omega}_E \wedge \mathbf{A}) \cdot \boldsymbol{\varphi}_{\mathbf{v}_E}^{(i)}$ .

$\varphi_{\mathbf{v}_E}^{(i)} = (\mathbf{A} \wedge \varphi_{\mathbf{v}_E}^{(i)}) \cdot \boldsymbol{\omega}_E$ . Cette première opération permet ainsi d'écrire le terme sous l'intégrale du second terme de ② sous la forme

$$(\boldsymbol{\omega}_E \wedge (\boldsymbol{\omega}_E \wedge \tilde{\mathbf{x}}_E)) \cdot \varphi_{\mathbf{v}_E}^{(i)} = \boldsymbol{\omega}_E \cdot [(\boldsymbol{\omega}_E \wedge \tilde{\mathbf{x}}_E) \wedge \varphi_{\mathbf{v}_E}^{(i)}]. \quad (6.46)$$

Bien que le premier vecteur  $\boldsymbol{\omega}_E$  puisse être extrait de l'intégrale, le terme entre crochet dans l'équation (6.46) dépend encore de la vitesse angulaire  $\boldsymbol{\omega}_E$  qui est un terme instationnaire. L'expression précédente n'est donc pas encore totalement satisfaisante et pour pallier ce problème il est nécessaire d'isoler les deux vecteurs vitesse angulaire. Le double produit vectoriel est alors développé à l'aide de la formule  $(\mathbf{u} \wedge \mathbf{v}) \wedge \mathbf{w} = (\mathbf{u} \cdot \mathbf{w}) \mathbf{v} - (\mathbf{v} \cdot \mathbf{w}) \mathbf{u}$  qui, appliquée à l'équation (6.46), donne

$$\begin{aligned} (\boldsymbol{\omega}_E \wedge (\boldsymbol{\omega}_E \wedge \tilde{\mathbf{x}}_E)) \cdot \varphi_{\mathbf{v}_E}^{(i)} &= \boldsymbol{\omega}_E \cdot [(\boldsymbol{\omega}_E \cdot \varphi_{\mathbf{v}_E}^{(i)}) \tilde{\mathbf{x}}_E - (\tilde{\mathbf{x}}_E \cdot \varphi_{\mathbf{v}_E}^{(i)}) \boldsymbol{\omega}_E] \\ &= (\boldsymbol{\omega}_E \cdot \tilde{\mathbf{x}}_E) (\boldsymbol{\omega}_E \cdot \varphi_{\mathbf{v}_E}^{(i)}) - (\tilde{\mathbf{x}}_E \cdot \varphi_{\mathbf{v}_E}^{(i)}) (\boldsymbol{\omega}_E \cdot \boldsymbol{\omega}_E). \end{aligned} \quad (6.47)$$

Le dernier terme dans la seconde égalité présente une forme satisfaisante puisque le vecteur vitesse angulaire apparaît de manière isolée. En revanche le vecteur vitesse angulaire  $\boldsymbol{\omega}_E$  dans le premier terme est impliqué dans les deux produits scalaires impliquant chacun un vecteur dépendant de la position spatiale des points du domaine. Pour remédier à cela, il suffit de remarquer qu'un scalaire issu du produit  $(\mathbf{a} \cdot \mathbf{b})(\mathbf{c} \cdot \mathbf{d})$  résulte de  $(a_i b_i)(c_j d_j) = (a_i c_j)(b_i d_j)$ . Par conséquent on obtient le même résultat avec le double produit de contraction des tenseurs  $\mathbf{a} \otimes \mathbf{c}$  et  $\mathbf{b} \otimes \mathbf{d}$ . En appliquant ce résultat au premier terme de la dernière égalité de l'équation (6.47), on peut transformer le terme ② en

$$\textcircled{2} = -(\boldsymbol{\omega}_E \wedge \mathbf{s}_{0,E}) \cdot \int_{\Omega} \varphi_{\mathbf{v}_E}^{(i)} d\Omega - (\boldsymbol{\omega}_E \otimes \boldsymbol{\omega}_E) : \int_{\Omega} \varphi_{\mathbf{v}_E}^{(i)} \otimes \tilde{\mathbf{x}}_E d\Omega + (\boldsymbol{\omega}_E \cdot \boldsymbol{\omega}_E) \int_{\Omega} \tilde{\mathbf{x}}_E \cdot \varphi_{\mathbf{v}_E}^{(i)} d\Omega. \quad (6.48)$$

La projection du terme centrifuge sur chaque mode POD, définie à l'équation (6.44), s'écrit donc

$$\begin{aligned} \langle \mathbf{f}_{\text{cen}}, \varphi^{(i)} \rangle &= -\left(\frac{\partial \mathbf{s}_{0,E}}{\partial t} + \boldsymbol{\omega}_E \wedge \mathbf{s}_{0,E}\right) \cdot \int_{\Omega} \varphi_{\mathbf{v}_E}^{(i)} d\Omega - \frac{\partial \boldsymbol{\omega}_E}{\partial t} \cdot \int_{\Omega} \tilde{\mathbf{x}}_E \wedge \varphi_{\mathbf{v}_E}^{(i)} d\Omega \\ &\quad - (\boldsymbol{\omega}_E \otimes \boldsymbol{\omega}_E) : \int_{\Omega} \varphi_{\mathbf{v}_E}^{(i)} \otimes \tilde{\mathbf{x}}_E d\Omega + (\boldsymbol{\omega}_E \cdot \boldsymbol{\omega}_E) \int_{\Omega} \tilde{\mathbf{x}}_E \cdot \varphi_{\mathbf{v}_E}^{(i)} d\Omega, \end{aligned} \quad (6.49)$$

où intervient une somme de produits scalaires entre des coefficients indépendants du temps et les vecteurs des degrés de liberté qui décrivent l'évolution temporelle du mouvement du domaine. L'expression obtenue est donc mise sous la forme synthétique

$$\boxed{\langle \mathbf{f}_{\text{cen}}, \varphi^{(i)} \rangle = -\left(\frac{\partial \mathbf{s}_{0,E}}{\partial t} + \boldsymbol{\omega}_E \wedge \mathbf{s}_{0,E}\right) \cdot \underline{\mathcal{K}}_i^{\text{Cor},1} - \frac{\partial \boldsymbol{\omega}_E}{\partial t} \cdot \underline{\mathcal{K}}_i^{\text{Cor},2} - (\boldsymbol{\omega}_E \otimes \boldsymbol{\omega}_E) : \underline{\mathcal{K}}_i^{\text{Cor},3} + (\boldsymbol{\omega}_E \cdot \boldsymbol{\omega}_E) \mathcal{K}_i^{\text{Cor},4}}, \quad (6.50)$$

pour faire clairement apparaître le fait que des coefficients indépendants du temps ont pu être déterminés. La projection de l'accélération centrifuge conduit à une forme plus compliquée qui fait notamment intervenir pour chaque mode POD de projection un coefficient scalaire  $\mathcal{K}_i^{\text{Cor},4}$ , deux vecteurs de coefficients  $\underline{\mathcal{K}}_i^{\text{Cor},1}$  et  $\underline{\mathcal{K}}_i^{\text{Cor},2}$  et enfin une matrice de coefficients  $\underline{\mathcal{K}}_i^{\text{Cor},3}$  définis par

$$\boxed{\begin{aligned} \underline{\mathcal{K}}_i^{\text{cen},1} &= \int_{\Omega} \varphi_{\mathbf{v}_E}^{(i)} d\Omega & \underline{\mathcal{K}}_i^{\text{cen},2} &= \int_{\Omega} \tilde{\mathbf{x}}_E \wedge \varphi_{\mathbf{v}_E}^{(i)} d\Omega \\ \underline{\mathcal{K}}_i^{\text{cen},3} &= \int_{\Omega} \varphi_{\mathbf{v}_E}^{(i)} \otimes \tilde{\mathbf{x}}_E d\Omega & \mathcal{K}_i^{\text{cen},4} &= \int_{\Omega} \tilde{\mathbf{x}}_E \cdot \varphi_{\mathbf{v}_E}^{(i)} d\Omega \end{aligned}}. \quad (6.51)$$

Les expressions détaillées de ces différents coefficients sont données ci-dessous dans la base  $\mathcal{B}_E$  supposée cartésienne et orthonormée. Les vecteurs de coefficients  $\underline{\mathcal{K}}_i^{\text{cen},1}$  et  $\underline{\mathcal{K}}_i^{\text{cen},2}$  s'écrivent respectivement

$$\underline{\mathcal{K}}_i^{\text{cen},1} = \begin{bmatrix} \int_{\Omega} \varphi_u^{(i)} d\Omega \\ \int_{\Omega} \varphi_v^{(i)} d\Omega \\ \int_{\Omega} \varphi_w^{(i)} d\Omega \end{bmatrix} \quad \text{et} \quad \underline{\mathcal{K}}_i^{\text{cen},2} = \begin{bmatrix} \int_{\Omega} y_E \varphi_w^{(i)} - z_E \varphi_v^{(i)} d\Omega \\ \int_{\Omega} z_E \varphi_u^{(i)} - x_E \varphi_w^{(i)} d\Omega \\ \int_{\Omega} x_E \varphi_v^{(i)} - y_E \varphi_u^{(i)} d\Omega \end{bmatrix}, \quad (6.52)$$

le scalaire  $\mathcal{K}_i^{\text{cen},4}$  prend la forme

$$\mathcal{K}_i^{\text{cen},4} = \int_{\Omega} x_E \varphi_u^{(i)} + y_E \varphi_v^{(i)} + z_E \varphi_w^{(i)} d\Omega, \quad (6.53)$$

et enfin les termes de la matrice  $\underline{\mathcal{K}}_i^{\text{cen},3}$  sont donnés par

$$\underline{\mathcal{K}}_i^{\text{cen},3} = \begin{bmatrix} \int_{\Omega} \varphi_u^{(i)} x_E d\Omega & \int_{\Omega} \varphi_u^{(i)} y_E d\Omega & \int_{\Omega} \varphi_u^{(i)} z_E d\Omega \\ \int_{\Omega} \varphi_v^{(i)} x_E d\Omega & \int_{\Omega} \varphi_v^{(i)} y_E d\Omega & \int_{\Omega} \varphi_v^{(i)} z_E d\Omega \\ \int_{\Omega} \varphi_w^{(i)} x_E d\Omega & \int_{\Omega} \varphi_w^{(i)} y_E d\Omega & \int_{\Omega} \varphi_w^{(i)} z_E d\Omega \end{bmatrix}. \quad (6.54)$$

Finalement, le modèle d'ordre réduit des équations de Navier-Stokes construit à partir de la formulation en repère entraîné avec les composantes de la vitesse relative s'écrit

$$\begin{aligned} \dot{a}_i(t) &= \mathcal{K}_i + \sum_{j=1}^q \mathcal{L}_{ij} a_j(t) + \sum_{j,k=1}^q \mathcal{Q}_{jik} a_j(t) a_k(t) \\ &- 2 \boldsymbol{\omega}_E(t) \cdot \underline{\mathcal{K}}_i^{\text{Cor}} - 2 \sum_{j=1}^q \left( \boldsymbol{\omega}_E(t) \cdot \underline{\mathcal{L}}_{ij}^{\text{Cor}} \right) a_j(t) \\ &- \left( \frac{\partial \mathbf{s}_{0,E}(t)}{\partial t} + \boldsymbol{\omega}_E(t) \wedge \mathbf{s}_{0,E}(t) \right) \cdot \underline{\mathcal{K}}_i^{\text{cen},1} - \frac{\partial \boldsymbol{\omega}_E(t)}{\partial t} \cdot \underline{\mathcal{K}}_i^{\text{cen},2} \\ &- (\boldsymbol{\omega}_E(t) \otimes \boldsymbol{\omega}_E(t)) : \underline{\mathcal{K}}_i^{\text{cen},3} + (\boldsymbol{\omega}_E(t) \cdot \boldsymbol{\omega}_E(t)) \mathcal{K}_i^{\text{cen},4} \end{aligned} \quad (6.55)$$

Les accélérations de Coriolis et centrifuge induisent six termes supplémentaires pour lesquels un nombre important de coefficients doit être déterminé. Ainsi, lorsque la base de projection comprend  $q$  modes POD, le nombre de coefficients autonomes s'élève à  $N_c^a = q + q^2 + q^3$  pour les opérateurs  $\mathcal{K}$ ,  $\mathcal{L}$  et  $\mathcal{Q}$ , auxquels viennent s'ajouter  $3q + 3q^2$  coefficients pour les termes  $\underline{\mathcal{K}}_i^{\text{Cor}}$  et  $\underline{\mathcal{L}}_{ij}^{\text{Cor}}$  relatifs à l'accélération de Coriolis, et encore  $3q + 3q + 9q + q$  coefficients pour les termes  $\underline{\mathcal{K}}_i^{\text{cen},1}$ ,  $\underline{\mathcal{K}}_i^{\text{cen},2}$ ,  $\underline{\mathcal{K}}_i^{\text{cen},3}$  et  $\mathcal{K}_i^{\text{cen},4}$  relatifs à l'accélération centrifuge, soit un total de  $N_c^{\text{na}} = 19q + q^2$  coefficients non-autonomes. Au total, le modèle d'ordre réduit des équations en repère entraîné avec les composantes de la vitesse relative requiert le calcul de  $20q + 4q^2 + q^3$  coefficients.

### 6.3.3 Construction de modèles d'ordre réduit des équations de Navier-Stokes par identification des coefficients

Au lieu de construire le modèle d'ordre réduit à l'aide des expressions analytiques développées précédemment pour les formulations en vitesse absolue et relative, il est possible d'identifier directement les coefficients à partir des clichés selon le principe d'identification exposé au paragraphe 2.4.4. Cette technique devient d'autant plus intéressante que le nombre de coefficients à évaluer pour ces nouvelles formulations du modèle d'ordre réduit a été augmenté.

La méthode d'identification des coefficients du modèle d'ordre réduit a été présentée au paragraphe 2.4.4 dans le cas d'un système autonome similaire à celui obtenu par la projection de Galerkin des équations de Navier-Stokes dans un domaine fluide de frontières fixes. Or la projection de Galerkin des équations de Navier-Stokes formulées avec les composantes de la vitesse absolue ou relative avait conduit aux systèmes non-autonomes (6.29) et (6.55) respectivement. Connaissant la forme des modèles d'ordre réduit, il est donc possible d'étendre le problème d'identification (2.65) afin de déterminer les coefficients supplémentaires provenant du mouvement du domaine.

L'extension aux modèles d'ordre réduit (6.29) et (6.55) est relativement triviale une fois que la forme du modèle d'ordre réduit est connue. Il suffit alors de former un nouveau vecteur d'inconnues  $\mathcal{X}_i$  contenant à la fois les coefficients du modèle d'ordre réduit autonome et les coefficients supplémentaires qui rendent le système non-autonome. Par ailleurs, la matrice  $\mathcal{A}$  doit aussi être adaptée pour tenir compte des corrélations supplémentaires. Le problème d'identification (2.65) conserve donc la même forme et la résolution est menée au moyen de la technique présentée au paragraphe A.3. La seule différence réside dans la définition du vecteur  $\mathcal{X}_i$  et de la matrice  $\mathcal{A}$  qui sont décomposés en deux contributions telles que

$$\mathcal{X}_i = [\mathcal{X}_i^a \quad \mathcal{X}_i^{\text{na}}]^\top \quad \text{et} \quad \mathcal{A} = [\mathcal{A}^a \quad \mathcal{A}^{\text{na}}]. \quad (6.56)$$

Les contributions  $\mathcal{X}_i^a$  et  $\mathcal{A}^a$  contiennent les termes liés à la partie autonome du modèle d'ordre réduit et sont ainsi données par les relations (2.64) et (2.66) respectivement. Le terme  $\mathcal{B}_i$  reste quant à lui inchangé. Les termes  $\mathcal{X}_i^{\text{na}}$  et  $\mathcal{A}^{\text{na}}$  apparaissent en raison du mouvement du domaine et leurs expressions dépendent de la formulation choisie pour les composantes de la vitesse. Les détails sont fournis ci-dessous lorsque la vitesse absolue ou relative est considérée.

**Formulation en vitesse absolue** Dans ce cas, la formulation du modèle d'ordre réduit (6.29) introduit  $6 + 6q$  coefficients supplémentaires pour chaque équation gouvernant l'amplitude modale  $a_i$ . Par conséquent, le vecteur des coefficients non-autonomes  $\mathcal{X}_i^{\text{na}}$  contient  $N_{c,i}^{\text{na}} = 6 + 6q$  coefficients provenant à la fois des contributions constantes  $\mathcal{K}_i^{ET}$  et  $\mathcal{K}_i^{ER}$  et des contributions linéaires  $\mathcal{L}_{ij}^{ET}$  et  $\mathcal{L}_{ij}^{ER}$ . En notant  $\mathcal{L}_{ij_0}^{ET}$  le vecteur colonne relatif à l'indice  $j_0$  on parvient à l'expression

$$\mathcal{X}_i^{\text{na}} = \left[ \left( \mathcal{K}_i^{ET} \right)^\top \quad \left( \mathcal{K}_i^{ER} \right)^\top \quad \left( \mathcal{L}_{i1}^{ET} \right)^\top \dots \left( \mathcal{L}_{iq}^{ET} \right)^\top \quad \left( \mathcal{L}_{i1}^{ER} \right)^\top \dots \left( \mathcal{L}_{iq}^{ER} \right)^\top \right]^\top. \quad (6.57)$$

La matrice  $\mathcal{A}^{\text{na}} \in \mathcal{M}(q, N_{c,i}^{\text{na}})$  prend une forme similaire notée ici

$$\mathcal{A}^{\text{na}} = \left[ \mathcal{A}^{\mathcal{K},ET}(\mathbf{s}_{0,E}) \quad \mathcal{A}^{\mathcal{K},ER}(\boldsymbol{\omega}_E) \quad \mathcal{A}^{\mathcal{L},ET}(\mathbf{s}_{0,E}, \mathbf{a}^{(m)}) \quad \mathcal{A}^{\mathcal{L},ER}(\boldsymbol{\omega}_E, \mathbf{a}^{(m)}) \right]^\top. \quad (6.58)$$

**Formulation en vitesse relative** Pour terminer, la formulation du modèle d'ordre réduit (6.55) avec les vitesses relatives introduit  $N_{c,i}^{\text{na}} = 19 + 3q$  coefficients supplémentaires pour chaque équation gouvernant l'amplitude modale  $a_i$ . Le vecteur des coefficients non-autonomes  $\mathcal{X}_i^{\text{na}}$  contient ainsi  $3 + 3q$  coefficients provenant des contributions constantes  $\mathcal{K}_i^{\text{Cor}}$  et linéaires  $\mathcal{L}_{ij}^{\text{Cor}}$  dues à l'accélération de Coriolis, puis encore 16 coefficients résultants de l'accélération centrifuge caractérisée par les coefficients  $\mathcal{K}_i^{\text{cen},1}$ ,  $\mathcal{K}_i^{\text{cen},2}$ ,  $\mathcal{K}_i^{\text{cen},3}$  et  $\mathcal{K}_i^{\text{cen},4}$ . L'expression du vecteur  $\mathcal{X}_i$  est alors donnée par

$$\mathcal{X}_i^{\text{na}} = \left[ \begin{array}{cccc} \left( \mathcal{K}_i^{\text{Cor}} \right)^\top & \left( \mathcal{L}_{i1}^{\text{Cor}} \right)^\top & \dots & \left( \mathcal{L}_{iq}^{\text{Cor}} \right)^\top \quad \left( \mathcal{K}_i^{\text{cen},1} \right)^\top \\ \left( \mathcal{K}_i^{\text{cen},2} \right)^\top & \left( \mathcal{K}_i^{\text{cen},3} \right)^\top & \left( \mathcal{K}_i^{\text{cen},3} \right)^\top & \left( \mathcal{K}_i^{\text{cen},3} \right)^\top \quad \mathcal{K}_i^{\text{cen},4} \end{array} \right], \quad (6.59)$$

et celle de la matrice  $\mathcal{A}^{\text{na}}$  pour la formulation en vitesse relative s'écrit

$$\mathcal{A}^{\text{na}} = \left[ \begin{array}{ccc} \mathcal{A}^{\mathcal{K},\text{Cor}}(\boldsymbol{\omega}_E) & \mathcal{A}^{\mathcal{L},\text{Cor}}(\boldsymbol{\omega}_E, \mathbf{a}^{(m)}) & \mathcal{A}^{\mathcal{K},\text{cen},1}(\mathbf{s}_{0,E}, \boldsymbol{\omega}_E) \\ \mathcal{A}^{\mathcal{K},\text{cen},2}(\boldsymbol{\omega}_E) & \mathcal{A}^{\mathcal{K},\text{cen},3}(\boldsymbol{\omega}_E) & \mathcal{A}^{\mathcal{K},\text{cen},4}(\boldsymbol{\omega}_E) \end{array} \right]. \quad (6.60)$$

### 6.3.4 Résumé et comparaison qualitative des différents modèles d'ordre réduit POD-Galerkin développés

Plusieurs modèles d'ordre réduit basés sur des formulations différentes des équations de Navier-Stokes ont été développés dans les chapitres 5 et 6. Dans tous les cas, ces modèles d'ordre réduit sont valables pour reproduire l'écoulement instationnaire d'un fluide compressible visqueux dans lequel des phénomènes non-linéaires peuvent être rencontrés. Trois formulations ont été développées afin de pouvoir tenir compte de l'effet d'une structure placée dans le fluide. Ces formulations se distinguent d'abord par le repère dans lequel sont formulées les équations projetées et ensuite par les composantes de la vitesse employée. Ainsi, lorsque la structure est fixe, la description eulérienne de l'écoulement dans le repère absolu  $\mathcal{R}_A$  avec les composantes de la vitesse absolue exprimée dans ce repère est employée, d'où la dénomination FEUL-RAUA pour ce modèle d'ordre réduit. Deux formulations du modèle d'ordre réduit des équations de Navier-Stokes dans le repère entraîné  $\mathcal{R}_E$  ont ensuite été développées : la première repose sur l'emploi des composantes de la vitesse absolue exprimée dans  $\mathcal{R}_E$  ce qui conduit au modèle d'ordre réduit FALE-REUE, tandis que la seconde repose sur les composantes de la vitesse relative exprimée dans  $\mathcal{R}_E$  ce qui conduit au modèle d'ordre réduit FALE-REVE. Dans les deux cas, on a supposé en outre que la structure est mobile mais non déformable afin de s'affranchir des problèmes d'incohérence dans la formulation de la décomposition orthogonale propre sur un domaine déformable. Le tableau 6.1 résume les différentes formulations des modèles d'ordre réduit développés et indique certaines de leurs spécificités.

Le premier modèle d'ordre réduit développé conduit à un système d'équations différentielles ordinaires quadratiques autonomes. Dans ce cas, l'instationnarité de l'écoulement doit être d'origine purement aérodynamique puisque la structure est supposée fixe. Deux autres formulations du modèle d'ordre réduit, toujours basées sur les équations de Navier-Stokes, reposent sur l'écriture des équations dans le repère entraîné avec les composantes de la vitesse absolue ou relative. Dans ce cas, les modèles d'ordre réduit obtenus ne sont plus autonomes puisque des termes sources supplémentaires liés à la prise en compte du mouvement du domaine apparaissent. Ces termes permettent ainsi d'obtenir la réponse avec le modèle d'ordre réduit pour un certain déplacement du domaine qui correspond au mouvement imposé à la structure puisque les équations ont été formulées dans le repère entraîné attaché à la structure. Si le mouvement imposé à la structure est un mouvement d'oscillation harmonique de l'angle d'incidence par exemple, il suffit de changer la fréquence ou l'amplitude des oscillations pour obtenir la réponse sur un certain intervalle de fréquence et/ou d'amplitude avec le modèle d'ordre réduit. Pour cela, il suffit simplement de modifier le vecteur de la vitesse angulaire  $\omega_E$  au second membre du modèle d'ordre réduit.

Les formulations du modèle d'ordre réduit en repère entraîné ont été développées pour un mouvement impliquant une seule rotation et une seule translation du domaine. Une généralisation au cas de mouvement plus complexes impliquant un nombre quelconque de repère pourrait être établie mais on se contentera dans le cadre de cette thèse d'un mouvement simple impliquant un seul repère entraîné.

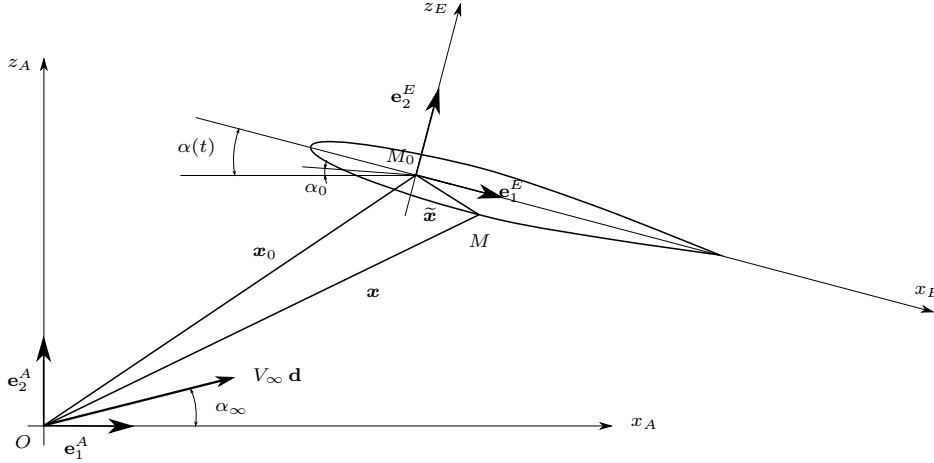
Formulation	Repère	Vitesse	Mouvement de la structure	Expression du modèle d'ordre réduit	Nombre de coefficients
FEUL-RAUA	$\mathcal{R}_A$	$\mathbf{u}_A$	Structure fixe	$\dot{a}_i = \mathcal{K}_i + \sum_{j=1}^q \mathcal{L}_{ij} a_j + \sum_{j,k=1}^q \mathcal{Q}_{jik} a_j a_k$	$q + q^2 + q^3$
FALE-REUE	$\mathcal{R}_E$	$\mathbf{u}_E$	Mouvement de corps rigide	$\begin{aligned} \dot{a}_i &= \mathcal{K}_i + \sum_{j=1}^q \mathcal{L}_{ij} a_j + \sum_{j,k=1}^q \mathcal{Q}_{jik} a_j a_k \\ &+ \mathbf{s}_{0,E} \cdot \mathcal{K}_i^{ET} + \boldsymbol{\omega}_E \cdot \mathcal{K}_i^{ER} \\ &+ \sum_{j=1}^q \left( \mathbf{s}_{0,E} \cdot \mathcal{L}_{ij}^{ET} + \boldsymbol{\omega}_E \cdot \mathcal{L}_{ij}^{ER} \right) a_j \end{aligned}$	$7q + 7q^2 + q^3$
FALE-REVE	$\mathcal{R}_E$	$\mathbf{v}_E$	Mouvement de corps rigide	$\begin{aligned} \dot{a}_i &= \mathcal{K}_i + \sum_{j=1}^q \mathcal{L}_{ij} a_j + \sum_{j,k=1}^q \mathcal{Q}_{jik} a_j a_k \\ &- 2 \boldsymbol{\omega}_E \cdot \mathcal{K}_i^{\text{Cor}} - 2 \sum_{j=1}^q \left( \boldsymbol{\omega}_E(t) \cdot \mathcal{L}_{ij}^{\text{Cor}} \right) a_j \\ &- \left( \frac{\partial \mathbf{s}_{0,E}}{\partial t} + \boldsymbol{\omega}_E \wedge \mathbf{s}_{0,E} \right) \cdot \mathcal{K}_i^{\text{cen},1} - \frac{\partial \boldsymbol{\omega}_E}{\partial t} \cdot \mathcal{K}_i^{\text{cen},2} \\ &- (\boldsymbol{\omega}_E \otimes \boldsymbol{\omega}_E) : \mathcal{K}_i^{\text{cen},3} + (\boldsymbol{\omega}_E \cdot \boldsymbol{\omega}_E) \mathcal{K}_i^{\text{cen},4} \end{aligned}$	$20q + 4q^2 + q^3$

**Tableau 6.1 :** Tableau comparatif des différents modèles d'ordre réduit développés à partir des équations de Navier-Stokes. Pour chacun d'entre eux, on indique dans la première colonne le nom de la formulation qui repose sur le repère dans lequel sont formulées les équations et sur les composantes de la vitesse utilisées. Ces deux paramètres sont précisés dans les colonnes suivantes :  $\mathcal{R}_A$  désigne le repère absolu tandis que  $\mathcal{R}_E$  désigne le repère entraîné ; le vecteur  $\mathbf{u}_A$  représente les composantes de la vitesse absolue exprimée dans  $\mathcal{R}_A$  alors que le vecteur  $\mathbf{u}_E$  représente les composantes de la vitesse absolue exprimée dans  $\mathcal{R}_E$ , et  $\mathbf{v}_E$  est le vecteur contenant les composantes de la vitesse relative exprimée dans  $\mathcal{R}_E$ . Le type de mouvement pris en compte par chaque modèle d'ordre réduit est donné dans la 4<sup>e</sup> colonne. Enfin l'expression du modèle d'ordre réduit est fournie dans la 5<sup>e</sup> colonne et le nombre de coefficients  $N_c$  à évaluer pour déterminer complètement le modèle d'ordre réduit est indiqué dans la dernière colonne.

## 6.4 Calcul des modes POD de l'écoulement transsonique autour d'un profil NACA0064 oscillant

### 6.4.1 Simulation numérique de l'écoulement transsonique autour du profil NACA0064 pour la génération des clichés

Un premier pas vers la modélisation d'un écoulement supportant des phénomènes aéroélastiques consiste à étudier le mouvement d'une structure mobile supposée rigide placée dans le fluide. Le problème est encore simplifié en procédant à un couplage faible dans lequel le mouvement de la structure est imposé. Dans le cas présent on considère un profil NACA0064 bidimensionnel, animé d'un mouvement forcé de corps rigide. Ce mouvement combine à la fois une translation qui représente l'avancement du profil<sup>3</sup>, et une oscillation autour de la position d'équilibre qui traduit la torsion d'une voilure sous l'effet des charges aérodynamiques. Le repère absolu  $\mathcal{R}_A$  est caractérisé par l'origine  $O$  et la base orthonormée directe de vecteurs  $\mathcal{B}_A = (\mathbf{e}_1^A, \mathbf{e}_2^A, \mathbf{e}_3^A)$  pointant dans les directions  $x_A, y_A$  et  $z_A$ . Le repère entraîné  $\mathcal{R}_E$  d'origine  $M_0$  est quant à lui attaché au mouvement de corps rigide du profil et le suit donc dans son mouvement de translation et de rotation. Le schéma de la figure 6.2 représente le système modélisé.



**Figure 6.2 :** Schéma du profil NACA64 en mouvement de translation uniforme à la vitesse  $V_\infty$  et en rotation autour de l'axe  $y_E$  d'un angle  $\alpha(t)$  par rapport à la direction horizontale  $x_A$ . Le domaine d'étude est repéré dans la base  $\mathcal{B}_E$  entraînée par le mouvement de corps rigide, de sorte que la distance relative  $\tilde{\mathbf{x}}$  d'un point  $M$  soit constante dans cette base.

Le mouvement de translation est supposé uniforme et s'effectue dans la direction  $\mathbf{d}$  à la vitesse constante  $V_\infty$  dont la valeur est donnée par les conditions expérimentales pour un nombre de Mach à l'infini  $Ma_\infty$  et la célérité du son à l'infini  $c_\infty$ . La direction de translation est éventuellement inclinée d'un angle d'incidence  $\alpha_\infty$ , si bien que le vecteur vitesse de translation du repère entraîné est donné par

$$\mathbf{s}_{0,A} = V_\infty \mathbf{d} = V_\infty \begin{bmatrix} \cos \alpha_\infty \\ 0 \\ \sin \alpha_\infty \end{bmatrix}. \quad (6.61)$$

La translation s'effectue donc à vitesse constante ce qui simplifiera en partie les termes induits par le mouvement d'entraînement dans la formulation des équations du fluide en repère entraîné. En revanche, un mouvement d'oscillation harmonique instationnaire est imposé au profil. Celui-ci est décrit par l'angle d'incidence  $\alpha(t)$  dont la loi d'évolution est

<sup>3</sup>Le mouvement de translation du profil est ici associé au déplacement du repère  $\mathcal{R}_E$  par rapport au repère  $\mathcal{R}_A$ . Une autre manière de procéder consiste à reporter le mouvement de translation du profil dans les conditions limites de l'écoulement à l'infini. Dans ce cas, le vecteur vitesse de translation est nul :  $\mathbf{s}_0 = \mathbf{0}$ .

$$\alpha(t) = \alpha_0 + \alpha_m \sin(\omega_\alpha t). \quad (6.62)$$

À l'équilibre, le profil est incliné d'un angle  $\alpha_0$  et les oscillations d'amplitude maximale  $\alpha_m$  s'effectuent autour de cette position d'équilibre à la fréquence  $f_\alpha = \omega_\alpha/(2\pi)$ . La période des oscillations forcées est donc  $T_\alpha = 1/f_\alpha$ . Bien souvent, la fréquence circulaire  $\omega_\alpha$  se déduit de la pulsation réduite

$$k = \frac{\omega_\alpha}{V_\infty} \frac{L_0}{2} \quad (6.63)$$

définie comme paramètre du modèle avec une longueur de référence  $L_0$  souvent prise égale à la corde du profil. La rotation du profil autour de l'axe  $y_E$  est caractérisée par l'angle  $\alpha(t)$ . Cependant, l'angle d'incidence est défini pour un profil pointant « vers le haut » pour reprendre la modélisation adoptée dans *elsA*, et par conséquent l'angle de rotation entre les bases  $\mathcal{B}_E$  et  $\mathcal{B}_A$  est en réalité  $\alpha' = 2\pi - \alpha$ . La matrice de passage associée au mouvement de rotation correspond à l'expression de  $\mathbf{R}_y$  donnée à l'équation (E.6) pour un angle  $\alpha'$ . Son expression en fonction de  $\alpha$  s'écrit alors

$$\mathbf{R}(\alpha) = \begin{bmatrix} \cos \alpha(t) & 0 & \sin \alpha(t) \\ 0 & 1 & 0 \\ -\sin \alpha(t) & 0 & \cos \alpha(t) \end{bmatrix}. \quad (6.64)$$

La matrice de la vitesse angulaire  $\boldsymbol{\Omega}_E$  est définie selon l'expression (E.40) par  $\boldsymbol{\Omega}_E = \mathbf{R}^\top \dot{\mathbf{R}}$ . La dérivée temporelle  $\dot{\mathbf{R}}$  de la matrice de rotation est

$$\frac{d\mathbf{R}}{dt} = \dot{\alpha}(t) \begin{bmatrix} -\sin \alpha(t) & 0 & \cos \alpha(t) \\ 0 & 0 & 0 \\ -\cos \alpha(t) & 0 & -\sin \alpha(t) \end{bmatrix}, \quad (6.65)$$

avec  $\dot{\alpha}(t) = \alpha_m \omega_\alpha \cos(\omega_\alpha t)$  selon l'expression (6.62). La matrice de la vitesse angulaire  $\boldsymbol{\Omega}_E$  et le vecteur vitesse angulaire  $\boldsymbol{\omega}_E$  défini à l'équation (E.40) sont donc donnés respectivement par

$$\boldsymbol{\Omega}_E = \dot{\alpha}(t) \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \quad \text{et} \quad \boldsymbol{\omega}_E = \begin{bmatrix} 0 \\ \dot{\alpha}(t) \\ 0 \end{bmatrix}. \quad (6.66)$$

Pour terminer, il est nécessaire de transformer l'expression (6.61) de la vitesse de translation puisque celle-ci est exprimée dans la base  $\mathcal{B}_A$  alors que le modèle d'ordre réduit requiert son expression dans  $\mathcal{B}_E$ . Pour cela, il suffit de projeter l'équation (6.61) sur la matrice de passage  $\mathbf{R}$  au moyen de la relation de changement de base  $\mathbf{s}_{0,E} = \mathbf{R}^\top \mathbf{s}_{0,A}$ , ce qui donne

$$\mathbf{s}_{0,E} = V_\infty \begin{bmatrix} \cos \alpha_\infty \cos \alpha(t) - \sin \alpha_\infty \sin \alpha(t) \\ 0 \\ \cos \alpha_\infty \sin \alpha(t) + \sin \alpha_\infty \cos \alpha(t) \end{bmatrix}. \quad (6.67)$$

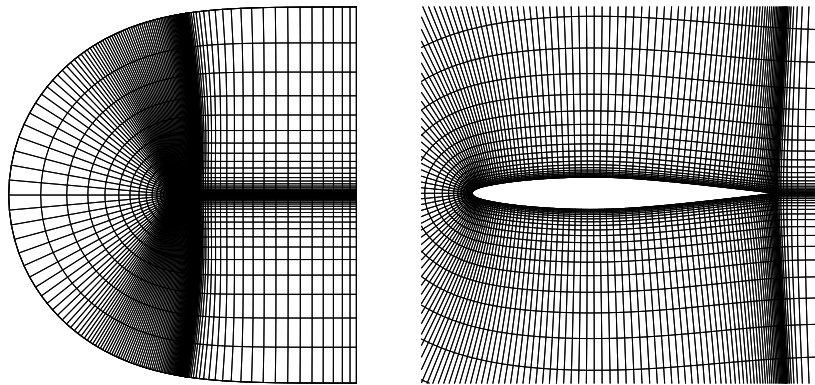
Dans l'exemple traité ici, on supposera que la translation s'effectue horizontalement dans la direction  $x_A$ , si bien que  $\alpha_\infty = 0$ . En outre, l'angle d'incidence moyen  $\alpha_0$  sera aussi supposé nul. Par conséquent, les expressions de la vitesse de translation et de la vitesse angulaire correspondent à

$$\boxed{\mathbf{s}_{0,E} = V_\infty \begin{bmatrix} \cos \alpha(t) \\ 0 \\ \sin \alpha(t) \end{bmatrix} \quad \text{et} \quad \boldsymbol{\omega}_E = \begin{bmatrix} 0 \\ \dot{\alpha}(t) \\ 0 \end{bmatrix} \quad \text{avec} \quad \alpha(t) = \alpha_m \sin(\omega_\alpha t)}. \quad (6.68)$$

L'amplitude maximale des oscillations est fixée à  $\alpha_m = -1,01^\circ$  et la pulsation réduite vaut  $k = 0,202$ . La vitesse à l'infini du fluide est calculée par la relation  $V_\infty = \text{Ma}_\infty c_\infty$ . Le modèle étudié dépend donc principalement de trois paramètres qui définissent complètement les degrés de liberté du mouvement du profil : le nombre Mach de l'écoulement à l'infini  $\text{Ma}_\infty$ , l'amplitude  $\alpha_m$  des oscillations du profil et leur fréquence  $f_\alpha$ .

L'écoulement autour du profil est supposé parfait, compressible et transsonique puisque le nombre de

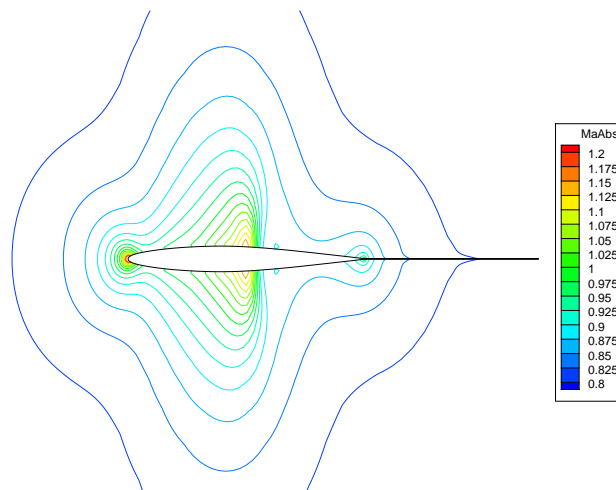




**Figure 6.3 :** Maillage du domaine fluide utilisé pour le calcul de l'écoulement autour du profil NACA0064 rigide en translation uniforme et animé d'un mouvement d'oscillation en tangage.

Mach pour lequel l'écoulement est simulé est  $Ma_\infty = 0,796$ . Par conséquent le champ aérodynamique supportera des ondes de choc dont la position oscille en suivant le mouvement imposé au profil. La résolution numérique repose sur les équations d'Euler formulées dans le repère entraîné avec les composantes de la vitesse absolue. Ce choix se justifie par le fait que l'écoulement est ici modélisé dans un domaine infini pour lequel la formulation en vitesse relative est peu adaptée.

L'exemple traité dans ce chapitre est un cas test développé pour valider la formulation des équations de Navier-Stokes en repère entraîné dans le logiciel *elsA*<sup>4</sup> et par conséquent le modèle numérique est relativement léger. Le maillage 2D en C représenté sur la figure 6.3 comporte en effet seulement  $N = 8\,192$  cellules de contrôle mais cela suffit pour capter les oscillations du choc sur la paroi du profil.

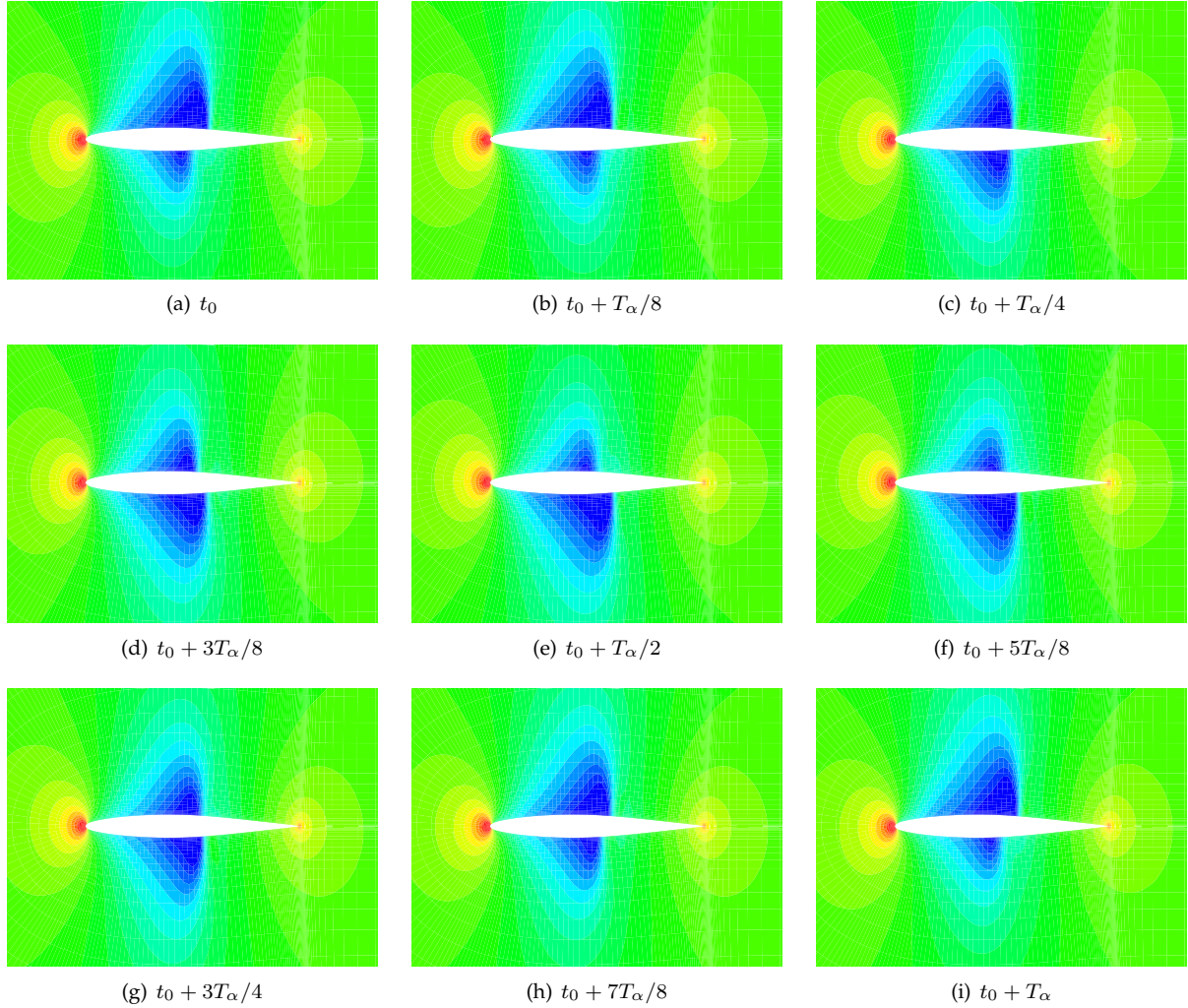


**Figure 6.4 :** Isolignes du nombre de Mach du champ aérodynamique stationnaire autour du profil NACA0064.

La résolution des équations effectuée par *elsA* fournit les champs aérodynamiques pour les variables conservatives. La discrétisation repose sur le schéma de Jameson [120] et, lorsqu'une simulation instationnaire est pratiquée, l'intégration en temps des équations est menée avec un schéma de Runge-Kutta d'ordre 4. Avant d'effectuer une simulation instationnaire en imposant un mouvement au profil, le champ stationnaire de l'écoulement est calculé. Les isolignes de Mach représentées sur la figure 6.4 indiquent clairement la présence d'un choc symétrique. La vitesse de l'écoulement est légèrement subsonique en amont du profil

<sup>4</sup>Les détails du modèle et le script de lancement de la simulation sont disponibles dans la section *naca-rigid* sur la page <http://elsa.onera.fr/elsA/validation/valid.html>.

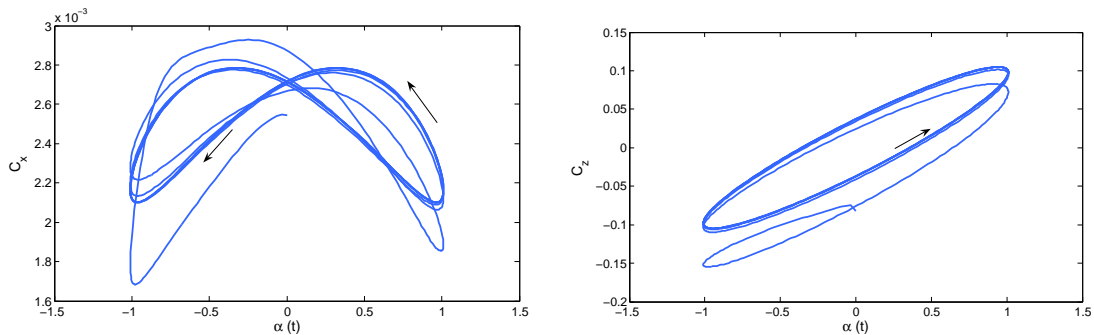
( $Ma_\infty = 0,796$  pour l'écoulement à l'infini) puis des zones de compression se forment et donnent naissance à un choc à environ 54% de la corde du profil. Le caractère symétrique des discontinuités sur l'intrados et l'extrados s'explique par le fait que le profil se déplace à l'horizontale ( $\alpha_\infty = 0$ ) et que l'angle d'incidence moyen est nul ( $\alpha_0 = 0$ ).



**Figure 6.5 :** Clichés du champ de pression sur une période d'oscillation du profil NACA0064. La position des chocs sur l'intrados et l'extrados oscille périodiquement en suivant le mouvement de rotation imposé au profil.

Une simulation instationnaire est ensuite réalisée en imposant au profil le mouvement de corps rigide caractérisé par les vitesses données à l'équation (6.68). Le champ aérodynamique est initialisé avec la solution stationnaire. La mise en mouvement du profil provoque alors une oscillation périodique de la position du choc sur la paroi à la fréquence  $f_\alpha = 1/T_\alpha$ . L'évolution du champ de pression enregistré à différents instants au cours d'une période d'oscillation est donnée sur la figure 6.5. À l'instant  $t_0$ , l'angle d'incidence atteint la valeur maximale  $\alpha(t_0) = \alpha_m$  et le choc sur l'extrados est positionné vers l'arrière du profil à environ 58% de la corde, tandis que le choc sur l'intrados se situe vers le bord d'attaque à environ 50% de la corde. Lorsque l'angle d'incidence diminue le choc de l'extrados se déplace vers le bord d'attaque alors que le choc sur l'intrados suit la tendance inverse. À  $t = t_0 + T_\alpha/4$ , l'angle d'incidence devient nul et le profil est donc orienté à l'horizontale ; le champ aérodynamique se confond alors avec le champ stationnaire puisque la configuration géométrique est identique. Le processus se poursuit et le choc sur l'intrados atteint sa position maximale à  $t = t_0 + T_\alpha/2$ . Le même processus se répète ensuite en sens inverse au cours de la demi-période restante jusqu'à retrouver le même champ aérodynamique à  $t_0 + T_\alpha$ .

Enfin, les graphiques de la figure 6.6 montrent les courbes de Lissajous des coefficients aérodynamiques  $C_x$  et  $C_z$ . Sur ces courbes, les efforts aérodynamiques dans les directions  $x_E$  et  $z_E$  sont représentés en fonction de l'excitation harmonique paramétrée par l'angle d'incidence  $\alpha(t)$  défini à l'équation (6.68). On peut noter pour commencer que les forces sont principalement orientées dans la direction verticale, ce qui s'explique d'une part puisque l'amplitude maximale de l'angle d'incidence est très faible et d'autre part parce que le fluide est supposé parfait. Ces courbes permettent par ailleurs d'interpréter les échanges d'énergie qui s'effectuent entre le fluide et la structure en fonction du sens dans lequel est parcouru le cycle. C'est la raison pour laquelle le sens de rotation est indiqué à chaque fois par des flèches. Lorsque le cycle est parcouru dans le sens trigonométrique, le système tend à être amorti si l'excitation n'est pas entretenue puisque les efforts aérodynamiques sont en retard sur l'excitation. C'est le cas notamment pour les efforts dans la direction  $z_E$ , ce qui est logique puisque dans l'exemple traité ici le mouvement est imposé à la structure qui est donc à l'origine des efforts aérodynamiques. À l'inverse, si le cycle est parcouru dans le sens horaire, les efforts aérodynamiques sont en avance sur l'excitation ce qui signifie que les efforts aérodynamiques peuvent provoquer la mise en mouvement de la structure. Par conséquent une amplification du mouvement conduisant à la divergence est possible. La courbe de Lissajous du  $C_x$  est plus complexe : elle laisse apparaître deux lobes qui sont parcourus chacun dans un sens différent. Lorsque l'angle d'incidence est positif, les forces tendent à être amorties puisque le cycle est parcouru dans le sens trigonométrique, tandis que lorsque l'angle d'incidence devient négatif, une amplification des forces est possible.



**Figure 6.6 :** Courbes de Lissajous des coefficients aérodynamiques  $C_x$  et  $C_z$  associés au mouvement d'oscillation harmonique imposé au profil NACA0064 en tangage.

Les résultats de cette simulation numérique seront exploités par la suite pour former la base des modes POD nécessaire pour la construction des modèles d'ordre réduit. Avant de présenter les résultats des différents modèles d'ordre réduit, on précise dans le paragraphe suivant les expressions des coefficients autonomes pour les équations d'Euler puis des coefficients non-autonomes pour le mouvement particulier traité ici.

#### 6.4.2 Détail des coefficients du modèle d'ordre réduit pour les équations d'Euler dans un domaine en translation uniforme et en rotation simple

Au paragraphe 6.3.1, un modèle d'ordre réduit POD-Galerkin des équations de Navier-Stokes formulées dans le repère entraîné avec les composantes de la vitesse absolue avait été développé. Ce modèle d'ordre réduit est légèrement simplifié dans le cas présent puisque l'écoulement étudié est modélisé par les équations d'Euler. Celles-ci se déduisent des équations de Navier-Stokes en éliminant les termes diffusifs associés aux processus visqueux. Les équations du fluide conservent la forme d'un système polynomial quadratique identique à celui de l'équation (6.12), avec les termes diffusifs  $\mathbf{Q}^D(\mathbf{q}, \mathbf{q})$  en moins. Les coefficients convectifs  $\mathbf{Q}^C$  sont encore donnés par l'équation (6.13). L'expression (6.14) du modèle d'ordre réduit est donc toujours valable, mais les coefficients autonomes  $\mathcal{K}_i$ ,  $\mathcal{L}_{ij}$  et  $\mathcal{Q}_{ijk}$  sont à présent définis par

$$\begin{aligned}
 \mathcal{K}_i &= \mathcal{K}_i^C = \left\langle \mathbf{Q}^C(\bar{\mathbf{q}}, \bar{\mathbf{q}}), \varphi^{(i)} \right\rangle, \\
 \mathcal{L}_{ij} &= \mathcal{L}_{ij}^C = \left\langle \mathbf{Q}^C(\bar{\mathbf{q}}, \varphi^{(j)}) + \mathbf{Q}^C(\varphi^{(j)}, \bar{\mathbf{q}}), \varphi^{(i)} \right\rangle, \\
 \mathcal{Q}_{ijk} &= \mathcal{Q}_{ijk}^C = \left\langle \mathbf{Q}^C(\varphi^{(j)}, \varphi^{(k)}), \varphi^{(i)} \right\rangle.
 \end{aligned} \tag{6.69}$$

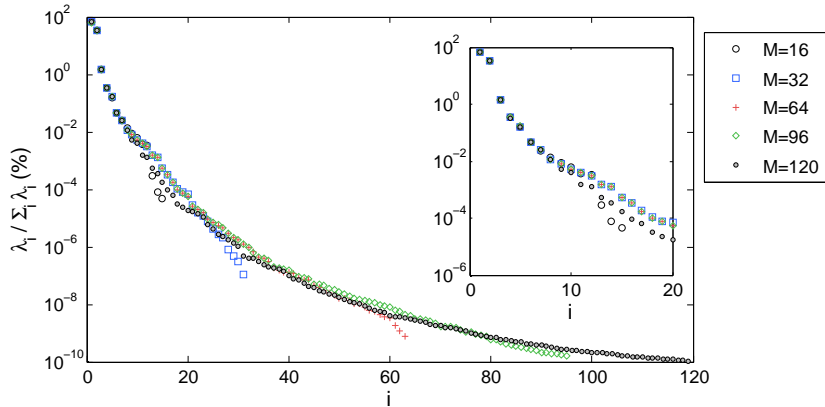
Les coefficients non-autonomes  $\mathcal{K}_i^E(t)$  et  $\mathcal{L}_{ij}^E(t)$  décomposés ensuite aux équations (6.28) et (6.27) peuvent aussi être simplifiés puisque le mouvement consiste dans cet exemple en une translation à vitesse uniforme et une rotation autour de l'axe  $y_E$ . En effet, les expressions de la vitesse de translation et de la vitesse angulaire fournies à l'équation (6.68) au début de ce paragraphe indiquent que seules trois composantes sont non nulles sur les six permettant de décrire le mouvement des deux degrés de liberté. Par conséquent, les produits scalaires dans les expressions (6.28) et (6.27) n'impliquent que ces trois composantes et les coefficients non-autonomes s'écrivent donc

$$\mathcal{K}_i^{ET} = \begin{bmatrix} \mathcal{K}_{i,x}^{ET} \\ 0 \\ \mathcal{K}_{i,z}^{ET} \end{bmatrix}, \quad \mathcal{L}_{ij}^{ET} = \begin{bmatrix} \mathcal{L}_{ij,x}^{ET} \\ 0 \\ \mathcal{L}_{ij,z}^{ET} \end{bmatrix}, \quad \mathcal{K}_i^{ER} = \begin{bmatrix} 0 \\ \mathcal{K}_{i,y}^{ER} \\ 0 \end{bmatrix} \quad \text{et} \quad \mathcal{L}_{ij}^{ER} = \begin{bmatrix} 0 \\ \mathcal{L}_{ij,y}^{ER} \\ 0 \end{bmatrix}. \tag{6.70}$$

Le nombre de coefficients non-autonomes à évaluer n'est donc plus  $N_c^{\text{na}} = 6q + 6q^2$  mais seulement  $N_c^{\text{na}} = 3q + 3q^2$  : il y a  $2q$  coefficients constants d'entraînement en translation à calculer et  $q$  coefficients pour la rotation, puis  $2q^2$  pour les coefficients linéaires d'entraînement en translation et  $q^2$  pour la rotation.

### 6.4.3 Calcul de la base des modes POD de l'écoulement autour du profil NACA0064 rigide en tangage

Dans un premier temps, un ensemble de clichés est extrait de la simulation instationnaire effectuée avec *elsA* sur l'intervalle de temps  $I_e = [t_0; t_0 + T_e]$ . La base des modes POD est alors calculée avec la méthode des clichés pour un ensemble de  $M = 16, 32, 64, 96$  et  $192$  clichés uniformément répartis sur une durée d'échantillonnage  $T_e \approx n_p T_\alpha$  couvrant environ une période et demi du mouvement d'oscillation forcé ( $n_p \approx 1, 5$ ). Un deuxième ensemble de  $M = 120$  clichés est sélectionné sur  $n_p \approx 2, 5$  périodes d'oscillation afin de vérifier si le spectre des valeurs propres reste proche lorsque la durée d'échantillonnage est allongée. Le résultat est représenté sur la figure 6.7 où le pourcentage d'énergie captée par chaque valeur propre est tracé pour différentes valeurs du nombre de clichés.

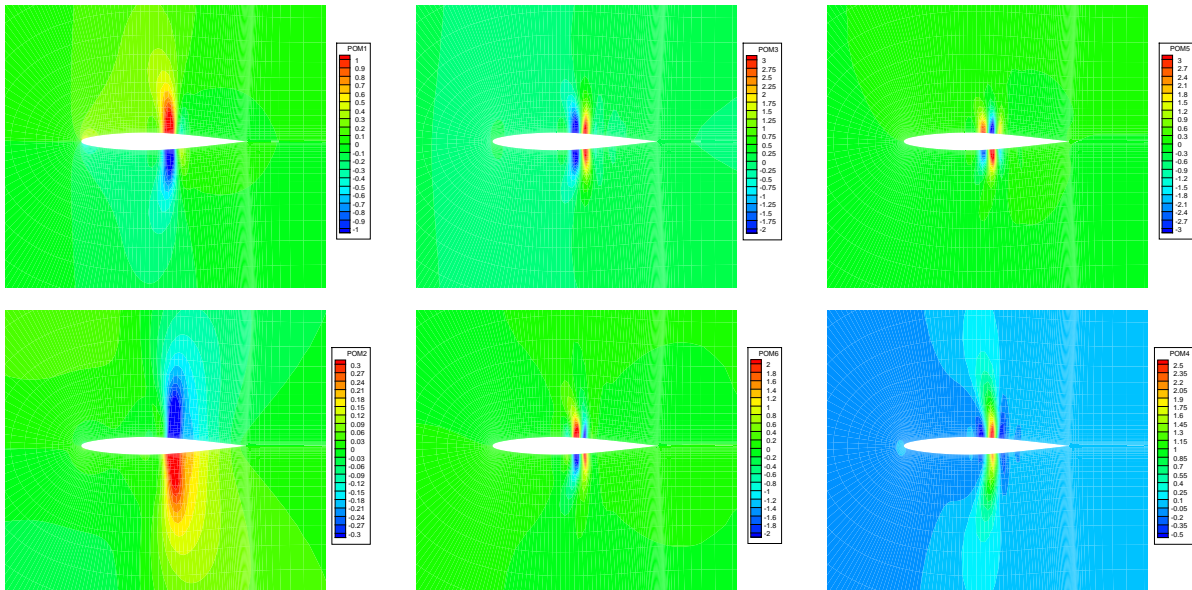


**Figure 6.7 :** Spectre des valeurs propres associées aux modes POD aérodynamiques de l'écoulement autour du profil NACA0064 animé d'un mouvement de corps rigide. Le spectre est tracé pour différentes valeurs  $M$  du nombre de clichés échantillonnés sur  $n_p \approx 1, 5$  période, sauf pour  $M = 120$  où les clichés sont prélevés sur  $n_p \approx 2, 5$  périodes. L'encadré correspond au zoom du spectre sur les 20 premières valeurs propres.

La durée d'échantillonnage conduit à des spectres relativement proches lorsque  $n_p \approx 1, 5$  et  $2, 5$  et que le nombre de clichés employés est du même ordre de grandeur. Ceci s'explique par le fait que l'écoulement

considéré est encore une fois périodique et que l'allongement de la durée d'échantillonnage n'apporte pas d'information supplémentaire au delà d'une période complète d'oscillation  $T_\alpha$ , comme on l'avait déjà observé dans l'annexe C.1. Le nombre  $M$  de clichés influe toutefois légèrement sur la forme du spectre, en particulier sur les dernières valeurs propres. Au-delà de  $M = 32$  clichés, les 20 premières valeurs propres sont pratiquement confondues et par conséquent on prendra pour la suite des ensembles de clichés comportant au moins  $M = 32$  éléments. Dans cet exemple, le spectre décroît de manière moins significative que pour l'écoulement autour du profil NACA0012 en forte incidence traité au chapitre 5. En particulier, on n'observe pas de coupure nette dans la pente du spectre, qui permettrait de décider du nombre de modes POD minimal à conserver dans la base de projection. Toutefois, le zoom sur les 20 premières valeurs propres tracé sur la figure 6.7 met en évidence une légère discontinuité de la pente au niveau de la 12<sup>e</sup> valeur propre. Pour la suite, on peut considérer qu'une dizaine de modes POD devrait encore suffire pour reproduire correctement l'écoulement puisque le pourcentage d'énergie captée par les  $q = 10$  premiers modes POD lorsque le nombre de clichés employés est  $M = 64$  s'élève à  $\eta_q = 99,99\%$ .

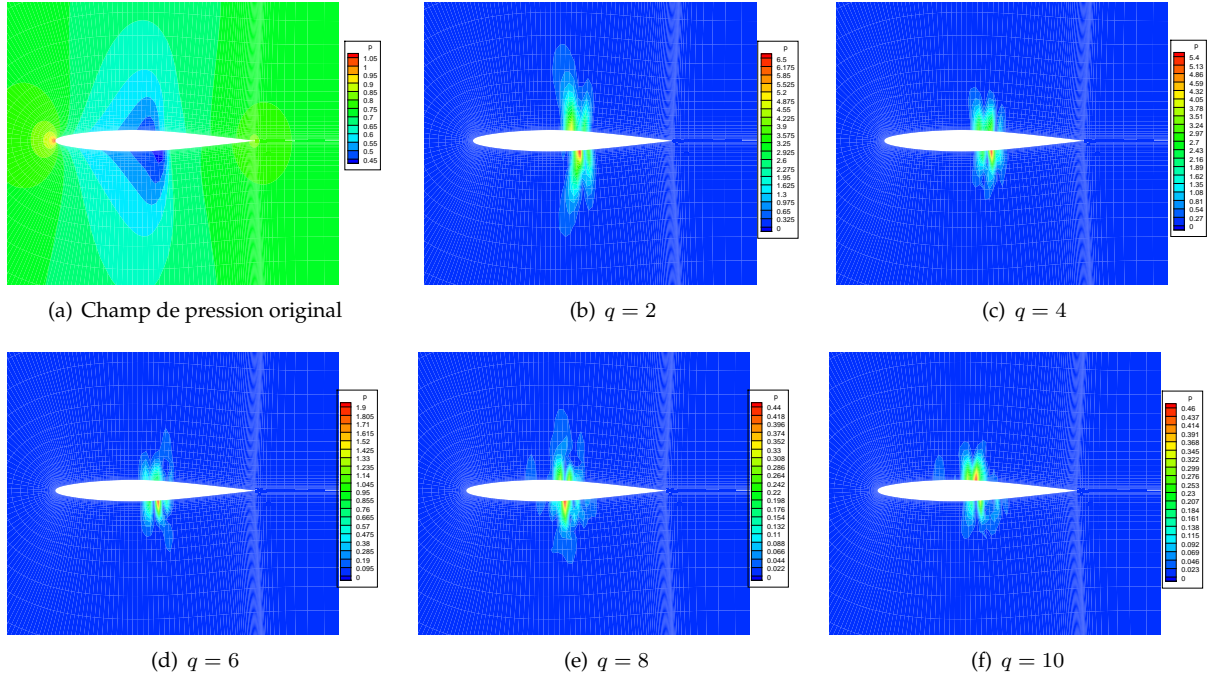
La figure 6.8 donne l'allure de la contribution en pression  $\varphi_p^{(j)}$  des six premiers modes POD de l'écoulement autour du profil NACA0064. Comme pour les modes POD de l'écoulement autour du profil NACA0012, les modes POD peuvent être regroupés par paires mais le classement ne s'effectue pas directement en fonction de l'indice des modes : sur les six premiers modes, on peut distinguer les paires (1-2), (3-6) et (4-5). Les modes obtenus font clairement apparaître des discontinuités représentatives du choc généré sur l'intrados et l'extrados. Les deux premiers modes POD mettent en évidence une seule zone de discontinuité asymétrique qui s'étend de part et d'autre du profil avec un signe opposé côté intrados et côté extrados. Les modes POD suivants font apparaître des structures de taille inférieure présentant encore des discontinuités. Les paires de modes (3-6) et (4-5) font apparaître respectivement deux et trois zones de discontinuité où les signes alternent. Les modes suivants peuvent encore être regroupés mais l'association des modes entre eux devient de plus en plus difficile puisque les structures ont des tailles de plus en plus faibles et relativement proches les unes des autres.



**Figure 6.8 :** Allure des six premiers modes POD de pression calculés à partir de  $M = 64$  clichés répartis uniformément sur environ une période et demi d'oscillation du mouvement harmonique imposé au profil NACA0064 en tangage. Les modes sont regroupés par paires dans chaque colonne en fonction des structures mises en évidence.

Le phénomène d'oscillation du choc doit pouvoir être reproduit à l'aide de la base POD en conservant assez de modes pour que la position de l'onde de choc s'écrive correctement comme une combinaison linéaire des modes. Dans un premier temps, on cherche à vérifier ceci à un instant donné, choisi de manière arbitraire comme l'instant final  $t_M = t_0 + T_e$  de la fenêtre d'échantillonnage. Le champ de pression correspondant est représenté sur la figure 6.9(a) à titre indicatif. Le champ de pression  $p_r(t_M) = \sum_{j=1}^q a_j^{(M)} \varphi_p^{(j)}$

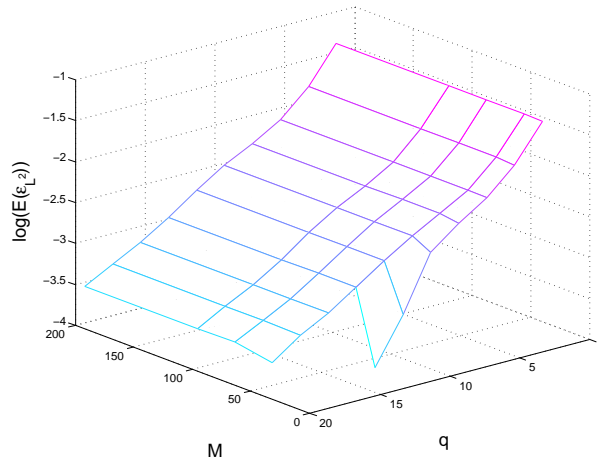
est ensuite reconstruit à l'aide des amplitudes modales  $a_j^{(M)}$  obtenues par projection du champ de pression sur une base de  $q$  modes POD  $\varphi_p^{(j)}$  de la base de projection. L'erreur relative de reconstruction commise en conservant un nombre croissant de modes POD dans la base de projection est représentée sur les figures 6.9(b) à (f). On constate bien une diminution de l'erreur lorsque le nombre  $q$  de modes POD conservés dans la base de projection augmente. L'erreur relative maximale qui est de l'ordre de 6,5% lorsque seuls deux modes POD sont conservés dans la base de projection chute à seulement 0,46% lorsque la base POD est constituée de 10 modes. Comme pour l'exemple du profil NACA0012 en forte incidence, les zones où l'erreur est importante correspondent bien aux modes qui ont été tronqués dans la base de projection et sont par conséquent atténuées lorsque les modes POD suivants sont introduits dans la base de projection.



**Figure 6.9 :** La figure (a) représente le champ de pression à l'instant  $t_M = t_0 + T_e$  qui est ensuite reconstruit au moyen de la POD. L'erreur relative de reconstruction entre le champ de pression représenté en (a) et le champ de pression reconstruit avec un nombre croissant  $q$  de modes POD dans la base de projection est représentée sur les figures (b) à (f).

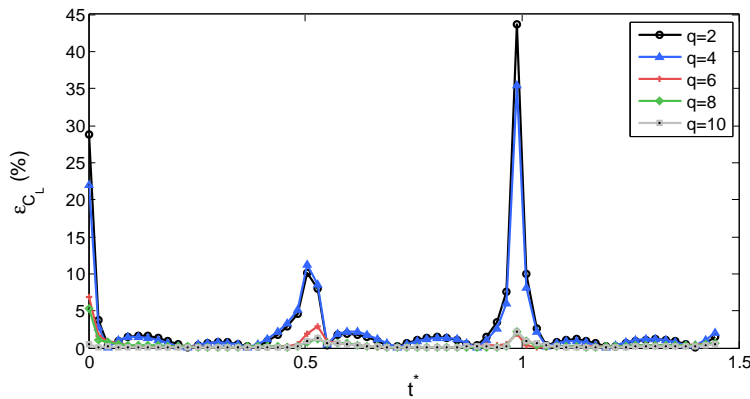
Les erreurs représentées sur les figures précédentes sont établies pour un instant donné au cours de la fenêtre d'échantillonnage et fournissent une description spatiale des zones où se concentrent les défauts de reconstruction. D'une manière plus générale, on définit une erreur de reconstruction spatio-temporelle globale pour chaque couple de paramètres  $(M, q)$  : l'erreur spatiale globale à chaque instant  $t_m$  est définie par la norme  $L^2$  avec la relation (5.51), puis l'erreur moyenne sur la fenêtre d'échantillonnage est obtenue par application de l'opérateur de moyenne temporelle  $E[\cdot]$  à l'erreur précédente. Le graphique de la figure 6.10 résume le comportement de l'erreur en fonction du nombre  $M$  de clichés et du nombre  $q$  de modes POD conservés dans la base de projection. L'erreur se comporte de la même manière que pour le profil NACA0012 : on observe une décroissance monotone de l'erreur lorsque le nombre de modes POD augmente, mais le nombre de clichés prélevés lors de l'échantillonnage n'a pratiquement pas d'influence sur l'erreur de reconstruction globale. Il faut cependant remarquer que l'erreur décroît cette fois de façon moins accentuée que pour le profil NACA0012. Ceci illustre le fait que l'écoulement est plus complexe et corrobore l'allure du spectre des valeurs propres tracé sur la figure 6.7 pour lequel on avait signalé que la pente était moins prononcée que pour l'exemple du profil NACA0012.

Bien que l'écoulement semble cette fois plus complexe, il semble d'après les résultats précédents que la réponse du système puisse être reproduite — au moins globalement — avec un faible nombre de modes POD puisque les premiers modes permettent déjà de représenter le phénomène d'oscillation du choc. Si l'objectif était de reproduire de façon très précise la position du choc, un nombre important de modes serait



**Figure 6.10 :** Convergence de l'erreur spatio-temporelle globale  $E [\varepsilon_{L^2, q}]$  en fonction du nombre  $M$  de clichés et du nombre  $q$  de modes POD contenus dans la base de projection pour l'écoulement autour du profil NACA0064. L'erreur est tracée en échelle logarithmique pour mettre en évidence la décroissance.

certainement nécessaire, mais ici, c'est principalement la résultante des efforts sur la paroi du profil qui importe puisque le solide est supposé rigide. C'est la raison pour laquelle on se contentera par la suite d'une dizaine de modes POD dans la base de projection. Le graphique de la figure 6.11 permet de vérifier que l'erreur relative sur le coefficient de portance est très faible sur la fenêtre d'échantillonnage, dès lors qu'une dizaine de modes POD a été conservée dans la base de projection.



**Figure 6.11 :** Erreur relative de reconstruction du coefficient de portance du profil NACA0064 oscillant en fonction du nombre  $q$  de modes POD dans la base de projection. L'erreur relative est tracée en fonction du temps rendu adimensionnel avec la période d'oscillation  $T_\alpha$  pour chaque cliché de la fenêtre d'échantillonnage.

Dans ce premier paragraphe, l'allure des modes POD a été présentée et les structures spatiales mises en évidence permettent de caractériser l'oscillation des chocs sur la paroi du profil. Les erreurs de reconstruction calculées à l'aide des amplitudes modales issues de la projection des modes POD sur les clichés ont permis de montrer qu'un faible nombre de modes POD devrait ici encore suffire pour reproduire l'écoulement, bien que la décroissance du spectre des valeurs propres indique que celui-ci semble plus riche. Avec seulement une dizaine de modes POD, le coefficient de portance est très bien reproduit et par conséquent on peut espérer que la réponse produite par l'intégration en temps du modèle d'ordre réduit sera acceptable.

## 6.5 Réponse du modèle d'ordre réduit de l'écoulement autour du profil NACA0064

Le modèle d'ordre réduit POD-Galerkin est à présent employé afin de reproduire le champ aérodynamique autour du profil, et en particulier le coefficient de portance généré. Les équations (6.14) sont donc intégrées en temps et les amplitudes modales  $a_i(t)$  obtenues sont comparées aux amplitudes de référence  $a_i^{(m)}$  issues de la projection sur les modes POD  $\varphi^{(i)}$  des clichés connus aux instants  $t_m$ . Le modèle d'ordre réduit souffre cependant d'un manque de stabilité encore plus prononcé que dans l'exemple du profil NACA0012 traité au chapitre 5. En effet, l'intégration en temps conduit à une réponse divergente bien avant d'atteindre la fin de la fenêtre d'échantillonnage avec  $M = 64$  clichés échantillonnés sur  $n_p \approx 1,5$  périodes et  $q = 10$  modes POD. L'augmentation du nombre  $q$  de modes POD compris dans l'intervalle  $[[1; 20]]$  n'y change rien. Aucune amélioration significative n'est non plus obtenue avec l'augmentation du nombre de clichés employés pour calculer les modes POD et qui s'élève à  $M = 192$  ou  $M = 120$  avec dans ce dernier cas un échantillonnage sur  $n_p \approx 2,5$  périodes.

L'extrême instabilité du modèle d'ordre réduit POD-Galerkin peut s'expliquer par le fait que le fluide a été supposé parfait et qu'il n'existe donc dans les équations continues aucune source de dissipation permettant de stabiliser la réponse du modèle d'ordre réduit. Dans la pratique, la résolution des équations d'Euler ou de Navier-Stokes pour les écoulements compressibles requiert l'introduction d'une certaine quantité de dissipation artificielle afin de stabiliser la réponse (voir l'annexe E.3.1.2 p.264). Cette viscosité supplémentaire associée à la discrétisation des équations est d'origine purement numérique [24, 157] mais le système d'équations la gouvernant ne se prête pas à une projection sur une base modale puisqu'il ne s'agit pas d'un système d'équations aux dérivées partielles polynomiales. Aucune modélisation de cette quantité (artificielle) n'a donc été introduite dans le modèle d'ordre réduit. Par ailleurs, l'hypothèse d'une dissipation (physique) constante et égale à la viscosité du fluide au repos qui s'était révélée acceptable dans l'exemple de l'écoulement visqueux et faiblement compressible autour du profil NACA0012 n'a plus lieu d'être ici puisque le fluide est parfait. Par conséquent, lorsque les équations d'Euler sont employées pour un écoulement transsonique tel que celui traité dans cette partie, l'absence totale de dissipation semble compromettre la stabilité du modèle d'ordre réduit, même à très court terme. Il devient donc absolument indispensable de corriger le modèle d'ordre réduit au moyen de techniques similaires à celles mises en œuvre dans le chapitre 5 pour tenter d'améliorer la réponse.

### 6.5.1 Réponse du modèle d'ordre réduit POD-Galerkin corrigé

#### 6.5.1.1 Correction de Cazemier *et al.* du modèle d'ordre réduit POD-Galerkin

On considère tout d'abord la méthode de correction de Cazemier *et al.* [42] à l'aide de laquelle un opérateur d'amortissement linéaire diagonal est ajouté dans les équations du modèle d'ordre réduit. Le principe rappelé au paragraphe 5.3.2.2 consiste à déterminer les coefficients d'amortissement de manière à ce que la moyenne du taux de variation de l'énergie dans la direction de chaque mode POD soit conservée. Dans le cas d'un système dynamique autonome, les coefficients d'amortissement sont donnés par la relation (5.33) mais il convient de modifier ici l'expression pour tenir compte des termes non-autonomes associés au mouvement du domaine. En suivant le même procédé que pour le système autonome, les coefficients d'amortissement s'écrivent

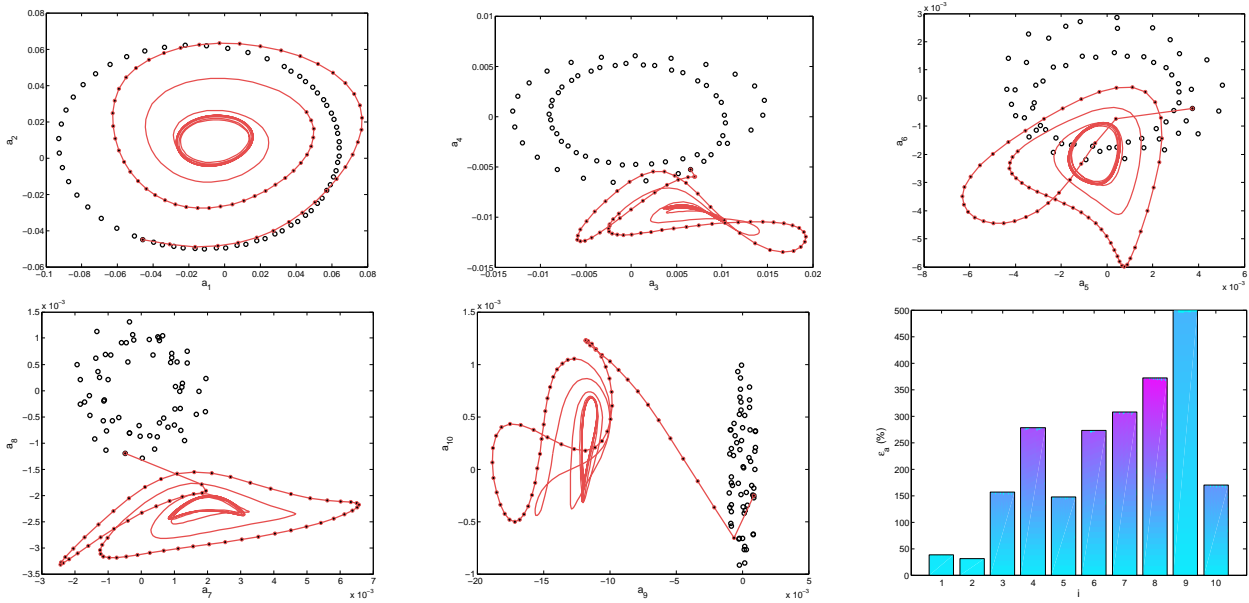
$$\mathcal{D}_{ii} = -\frac{1}{\lambda_i} \sum_{j,k=1}^q \mathcal{Q}_{ijk} E(a_i a_j a_k) - \mathcal{L}_{ii} - \frac{1}{\lambda_i} \left[ E(\mathcal{K}_i^E(t_m) a_i) + \sum_{j=1}^q E(\mathcal{L}_{ij}^E(t_m) a_i a_j) \right]. \quad (6.71)$$

L'expression des termes de correction fait à présent apparaître les contributions  $\mathcal{K}_i^E(t_m)$  et  $\mathcal{L}_{ij}^E(t_m)$  relatives au mouvement imposé. Ces contributions instationnaires sont évaluées au moyen des relations (6.28) et (6.27) à chaque instant  $t_m$  où les clichés sont extraits.

Utilisée telle quelle, la correction précédente ne permet pas d'améliorer la réponse du modèle d'ordre réduit de manière significative avec  $M = 64$  et  $q = 10$ . Pour améliorer significativement la stabilité à long terme, on peut forcer artificiellement l'amortissement en multipliant les coefficients de la correction de Cazemier *et al.* par un facteur de pondération. Ce facteur peut être identique pour chaque coefficient de



la diagonale de l'opérateur  $\mathcal{D}$  mais il est préférable de fournir un facteur différent pour chaque coefficient. Dans les deux cas, la réponse peut être stabilisée, au moins sur la durée d'échantillonnage des clichés.

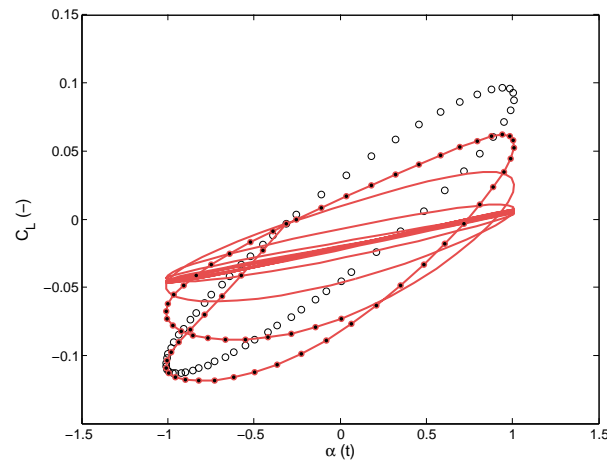


**Figure 6.12 :** Comparaison des portraits de phase des amplitudes modales calculées par intégration du modèle d'ordre réduit corrigé par la méthode de Cazemier et al. sur la fenêtre d'échantillonnage des clichés ( $\bullet$ ) et des amplitudes modales de référence issues de la projection des clichés sur les modes POD ( $\circ$ ). La réponse du modèle d'ordre réduit à long terme tracée en traits plein ( $\text{—}$ ) montre la convergence de la solution vers un cycle limite. Le dernier graphique représente l'erreur  $\bar{\varepsilon}_{\alpha_j}$  en norme  $L^2$  sur l'intervalle de temps  $I_e$  par rapport aux amplitudes modales de référence.

La réponse obtenue est illustrée par les portraits de phase des amplitudes modales tracés sur la figure 6.12 lorsque les termes d'amortissement introduits sont pondérés par un facteur variant linéairement en fonction de l'indice  $i$  correspondant à l'équation gouvernant l'amplitude  $a_i$ . Sur l'intervalle de temps  $I_e = [t_0; t_0 + T_e]$ , les erreurs sont relativement importantes : l'erreur globale sur les amplitudes modales définie à l'équation (5.53) atteint  $\varepsilon_A = 45,0\%$ . Le détail des erreurs pour chaque amplitude modale est donné sur le dernier graphique de la figure 6.12. On observe en particulier que la réponse de la neuvième amplitude est complètement erronée. Les erreurs pour les premières amplitudes modales restent quant à elles raisonnables, si bien que la réponse pour des grandeurs macroscopiques telles que le coefficient de portance  $C_L$  est passable. Ainsi, l'erreur moyenne en temps  $\bar{\varepsilon}_{C_L}$  au sens de la norme  $L^2$  définie par (5.52) s'élève à  $\bar{\varepsilon}_{C_L} = 33,7\%$ .

Le portrait de phase du coefficient de portance  $C_L$  est tracé sur la figure 6.13 en fonction de l'angle d'incidence  $\alpha(t)$  imposé au profil. Sur l'intervalle  $I_e$ , le cycle limite n'est pas exactement reproduit mais l'orientation et l'amplitude du cycle sont qualitativement satisfaisants. Toutefois, lorsque l'intégration en temps du modèle d'ordre réduit est poursuivie, les amplitudes modales sont beaucoup trop amorties et le coefficient de portance converge vers un cycle limite beaucoup trop incliné et aplati. Le comportement à long terme du modèle d'ordre réduit est évalué au moyen de diverses erreurs. La première porte sur la fréquence d'oscillation  $\varepsilon_{f_\alpha}$  qui est relativement faible ici :  $\varepsilon_{f_\alpha} = 0,40\%$ . Ensuite, on définit des erreurs relatives pour le maximum et le minimum atteints par le coefficient de portance ( $\varepsilon_{C_L, \max} = 53,7\%$  et  $\varepsilon_{C_L, \min} = 88,9\%$ ) ainsi que l'erreur sur l'amplitude crête-à-crête ( $\varepsilon_{\Delta_{cc}} = 69,9\%$ ). Finalement l'évolution à long terme des oscillations est caractérisée par le décrément logarithmique moyen qui vaut ici  $\delta_{\log} = 0,111$ . Celui-ci est calculé à partir de la moyenne des décréments calculés pour les maxima du coefficient de portance sur les dernières oscillations par la relation

$$\delta_{\log} = \frac{1}{n_{\max}} \sum_{i=1}^{n_{\max}} \delta_{\log}(i) \quad \text{avec} \quad \delta_{\log}(i) = \log_{10} \frac{C_{L, \max}(i)}{C_{L, \max}(i+1)}. \quad (6.72)$$



**Figure 6.13** : Comparaison sur l'intervalle de temps  $I_e$  du portait de phase du coefficient de portance ( $\bullet$ ) calculé avec le modèle d'ordre réduit POD-Galerkin corrigé par la méthode de Cazemier *et al.* avec le coefficient de portance de référence ( $\circ$ ). La trajectoire du coefficient de portance à long terme tracée en trait plein ( $\text{—}$ ) converge vers un cycle limite erroné.

Le décrément caractérise la tendance de la réponse à la divergence ou à l'amortissement. Dans le cas présent, le décrément logarithmique doit être nul puisque le mouvement d'oscillation est imposé au profil, et par conséquent les oscillations sont nécessairement entretenues. Au contraire, une valeur négative indique une tendance à la divergence de la réponse, tandis qu'une valeur positive est le signe d'un amortissement conduisant éventuellement à la disparition du mouvement.

En conclusion, la correction de Cazemier *et al.* permet de stabiliser le modèle d'ordre réduit en introduisant un surplus d'amortissement, mais le comportement à long terme du modèle d'ordre réduit n'est pas satisfaisant. Il semble que l'introduction d'un amortissement linéaire diagonal ne soit pas suffisante, puisque la réponse du modèle d'ordre réduit est soit divergente, soit trop amortie en fonction de la valeur attribuée au facteur multiplicatif introduit pour pondérer la correction.

La correction n'est pas plus efficace pour d'autres valeurs de  $M$  et  $q$ . En particulier, lorsqu'un nombre plus important de modes POD est conservé dans la base de projection, la valeur du facteur multiplicatif doit aussi être augmenté pour éviter la divergence. Dans ce qui suit, on a donc recours aux méthodes de calibrage présentées au paragraphe 5.3.3.1 afin que la correction introduite puisse agir sur l'ensemble des termes du modèle d'ordre réduit et non plus seulement sur l'amortissement linéaire diagonal.

### 6.5.1.2 Calibrage des coefficients du modèle d'ordre réduit POD-Galerkin

Les résultats du paragraphe 5.5.3 concernant le calibrage du modèle d'ordre réduit pour l'exemple du profil NACA0012 conduisent à retenir la formulation de Floquet pour la fonctionnelle à minimiser et à employer l'algorithme de Levenberg-Marquardt pour chercher les coefficients optimaux.

Plusieurs difficultés apparaissent avec ce nouvel exemple. La première provient du fait que le modèle d'ordre réduit est extrêmement instable. Par conséquent, il n'est pas possible d'évaluer correctement la fonctionnelle à minimiser puisque celle-ci requiert les amplitudes modales calculées sur la durée d'échantillonnage  $I_e$ . Or au cours de l'intégration en temps du modèle d'ordre réduit non corrigé, la solution diverge bien avant d'atteindre l'instant final  $t_0 + T_e$ . La solution adoptée ici consiste donc à partir du modèle d'ordre réduit déjà corrigé avec la méthode de Cazemier *et al.* [42] de façon à disposer des amplitudes modales sur l'ensemble de l'intervalle de temps d'échantillonnage. Lorsque les coefficients à calibrer sont corrigés progressivement au fil des itérations de l'algorithme d'optimisation, il est possible que l'intégration en temps du modèle d'ordre réduit diverge encore, empêchant ainsi d'évaluer la fonctionnelle sur l'intervalle de temps  $I_e$  complet. L'optimisation est toutefois poursuivie en évaluant la fonctionnelle sur l'intervalle de temps sur lequel la simulation est stable, et en introduisant un facteur de pénalité important pour signaler la divergence avec le nouveau jeu de paramètres proposé par l'algorithme d'optimisation. La seconde difficulté concerne le choix des paramètres à calibrer. Dans l'exemple du profil NACA0012, un

choix naturel consistait à calibrer le nombre de Reynolds en le modifiant pour chaque équation du modèle d'ordre réduit, de manière à introduire le taux de dissipation adéquat. Dans le cas présent, puisque le fluide est supposé parfait, cette solution est inapplicable. Il faut donc modifier directement certains coefficients du modèle d'ordre réduit, de manière à stabiliser la réponse à long terme. Le choix des coefficients est rendu encore plus difficile par le fait que le modèle d'ordre réduit contient cette fois  $N_c^{\text{na}} = 6q + 6q^2$  coefficients supplémentaires liés au mouvement imposé au domaine.

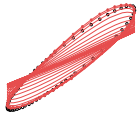
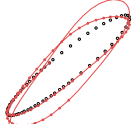
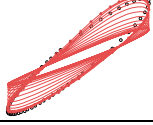
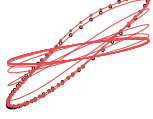
Identifiant	Coefficients	$p_c$
KdL	$\mathcal{K}_i, \mathcal{L}_{i,i}$	$2q$
K3dL	$\mathcal{K}_i, \mathcal{L}_{i,i}, \mathcal{L}_{i,i+1}, \mathcal{L}_{i,i-1}$	$2q + 2(q - 1)$
KdL_E	$\mathcal{K}_i, \mathcal{L}_{i,i}, s_{0,E,i}, \omega_{E,i}$	$2q + 6$
K3dL_E	$\mathcal{K}_i, \mathcal{L}_{i,i}, \mathcal{L}_{i,i+1}, \mathcal{L}_{i,i-1}, s_{0,E,i}, \omega_{E,i}$	$2q + 2(q - 1) + 6$

**Tableau 6.2 :** Résumé des jeux de coefficients calibrés par minimisation de la fonctionnelle d'erreur de la formulation de Floquet. La première colonne indique le nom du jeu de paramètres, la seconde les coefficients calibrés et la dernière le nombre  $p_c$  de coefficients calibrés.

Le tableau 6.2 résume les différents jeux de paramètres qui sont regroupés dans le vecteur  $\mathbf{w} \in \mathbb{R}^{p_c}$  intervenant dans la définition de la fonctionnelle du problème d'optimisation. Les deux premiers jeux de paramètres KdL et K3dL permettent de calibrer une partie des coefficients autonomes du modèle d'ordre réduit. L'ensemble des coefficients constants est corrigé pour pallier les erreurs relatives à la partie moyenne de la réponse du système. Avec la correction KdL, seuls les coefficients linéaires de la diagonale principale sont modifiés. Pour améliorer la reproduction de l'amortissement, la correction K3dL a été introduite pour calibrer les termes de couplage linéaires intervenant sur les deux diagonales voisines de la diagonale principale dans l'opérateur linéaire  $\mathcal{L}$ . Enfin, les jeux de paramètres KdL\_E et K3dL\_E ont été introduits afin de corriger les termes sources non-autonomes. Plutôt que de corriger directement chacun des  $N_c^{\text{na}}$  coefficients supplémentaires, on a choisi ici d'agir sur le vecteur vitesse de translation  $s_{0,E}$  et le vecteur vitesse angulaire  $\omega_E$ , ce qui permet de réduire à 6 le nombre de paramètres. Pour le mouvement particulier imposé, ce dernier type de calibrage n'introduit en réalité que 3 paramètres supplémentaires puisque le mouvement s'effectue dans le plan.

Les réponses du modèle d'ordre réduit POD-Galerkin dont certains coefficients ont été corrigés sont évaluées par rapport à la réponse de référence dans le tableau 6.3. Les différents jeux de coefficients présentés dans le tableau 6.2 sont utilisés afin de corriger le modèle d'ordre réduit construit dans un premier temps avec  $M = 64$  clichés prélevés sur  $n_p \approx 1,5$  période d'oscillation du système. La base de projection contient à chaque fois  $q = 10$  modes POD. La réponse est caractérisée par plusieurs types d'erreurs, relatives à la réponse à court terme ( $\varepsilon_A, \overline{\varepsilon_{C_L}}$ ) ou à long terme ( $\varepsilon_{f_\alpha}, \delta_{\log}$ ). À chaque fois, le temps de calcul nécessaire pour l'optimisation des coefficients est indiqué dans l'avant dernière colonne en secondes. Le portrait de phase du coefficient de portance tracé en fonction de l'angle d'incidence  $\alpha(t)$  est donné dans la dernière colonne.

D'une manière générale, le calibrage des coefficients du modèle d'ordre réduit permet d'améliorer notablement la réponse à court terme : parmi les quatre jeux de paramètres testés, l'erreur sur le coefficient de portance varie entre 5,56% et 19,3%, au lieu de 33,7% avec la correction de [Cazemier et al.](#) L'erreur sur les amplitudes modales est quant à elle aussi raisonnable. L'augmentation du nombre de paramètres calibrés n'est pas toujours synonyme d'une amélioration de l'erreur de reconstruction du modèle d'ordre réduit : en effet, l'erreur augmente considérablement entre les calibrages KdL et K3dL, alors que le nombre de coefficients calibrés  $p_c$  passe de 20 à 38. En outre, la solution devient instable très rapidement après la fenêtre d'échantillonnage. Une explication plausible quant à cet accroissement de l'erreur est que la recherche de l'optimum est de plus en plus difficile lorsque le nombre de paramètres augmente et il est fort probable que l'algorithme n'ait ici convergé que vers un optimum local. Par ailleurs, la diminution de l'erreur à court terme ne permet pas non plus d'affirmer que le comportement à long terme de la réponse sera meilleur, même si il s'agit d'un indice en faveur d'une probable amélioration. Enfin, le calibrage des coefficients non-autonomes est une piste pour l'amélioration de la réponse à court et/ou long terme. En effet, la correction KdL\_E produit un cycle limite plus proche de celui attendu, et ce malgré des erreurs à court terme plus élevées que pour les cas KdL et K3dL\_E.

$M$	$q$	$n_p$	Calibration	Erreurs	Temps CPU	Portrait de phase
64	10	1,5	KdL	$\varepsilon_{\mathbf{A}} = 15,7\%$ $\overline{\varepsilon_{C_L}} = 5,93\%$ $\varepsilon_{f_{\alpha}} = 0,37\%$ $\delta_{\log} = 0,040$	6 008	
64	10	1,5	K3dL	$\varepsilon_{\mathbf{A}} = 23,5\%$ $\overline{\varepsilon_{C_L}} = 19,3\%$ $\varepsilon_{f_{\alpha}} = \text{ND}$ $\delta_{\log} = \text{ND}$	21 597	
64	10	1,5	KdL_E	$\varepsilon_{\mathbf{A}} = 19,2\%$ $\overline{\varepsilon_{C_L}} = 7,48\%$ $\varepsilon_{f_{\alpha}} = 0,22\%$ $\delta_{\log} = 0,045$	10 139	
64	10	1,5	K3dL_E	$\varepsilon_{\mathbf{A}} = 13,0\%$ $\overline{\varepsilon_{C_L}} = 5,56\%$ $\varepsilon_{f_{\alpha}} = 0,83\%$ $\delta_{\log} = 0,101$	16 473	

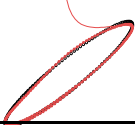
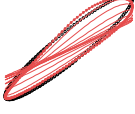
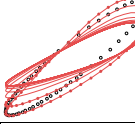
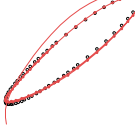
**Tableau 6.3 :** Réponses du modèle d'ordre réduit POD-Galerkin construit avec  $M = 64$  clichés sur  $n_p \approx 1,5$  période et calibré avec différents jeux de paramètres. Les premières colonnes indiquent respectivement le nombre  $M$  de clichés, le nombre  $q$  de modes POD dans la base de projection, la durée  $n_p$  en nombre de périodes sur laquelle les clichés sont prélevés, et le jeu de paramètres calibrés (cf. Table 6.2). La colonne suivante donne les différentes erreurs de reconstruction à court et long terme lorsqu'elles peuvent être calculées (ND signifie non défini). Le temps de calcul pour la résolution du problème d'optimisation est fourni dans l'avant-dernière colonne. Enfin le portrait de phase du coefficient de portance ( $\bullet$ ) calculé avec le modèle d'ordre réduit POD-Galerkin corrigé est comparé avec le coefficient de portance de référence ( $\circ$ ) et la trajectoire à long terme tracée en trait plein ( $\text{---}$ ).

D'autres jeux de paramètres que ceux présentés dans le tableau 6.3 ont été testés sans grand succès avec les mêmes caractéristiques du modèle d'ordre réduit ( $M = 64$ ,  $n_p \approx 1,5$ ,  $q = 10$ ). L'introduction d'un nombre  $p_c$  toujours plus important de paramètres n'est pas une solution souhaitable puisque l'algorithme d'optimisation devient rapidement incapable de trouver l'optimum. Ainsi le calibrage de type KL dans lequel tous les coefficients linéaires sont corrigés ne produit pas l'amélioration escomptée. En outre le temps de calcul devient rapidement réhibitoire puisqu'il atteint déjà plusieurs heures lorsque le nombre de paramètres est de l'ordre de  $\mathcal{O}(q)$ . Le calibrage de certains coefficients quadratiques pourrait être une piste intéressante mais les rares tentatives effectuées n'ont jamais conduit à une amélioration notable de la stabilité de la réponse. Il manque en outre un critère pour permettre de décider quels coefficients quadratiques doivent être corrigés.

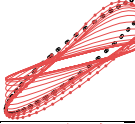
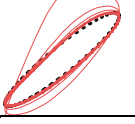

Les informations contenues dans le modèle d'ordre réduit sont alors complétées soit par l'introduction de nouveaux clichés toujours prélevés sur  $n_p \approx 1,5$  période, soit en complétant la base de projection avec des modes négligés précédemment. Le tableau 6.4 présente les résultats pour  $M = 192$  et  $q = 10$  puis pour  $M = 64$  et  $q = 15$ . Pour la correction K3dL, l'augmentation du nombre de clichés permet de réduire considérablement l'erreur à court terme puisque l'erreur sur le coefficient de portance chute de  $\overline{\varepsilon_{C_L}} = 19,3\%$  à seulement  $3,77\%$ . Toutefois, il ne s'agit pas d'une règle générale puisque pour la correction KdL\_E, l'augmentation du nombre de clichés produit une erreur légèrement plus importante. Lorsque le nombre de modes POD dans la base de projection est augmenté, l'erreur à court terme est bien réduite avec la correction KdL, mais elle augmente avec KdL\_E. Dans ce dernier cas, la réponse devient même instable et diverge rapidement. En conclusion, il n'est pas possible d'affirmer que l'augmentation du nombre de clichés ou de modes POD permet de réduire les erreurs et de stabiliser la réponse, bien que quelques améliorations sporadiques aient été constatées.

Enfin une dernière tentative est effectuée en prélevant  $M = 120$  clichés sur  $n_p \approx 2,5$  périodes d'oscillation. En effet, on peut espérer qu'en calibrant sur une durée plus longue l'amortissement du système sera mieux reproduit. Les résultats présentés dans le tableau 6.5 ne sont toujours pas satisfaisants. À nombre de modes POD équivalent, l'erreur est plus importante dans le cas KdL avec  $M = 120$  clichés qu'avec  $M = 64$ ,

6. CONSTRUCTION D'UN MODÈLE D'ORDRE RÉDUIT POD-GALERKIN POUR L'AÉROÉLASTICITÉ

$M$	$q$	$n_p$	Calibration	Erreurs	Temps CPU	Portrait de phase
192	10	1,5	K3dL	$\varepsilon_{\mathbf{A}} = 13,3\%$ $\overline{\varepsilon_{C_L}} = 3,77\%$ $\varepsilon_{f_\alpha} = \text{ND}$ $\delta_{\log} = \text{ND}$	16 206	
192	10	1,5	KdL_E	$\varepsilon_{\mathbf{A}} = 19,3\%$ $\overline{\varepsilon_{C_L}} = 9,59\%$ $\varepsilon_{f_\alpha} = 1,12\%$ $\delta_{\log} = 0,110$	18 708	
64	15	1,5	KdL	$\varepsilon_{\mathbf{A}} = 28,6\%$ $\overline{\varepsilon_{C_L}} = 18,0\%$ $\varepsilon_{f_\alpha} = 0,02\%$ $\delta_{\log} = 0,030$	21 113	
64	15	1,5	KdL_E	$\varepsilon_{\mathbf{A}} = 16,4\%$ $\overline{\varepsilon_{C_L}} = 5,56\%$ $\varepsilon_{f_\alpha} = \text{ND}$ $\delta_{\log} = \text{ND}$	36 630	

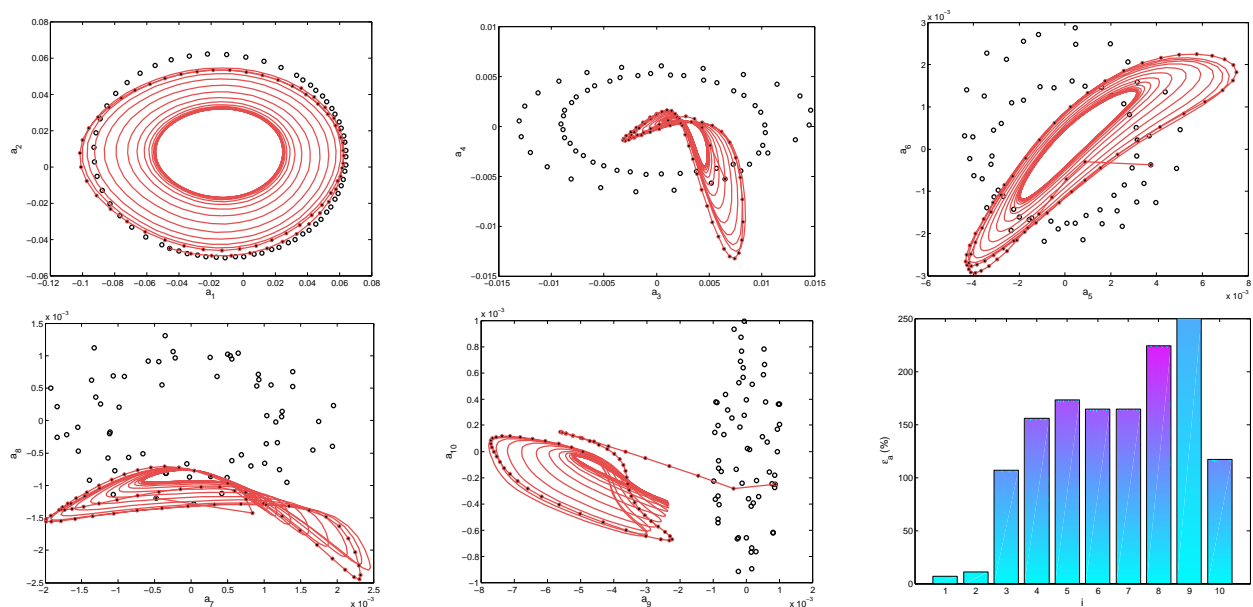
**Tableau 6.4 :** Réponses des modèles d'ordre réduit POD-Galerkin calibrés avec différents jeux de paramètres. Les modèles d'ordre réduit sont construits pour différentes valeurs du nombre  $M$  de clichés ou avec une base de projection contenant un nombre plus élevé  $q$  de modes POD. Les informations sont les mêmes que celles du tableau 6.3.

$M$	$q$	$n_p$	Calibration	Erreurs	Temps CPU	Portrait de phase
120	10	2,5	KdL	$\varepsilon_{\mathbf{A}} = 26,5\%$ $\overline{\varepsilon_{C_L}} = 21,9\%$ $\varepsilon_{f_\alpha} = 0,32\%$ $\delta_{\log} = 0,094$	13 763	
120	10	2,5	KdL_E	$\varepsilon_{\mathbf{A}} = 14,0\%$ $\overline{\varepsilon_{C_L}} = 5,30\%$ $\varepsilon_{f_\alpha} = 0,86\%$ $\delta_{\log} = -0,074$	22 119	
120	15	2,5	KdL_E	$\varepsilon_{\mathbf{A}} = 38,9\%$ $\overline{\varepsilon_{C_L}} = 35,4\%$ $\varepsilon_{f_\alpha} = 0,27\%$ $\delta_{\log} = 0,006$	41 302	

**Tableau 6.5 :** Réponses des modèles d'ordre réduit POD-Galerkin calibrés avec différents jeux de paramètres. Les modèles d'ordre réduit sont construits avec  $M = 120$  clichés prélevés sur  $n_p \approx 2,5$  périodes et  $q = 10$  ou 15 modes POD. Les informations sont les mêmes que celles du tableau 6.3.

ce qui peut se comprendre puisque l'erreur est évaluée sur une durée plus longue au cours de laquelle la déviation par rapport à la solution de référence a tendance à s'accroître. La correction  $\kappa\text{dL\_E}$  produit certes des erreurs nettement inférieures à court terme avec  $M = 120$  clichés, mais la réponse diverge rapidement en dehors de la fenêtre d'échantillonnage. Finalement, le cas  $\kappa\text{dL\_E}$  avec  $q = 15$  modes ne donne pas plus satisfaction, car même si la réponse reste stable à long terme, le cycle limite est totalement erroné.

Malgré ces résultats en demi-teinte, il n'en reste pas moins que la correction apportée par l'optimisation de certains coefficients permet d'améliorer considérablement la réponse par rapport à celle obtenue avec la correction pondérée de *Cazemier et al.* Parmi les nombreux jeux de coefficients calibrés, la correction  $\kappa\text{dL\_E}$  avec  $M = 64$  et  $q = 10$  fournit un compromis acceptable : l'erreur à court terme est assez faible, et la réponse à long terme converge vers un cycle limite qui se rapproche de celui de référence. Les amplitudes modales tracées sur la figure 6.14 font clairement apparaître l'amélioration obtenue en comparaison de la figure 6.12. En particulier l'erreur sur les premières amplitudes modales responsables de l'allure globale de la réponse a été nettement réduite. Les cycles limites pour les autres amplitudes restent parfois éloignés de ceux attendus mais la plupart des erreurs — à court terme — ont pu être diminuées, comme en atteste le dernier graphique de la figure 6.14.

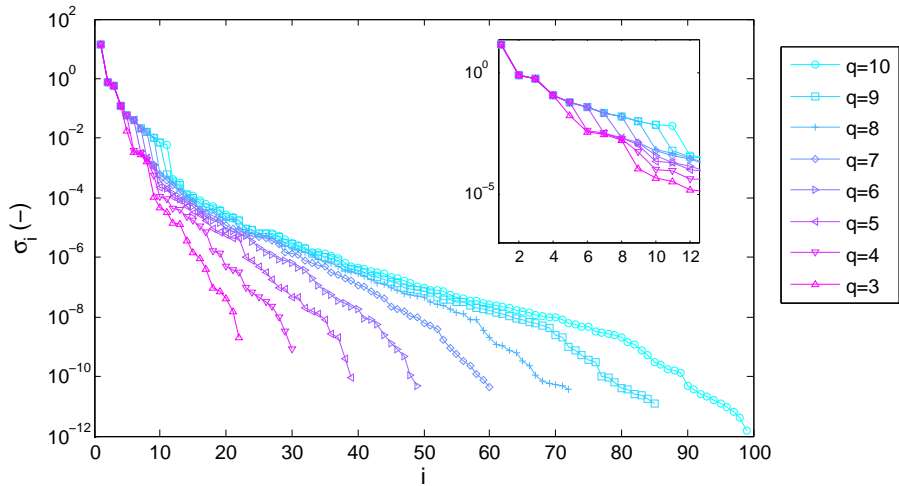


**Figure 6.14 :** Comparaison des portraits de phase des amplitudes modales calculées par intégration du modèle d'ordre réduit calibré pour le jeu de coefficients  $\kappa\text{dL\_E}$ . La légende est identique à celle de la figure 6.12 mais la correction (•) correspond cette fois au calibrage  $\kappa\text{dL\_E}$ .

En conclusion, il apparaît que le modèle d'ordre réduit POD-Galerkin étendu pour tenir compte du mouvement d'une structure est extrêmement instable. La raison la plus probable pour expliquer cela est l'absence de modélisation de la dissipation artificielle, ce qui est en accord avec les problèmes de dissipation observés dans la réponse du modèle d'ordre réduit. Les méthodes de correction mises en œuvre dans le chapitre précédent ne semblent cependant plus être en mesure de pallier les erreurs commises lors de la construction du modèle d'ordre réduit. Malgré toutes les tentatives menées, le modèle d'ordre réduit POD-Galerkin n'a jamais pu être corrigé de manière à reproduire correctement la réponse à long terme, même si la réponse à court terme peut être correctement calculée. Puisque le calcul des coefficients du modèle d'ordre réduit avec les expressions analytiques ne donne pas satisfaction, on cherche dans ce qui suit à identifier les coefficients selon le principe exposé au paragraphe 6.3.3.

### 6.5.2 Réponse du modèle d'ordre réduit POD-Galerkin identifié

Le nombre de coefficients à déterminer dans le problème d'identification 2.65 s'élève à présent à  $N_{c,i}^a + N_{c,i}^{na} = 4 + 4q + q(q + 1)/2$  et le problème devient sur-déterminé uniquement lorsque  $M > N_{c,i}$ . Le problème d'identification est alors résolu au moyen de la technique décrite au paragraphe 2.4.4 et dans l'annexe A.3. La troncature de la décomposition en valeurs singulières effectuée pour le calcul du pseudo-inverse est capitale car elle conditionne la stabilité de la réponse du modèle d'ordre réduit comme on le verra.



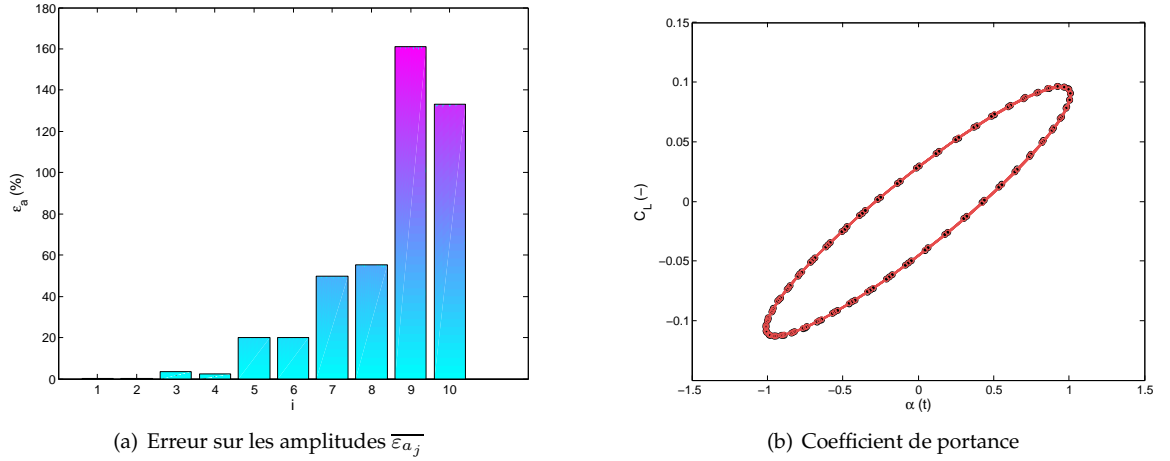
**Figure 6.15 :** Spectre des valeurs singulières de la matrice de corrélation  $\mathcal{A}$  calculée pour différentes valeurs du nombre  $q$  de modes POD employés dans la base de projection. Dans tous les cas,  $M = 120$  clichés ont été prélevés sur  $n_p \approx 2,5$  périodes d'oscillation du système. La discontinuité observée dans le spectre se déplace vers les indices inférieures lorsque  $q$  diminue.

Le spectre des valeurs singulières de la matrice  $\mathcal{A}$  est tracé sur la figure 6.15 pour différentes valeurs  $q$  du nombre de modes POD employés dans la base de projection et lorsque  $M = 120$  clichés ont été échantillonnés sur  $n_p \approx 2,5$  périodes. Le spectre suit une tendance similaire à celle observée sur le spectre des valeurs propres de l'opérateur de corrélation de la méthode des clichés (voir figure 6.7 p.187). En particulier pour  $q = 10$  le même décrochement entre la onzième et la douzième valeur singulière est observé. L'indice de troncature est sélectionné soit a priori en s'aidant du spectre des valeurs singulières de la matrice  $\mathcal{A}$  tracé sur la figure 6.15, soit au moyen de la méthode de la courbe en L (voir [105, 106] et le paragraphe A.3).

Le modèle d'ordre réduit est tout d'abord construit pour  $q = 10$  en conservant  $p = 9$  valeurs singulières pour approcher le pseudo-inverse. L'erreur relative commise par la troncature SVD est extrêmement faible :  $\varepsilon_{\mathcal{A}_p} = 100 \|\mathcal{A}_p - \mathcal{A}\| / \|\mathcal{A}\| = 0,05\%$ . La qualité de la matrice est par ailleurs fortement améliorée puisque le nombre de conditionnement passe de  $\kappa(\mathcal{A}) = 9,37 \times 10^{12}$  à seulement  $\kappa(\mathcal{A}_p) = 1379$ .

L'intégration en temps du modèle d'ordre réduit fournit une réponse en parfait accord avec celle de référence sur l'intervalle d'échantillonnage. L'erreur globale sur les amplitudes chute à  $\varepsilon_{\mathbf{A}} = 1,50\%$  et les portraits de phase des amplitudes modales tracés sur les graphiques du tableau 6.6 sont bien reproduits sur  $I_e$ . L'histogramme de la figure 6.16(a) indique une diminution spectaculaire de l'erreur  $\overline{\varepsilon_{a_j}}$  : l'erreur sur les premières amplitudes est inférieure à 0,2% et culmine au maximum à 161% pour la neuvième. À titre de comparaison, le modèle d'ordre réduit POD-Galerkin corrigé par la méthode de Cazemier *et al.* produisait une erreur aussi importante dès la troisième amplitude et pour toutes les suivantes, avec un record à plus de 2100% d'erreur pour la neuvième amplitude (voir l'histogramme de la figure 6.12).

À l'exception près des dernières amplitudes modales, le comportement à long terme est globalement bien reproduit. Ainsi, sur la période de temps considérée, le système reste stable et l'erreur sur la valeur maximale et minimale du coefficient de portance  $C_L$  reste inférieure à 0,5%, de même que pour l'erreur commise sur l'amplitude crête-à-crête. Par ailleurs le décroissement logarithmique est quasiment nul :  $\delta_{\log} = 6,10 \times 10^{-4}$ . Pour autant, cela ne signifie pas que la réponse soit stable sur le très long terme puisque les portraits de phase des dernières amplitudes du modèle d'ordre réduit présagent d'une éventuelle divergence (voir la figure 6.6, pour les amplitudes  $a_{j,j \geq 7}$ ). Celle-ci survient effectivement lorsque l'intégration



**Figure 6.16 :** L'histogramme (a) représente les erreurs  $\overline{\varepsilon_{a_j}}$  sur les amplitudes modales associées à chaque mode POD sur l'intervalle de temps  $I_c$  et la figure (b) fournit une comparaison sur le même intervalle de temps entre le portrait de phase du coefficient de portance ( $\bullet$ ) calculé avec l'identification des coefficients du modèle réduit et le coefficient de portance de référence ( $\circ$ ). La trajectoire du coefficient de portance à long terme est tracée en trait plein ( $\rightarrow$ ) sur une vingtaine de périodes.

est poursuivie sur une durée équivalente à une centaine de périodes.

Ces premiers résultats indiquent que l'approche d'identification des coefficients permet de construire un modèle d'ordre réduit extrêmement précis à court terme. La solution reconstruite est d'une qualité irréprochable sur l'intervalle de temps d'échantillonnage, aussi bien pour les amplitudes modales que pour les grandeurs macroscopiques mais la réponse à long terme laisse toutefois apparaître des signes d'instabilité qui se manifestent parfois par une divergence. Un comportement similaire avait déjà été évoqué par Sirisup et Karniadakis [216], mais aucune solution adéquate n'avait permis de stabiliser correctement le modèle d'ordre réduit pour des durées extrêmement longues. Dans ce qui suit, on cherche donc à établir un certain nombre de conditions relatives aux paramètres  $M$ ,  $q$ ,  $p$  et  $n_p$  pour lesquels la réponse reste stable à très long terme.

Dans un premier temps on considère le modèle d'ordre réduit construit avec  $M = 120$  clichés prélevés sur  $n_p \approx 2,5$  période d'oscillation avec une base de projection de dimension  $q = 10$ . Les coefficients du modèle d'ordre réduit sont alors calculés comme les solutions approchées définies à l'équation (2.67) où le pseudo-inverse est évalué en conservant un nombre fixe  $p$  de valeurs singulières. Les effets de la valeur de l'indice de troncature sont alors analysés lorsque  $p \in \llbracket 3; 11 \rrbracket$ . Au-delà de  $p = 11$ , les valeurs singulières ajoutées polluent la solution et la réponse du système diverge dans tous les cas. Pour chaque valeur de  $p$ , les coefficients du modèle d'ordre réduit sont identifiés et la réponse est intégrée en temps sur une vingtaine de périodes. Les erreurs à court terme pour les amplitudes modales et le coefficient de portance sont précisées, ainsi que les erreurs à long terme pour la fréquence et l'amortissement du système. Le tableau 6.6 résume ces erreurs et recense les portraits de phase des amplitudes  $a_j$  ainsi que du coefficient de portance.

Plusieurs remarques peuvent être formulées au vu des résultats. Tout d'abord, la réponse macroscopique diverge de manière exponentielle à long terme lorsque le pseudo-inverse est approché avec  $p = 6, 8$  et 10 valeurs singulières alors que les autres indices de troncature rendent le modèle d'ordre réduit stable. Cependant, lorsque l'intégration est poursuivie à très long terme, seule la troncature avec  $p = 7$  valeurs singulières conduit à une réponse stable. D'après les réponses obtenues, il s'avère que la troncature de la matrice  $\mathcal{A}$  avec des valeurs impaires de  $p$  améliore l'évaluation des coefficients. Il est fort probable qu'il existe un lien entre le caractère impair de l'indice de troncature et le regroupement par paire des modes POD qui avait été mis en évidence au paragraphe 6.4.3.

La stabilité du cycle limite du coefficient de portance est toutefois assurée au prix d'une divergence progressive des amplitudes modales  $a_{j,j \geq 7}$  : les oscillations sont en fait entretenues artificiellement par suite d'un accroissement progressif des contributions associées aux derniers modes POD de la base de projection. On peut aussi observer que la divergence des amplitudes diffère suivant le nombre de valeurs singulières



$p$	Erreurs	Amplitudes modales					$C_L$	
3	$\varepsilon_A = 21,0\%$ $\overline{\varepsilon_{C_L}} = 1,43\%$ $\varepsilon_{f_\alpha} = 0,020\%$ $\delta_{\log} = -0,0071$							
4	$\varepsilon_A = 96,4\%$ $\overline{\varepsilon_{C_L}} = 2,51\%$ $\varepsilon_{f_\alpha} = 0,021\%$ $\delta_{\log} = -0,0042$							
5	$\varepsilon_A = 6,72\%$ $\overline{\varepsilon_{C_L}} = 1,35\%$ $\varepsilon_{f_\alpha} = 0,005\%$ $\delta_{\log} = -0,0085$							
6	$\varepsilon_A = 32,6\%$ $\overline{\varepsilon_{C_L}} = 1,56\%$ $\varepsilon_{f_\alpha} = 0,001\%$ $\delta_{\log} = \text{ND}$							
7	$\varepsilon_A = 2,82\%$ $\overline{\varepsilon_{C_L}} = 0,13\%$ $\varepsilon_{f_\alpha} = 0,001\%$ $\delta_{\log} = 0,0007$							
8	$\varepsilon_A = 25,7\%$ $\overline{\varepsilon_{C_L}} = 0,22\%$ $\varepsilon_{f_\alpha} = 0,317\%$ $\delta_{\log} = -0,1290$							
9	$\varepsilon_A = 1,50\%$ $\overline{\varepsilon_{C_L}} = 0,13\%$ $\varepsilon_{f_\alpha} = 0,001\%$ $\delta_{\log} = 0,0006$							
10	$\varepsilon_A = 52,0\%$ $\overline{\varepsilon_{C_L}} = 2,64\%$ $\varepsilon_{f_\alpha} = 0,075\%$ $\delta_{\log} = \text{ND}$							
11	$\varepsilon_A = 1,44\%$ $\overline{\varepsilon_{C_L}} = 0,13\%$ $\varepsilon_{f_\alpha} = 0,014\%$ $\delta_{\log} = 0,0001$							

**Tableau 6.6 :** Réponses des modèles d'ordre réduit construits par identification des coefficients avec  $M = 120$  clichés prélevés sur  $n_p \approx 2,5$  périodes et  $q = 10$  modes POD. La résolution du problème inverse est effectuée à chaque fois en fixant un nombre constant  $p$  de valeurs singulières pour l'approximation du pseudo-inverse  $\mathcal{A}^+$ . Dans chaque cas, on donne les erreurs à court et long terme puis on présente les portraits de phase de toutes les amplitudes modales avant de fournir celui du coefficient de portance.

conservées. Ainsi, le second portrait de phase qui représente  $a_4$  en fonction de  $a_3$  est extrêmement mal reproduit pour  $p < 5$  mais pour  $5 \leq p \leq 10$  la forme du cycle est globalement bien captée. Le même phénomène s'observe pour le troisième portrait de phase — représentant  $a_6$  en fonction de  $a_5$  — qui est bien reproduit pour  $7 \leq p < 11$  et ainsi de suite. La généralisation de cette observation conduit à penser que l'indice de troncature  $p$  nécessaire pour capter correctement  $j$  amplitudes modales est donné par la relation  $j = 2\lfloor(p-1)/2\rfloor$  où  $\lfloor \cdot \rfloor$  désigne la partie entière.

En conclusion de cette première série de résultats, il paraît judicieux de choisir une valeur impaire de l'indice de troncature dans l'intervalle  $\llbracket 0; 11 \rrbracket$  lorsque la base de projection est de dimension  $q = 10$ . La borne supérieure correspond à la discontinuité observée sur le spectre des valeurs singulières tracé sur la figure 6.15 et la borne inférieure est dictée de manière qualitative par le minimum de vecteurs singuliers à conserver pour maintenir une approximation valable du pseudo-inverse  $\mathcal{A}^+$ . Par ailleurs, plus le nombre  $p$  de valeurs singulières est grand, plus le nombre  $j$  d'amplitudes modales correctement représenté est important. D'après la relation établie précédemment, il faudrait ici conserver un minimum de  $p = 11$  valeurs singulières pour que l'ensemble des  $q = 10$  amplitudes modales soit correctement reproduites. Toutefois, cet indice de troncature correspond à la discontinuité observée dans le spectre des valeurs singulières et il s'avère que dans ce cas la réponse est instable.

Le modèle d'ordre réduit est à présent construit avec un nombre  $q$  de modes POD inférieur à 10. Pour pratiquement toutes les valeurs de  $q$  dans l'intervalle  $\llbracket 3; 9 \rrbracket$ , il est possible d'identifier correctement les coefficients pour obtenir une réponse stable à long terme. Mieux encore, le cas  $p = 7$  fournit une réponse stable à très long terme sauf pour  $q = 3$  et 4. Le tableau 6.7 résume les résultats obtenus. Le nombre  $q$  de modes POD et le nombre de coefficients  $N_{c,i} = N_{c,i}^a + N_{c,i}^{na}$  à identifier est précisé dans chaque cas dans la première colonne avec l'indice de troncature  $p$ . Les erreurs commises sont à chaque fois très faibles, et en particulier le décrétement logarithmique est la plupart du temps de l'ordre de  $10^{-4}$  ce qui signifie que le modèle d'ordre réduit reproduit correctement les oscillations entretenues. Les portraits de phase des amplitudes modales sont représentés dans la colonne suivante ; lorsque le nombre  $q$  de modes POD est impair, le dernier graphique représente l'évolution de l'amplitude en fonction de l'angle d'incidence  $\alpha$  pour conserver la forme d'un portrait de phase. Enfin, les portraits de phase du coefficient de portance sont tracés dans la dernière colonne pour une centaine de périodes d'oscillation.

Les résultats obtenus confirment bien que la valeur impaire  $p = 7$  est propice à l'approximation du problème inverse et dans ce cas les  $j = \min(2\lfloor(p-1)/2\rfloor, q)$  premières amplitudes modales sont correctement reproduites, au moins lorsque le nombre  $q$  de modes POD dans la base de projection est dans un voisinage de  $p$  : la valeur  $p = 7$  donne des résultats optimaux pour  $q \in \llbracket 5; 10 \rrbracket$  mais pour de plus faibles valeurs de  $q$ , l'indice de troncature  $p$  chute à 4 ou 5. Cette dernière observation indique que le choix de l'indice de troncature n'est pas totalement décorrélié de la dimension  $q$  de la base de projection, bien que la relation établie entre le nombre  $j$  d'amplitudes modales correctement reproduites et l'indice de troncature  $p$  ne dépende pas directement du nombre  $q$  de modes POD. La seule contrainte est que l'indice de troncature doit être inférieur à l'indice maximal  $p = 11$  correspondant à la discontinuité sur le spectre des valeurs singulières. Or celle-ci se déplace bien lorsque le nombre de modes POD conservés dans la base de projection change comme l'atteste le graphique de la figure 6.15.

Pour conclure, l'influence du nombre  $M$  de clichés sur l'approximation des coefficients est analysée. Les réponses produites pour différentes valeurs de  $q$  sont présentées dans le tableau 6.8 pour  $M = 64$  et 192 clichés échantillonnés sur  $n_p \approx 1,5$  période d'oscillation du système. À chaque fois le pseudo-inverse  $\mathcal{A}^+$  est calculé avec  $p = 7$  valeurs singulières puisque cette valeur avait conduit dans les exemples précédents à des réponses de très bonne qualité lorsque les  $M = 120$  clichés étaient échantillonnés sur  $n_p \approx 2,5$  périodes (voir les tableaux 6.6 et 6.7). Le nombre  $M$  de clichés a peu d'influence : ainsi lorsqu'on compare le cas  $q = 10, p = 7$  avec  $M = 192, 64$  ou 120 clichés, on constate que les erreurs sont quasiment identiques à court terme. L'augmentation du nombre de clichés prélevés sur la même durée  $n_p$  permet de réduire très légèrement l'erreur  $\varepsilon_A$  à court terme, mais l'erreur sur le coefficient de portance  $\overline{\varepsilon_{C_L}}$  n'est pratiquement pas affectée. En revanche, cette erreur est beaucoup plus faible lorsque les  $M = 120$  clichés sont prélevés sur  $n_p \approx 2,5$  périodes. En outre, le décrétement logarithmique est bien meilleur lorsque l'échantillonnage est pratiqué sur une durée plus longue. Ces résultats s'observent aussi pour les cas  $q = 7$  et  $q = 5$ . Il est aussi important de signaler que les réponses sont pratiquement identiques indépendamment du fait que le problème est sur- ou sous-déterminé. En effet, si on considère une base de projection de dimension  $q = 10$ , le problème est sur-déterminé si l'ensemble de clichés contient  $M = 192$  éléments, tandis que le problème

devient sous-déterminé avec  $M = 64$ . Malgré ces différences quant au rang de la matrice  $\mathcal{A}$ , la résolution du problème n'est pas affectée de manière significative et les réponses restent globalement identiques.

$[q, p, N_{c,i}]$	Erreurs	Amplitudes modales					$C_L$
[9, 7, 85]	$\varepsilon_{\mathbf{A}} = 2,81\%$ $\overline{\varepsilon}_{C_L} = 0,14\%$ $\varepsilon_{f_{\alpha}} = 0,001\%$ $\delta_{\log} = 0,0004$						
[8, 7, 72]	$\varepsilon_{\mathbf{A}} = 2,71\%$ $\overline{\varepsilon}_{C_L} = 0,13\%$ $\varepsilon_{f_{\alpha}} = 0,001\%$ $\delta_{\log} = 0,0004$						
[7, 7, 60]	$\varepsilon_{\mathbf{A}} = 2,51\%$ $\overline{\varepsilon}_{C_L} = 0,13\%$ $\varepsilon_{f_{\alpha}} = 0,001\%$ $\delta_{\log} = 0,0004$						
[6, 7, 49]	$\varepsilon_{\mathbf{A}} = 1,32\%$ $\overline{\varepsilon}_{C_L} = 0,13\%$ $\varepsilon_{f_{\alpha}} = 0,001\%$ $\delta_{\log} = 0,0003$						
[5, 7, 39]	$\varepsilon_{\mathbf{A}} = 1,29\%$ $\overline{\varepsilon}_{C_L} = 0,18\%$ $\varepsilon_{f_{\alpha}} = 0,008\%$ $\delta_{\log} = 0,0003$						
[4, 5, 30]	$\varepsilon_{\mathbf{A}} = 1,86\%$ $\overline{\varepsilon}_{C_L} = 1,01\%$ $\varepsilon_{f_{\alpha}} = 0,004\%$ $\delta_{\log} = 0,0005$						
[3, 4, 22]	$\varepsilon_{\mathbf{A}} = 18,8\%$ $\overline{\varepsilon}_{C_L} = 1,03\%$ $\varepsilon_{f_{\alpha}} = 0,023\%$ $\delta_{\log} = 0,0010$						

**Tableau 6.7 :** Réponses des modèles d'ordre réduit construits par identification des coefficients avec  $M = 120$  clichés prélevés sur  $n_p \approx 2,5$  périodes. La base de projection contient un nombre variable  $q$  de modes POD et la résolution du problème inverse est effectuée à chaque fois en fixant un nombre constant  $p$  de valeurs singulières pour l'approximation du pseudo-inverse  $\mathcal{A}^+$ . Dans chaque cas, on donne les erreurs à court et long terme puis on présente les portraits de phase de toutes les amplitudes modales avant de fournir celui du coefficient de portance.

De nombreuses autres configurations de paramètres ont été testées en dehors de celles présentées ici. Celles-ci confirment que les valeurs impaires de l'indice de troncature fournissent en général des réponses de meilleure qualité. Parmi toutes les simulations effectuées ( $q \in \llbracket 3; 10 \rrbracket, p \in \llbracket 3; 11 \rrbracket, M = \{64, 192\}, n_p \approx 1,5$ ) très peu de cas ont fourni une réponse stable à très long terme en dehors des jeux de paramètres  $q = 3, p = 10$  et  $M = 64$  ou  $192$ . Au contraire, lorsque les  $M = 120$  clichés sont échantillonnés sur une durée plus longue  $n_p \approx 2,5$ , de nombreux jeux de paramètres conduisent à une réponse stable à très long terme.

Par conséquent, avec un choix adéquat des paramètres de construction ( $M, q, p, n_p$ ), l'identification des coefficients du modèle d'ordre réduit est une technique particulièrement efficace. D'après les résultats de ce paragraphe, il s'avère qu'une durée de simulation de l'ordre de  $n_p \approx 2,5$  périodes d'oscillation au lieu de  $n_p \approx 1,5$  conduise à de meilleurs résultats pour la stabilité à long terme de la réponse. Le nombre  $M$  de clichés a peu d'influence et ceci rejoint les analyses effectuées pour les modèles d'ordre réduit POD-Galerkin où seul le nombre de modes POD conservés dans la base de projection avait une influence significative. Le choix de l'indice de troncature  $p$  va de pair avec celui du nombre  $q$  de modes POD et doit être sélectionné dans un voisinage de la valeur de  $q$  en dessous de la discontinuité observée dans le spectre des valeurs singulières. On peut alors s'attendre à ce que le nombre  $j$  d'amplitudes modales correctement prédites soit donné par la relation  $j = \min(2\lfloor(p-1)/2\rfloor, q)$ .

À la différence des modèles d'ordre réduit POD-Galerkin, certaines amplitudes modales ont tendance à

6.5. Réponse du modèle d'ordre réduit de l'écoulement autour du profil NACA0064

$[q, p, M, N_{c,i}]$	Erreurs	Amplitudes modales					$C_L$
[10, 7, 64, 99]	$\varepsilon_{\mathbf{A}} = 3,16\%$ $\overline{\varepsilon_{C_L}} = 0,34\%$ $\varepsilon_{f_\alpha} = 0,013\%$ $\delta_{\log} = 0,0081$						
[10, 7, 192, 99]	$\varepsilon_{\mathbf{A}} = 2,97\%$ $\overline{\varepsilon_{C_L}} = 0,33\%$ $\varepsilon_{f_\alpha} = 0,009\%$ $\delta_{\log} = 0,0083$						
[7, 7, 64, 60]	$\varepsilon_{\mathbf{A}} = 2,53\%$ $\overline{\varepsilon_{C_L}} = 0,37\%$ $\varepsilon_{f_\alpha} = 0,013\%$ $\delta_{\log} = 0,0054$						
[7, 7, 192, 60]	$\varepsilon_{\mathbf{A}} = 2,48\%$ $\overline{\varepsilon_{C_L}} = 0,38\%$ $\varepsilon_{f_\alpha} = 0,094\%$ $\delta_{\log} = 0,0057$						
[5, 7, 64, 39]	$\varepsilon_{\mathbf{A}} = 1,28\%$ $\overline{\varepsilon_{C_L}} = 0,38\%$ $\varepsilon_{f_\alpha} = 0,059\%$ $\delta_{\log} = 0,0058$						
[5, 7, 192, 39]	$\varepsilon_{\mathbf{A}} = 1,25\%$ $\overline{\varepsilon_{C_L}} = 0,40\%$ $\varepsilon_{f_\alpha} = 0,053\%$ $\delta_{\log} = 0,0055$						

**Tableau 6.8 :** Réponses des modèles d'ordre réduit construits par identification des coefficients avec  $M = 64$  ou 192 clichés prélevés sur  $n_p \approx 1,5$  période. La base de projection contient un nombre variable  $q$  de modes POD et la résolution du problème inverse est effectuée à chaque fois en fixant un nombre constant  $p = 7$  de valeurs singulières pour l'approximation du pseudo-inverse  $\mathbf{A}^+$ . Dans chaque cas, on donne les erreurs à court et long terme puis on présente les portraits de phase de toutes les amplitudes modales avant de fournir celui du coefficient de portance.

diverger à long terme au lieu de se stabiliser sur un cycle limite. Malgré cela, la réponse macroscopique reste dans de nombreux cas stable à très long terme, même si une divergence en un temps fini n'est pas exclue. Ce phénomène disparaît parfois lorsqu'un très petit nombre d'équations est conservé dans le modèle d'ordre réduit : ainsi, avec  $q = 3, 4$  ou 5 et  $M = 120$  clichés, l'ensemble des amplitudes modales reste stable (voir tableau 6.7) et on peut considérer que ces modèles d'ordre réduit constituent une approximation précise du système fluide. L'identification des coefficients du modèle réduit est donc une technique très puissante puisque le système fluide caractérisé initialement par  $N_v = 32\,768$  degrés de liberté peut être finalement réduit à un système de 3, 4 ou 5 degrés de liberté. En outre, le temps de calcul pour la construction du modèle d'ordre réduit est considérablement réduit par rapport à l'approche POD-Galerkin puisque les coefficients sont directement identifiés à partir de la résolution du problème (2.65) dont le coût CPU est de l'ordre de la seconde alors que les étapes de calcul puis de correction des coefficients dans l'approche POD-Galerkin prenaient plusieurs heures.

## 6.6 Synthèse du chapitre 6

Dans ce dernier chapitre, le problème de la prise en compte du mouvement d'une structure dans le modèle d'ordre réduit a été traité. Les équations de Navier-Stokes qui avaient été formulées au chapitre 4 avec les variables primitives modifiées au moyen d'une description ALE ne sont pas compatibles avec la construction d'un modèle d'ordre réduit basé sur la décomposition orthogonale propre lorsque le domaine du fluide se déforme. Pour remédier à ce problème, les équations de Navier-Stokes ont donc été reformulées dans le repère entraîné associé au mouvement de la structure mobile puisque sous l'hypothèse d'un *mouvement de corps rigide*, le domaine devient invariant. Ainsi, la formulation POD développée au début de ce mémoire reste valable et une projection de Galerkin des équations de Navier-Stokes peut être pratiquée en conservant une cohérence mathématique. Deux modèles d'ordre réduit ont alors été construits en écrivant la projection de Galerkin formelle des équations de Navier-Stokes dans le repère entraîné avec les composantes de la vitesse absolue ou relative. Ces deux modèles d'ordre réduit permettent de traiter les écoulements internes ou externes respectivement, dans lesquels une structure mobile indéformable est la source d'une instationnarité.

Le modèle d'ordre réduit en formulation vitesse absolue a été mis en œuvre sur l'exemple d'un profil NACA0064 oscillant autour d'une position d'équilibre dans un écoulement parfait transsonique. Les équations non-linéaires d'Euler ont alors été employées pour construire le modèle d'ordre réduit et la réponse a été reproduite après que les coefficients aient été corrigés au moyen des méthodes développées au chapitre 5. Toutefois, la construction du modèle d'ordre réduit POD-Galerkin s'avère très coûteuse, notamment à cause du calibrage des coefficients. En outre, la correction introduite n'est pas en mesure de pallier les déficiences du modèle d'ordre réduit dont la réponse est extrêmement instable. Seule la réponse à court terme du modèle d'ordre réduit a pu être reproduite correctement. Une des causes possibles de l'instabilité du modèle d'ordre réduit est l'absence totale de dissipation dans le modèle d'ordre réduit puisque les équations d'Euler ont été employées pour la projection.

L'approche POD-Galerkin a donc été abandonnée au profit de la méthode d'identification des coefficients du modèle réduit à l'aide de laquelle il est possible d'évaluer correctement les coefficients. Ainsi, la réponse reste stable même à très long terme. Une comparaison des coefficients évalués à partir des expressions analytiques ou identifiés est fournie dans l'annexe F.2. Celle-ci révèle que l'identification des coefficients permet de filtrer en quelque sorte les erreurs d'évaluation commises à partir des expressions analytiques et accentue les contributions de certains coefficients. Par ailleurs, l'analyse des contributions relatives a permis de mettre en évidence la prépondérance des termes liés au mouvement de translation. Les termes autonomes et non-autonomes participent conjointement à la réponse du système lorsque les coefficients sont identifiés alors que leurs effets se compensent en grande partie si les coefficients sont évalués à partir des expressions analytiques. Ceci est donc un signe supplémentaire que les expressions analytiques des coefficients nécessitent d'être corrigées.

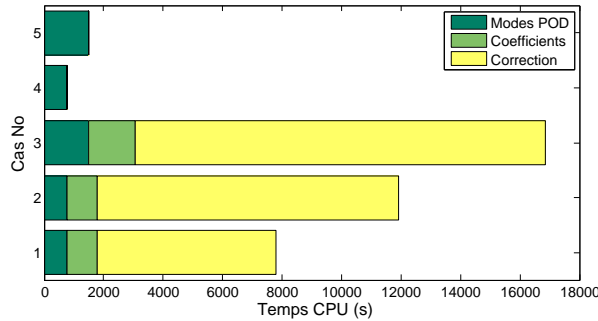
L'identification des coefficients du modèle d'ordre réduit a en outre l'avantage d'être une méthode particulièrement attractive en terme de coût de calcul puisque la détermination des coefficients du modèle d'ordre réduit est pratiquement immédiate et ne nécessite aucune correction ultérieure. Le graphique de la figure 6.17 indique ainsi le coût CPU des différentes étapes de construction<sup>5</sup> de plusieurs modèles d'ordre réduit POD-Galerkin dont les paramètres sont donnés dans le tableau 6.9 (Cas 1, 2 et 3). L'étape de calcul des coefficients est pratiquement négligeable lorsqu'ils sont identifiés alors que leur évaluation à partir des expressions analytiques établies au début de ce chapitre représente pratiquement autant de temps que celui nécessaire au calcul des modes POD. L'inconvénient de l'évaluation des expressions analytiques des coefficients provient cependant de l'étape de correction qui est de loin la plus coûteuse puisque l'algorithme d'optimisation doit faire face à un problème fortement non-linéaire comportant plusieurs dizaines de paramètres.

Une fois que le modèle d'ordre réduit est construit, le temps de simulation de l'écoulement transsonique autour d'un profil en tangage peut être réduit d'un facteur de l'ordre de  $\mathcal{O}(10^2)$  en terme de coût de calcul tout en conservant une précision excellente. Le nombre de degrés de liberté a ainsi été diminué de  $N_v = 32\,768$  à  $q \leq 10$  seulement, ce qui permet d'une part de réduire les coûts de calcul, mais aussi d'étudier plus simplement le fonctionnement du système. Les gains effectués sont résumés sur la figure 6.18 : trois

<sup>5</sup>Les temps CPU indiqués ne tiennent pas compte de la simulation préalable avec le modèle non-réduit qui fournit les clichés. En effet, ce temps de calcul est identique pour l'ensemble des modèles d'ordre réduit construits, quelle que soit l'approche considérée.

modèles d'ordre réduit construits par identification des coefficients avec  $q = 10, 7$  et  $3$  modes POD sont comparés au modèle non réduit au moyen du logarithme du nombre de degrés de liberté ( $n_{\text{ddls}} = N_v$  ou  $q$ ) et du logarithme du gain en coût CPU ( $\eta_{\text{CPU}} = \tau_{\text{CPU elsA}} / \tau_{\text{CPU ROM}}$ ). Les temps  $\tau_{\text{CPU elsA}}$  et  $\tau_{\text{CPU ROM}}$  correspondent respectivement aux temps CPU pour l'intégration en temps du système dynamique sur une période d'oscillation  $T_\alpha$  avec le modèle non-réduit (*elsA*) et avec le modèle d'ordre réduit.

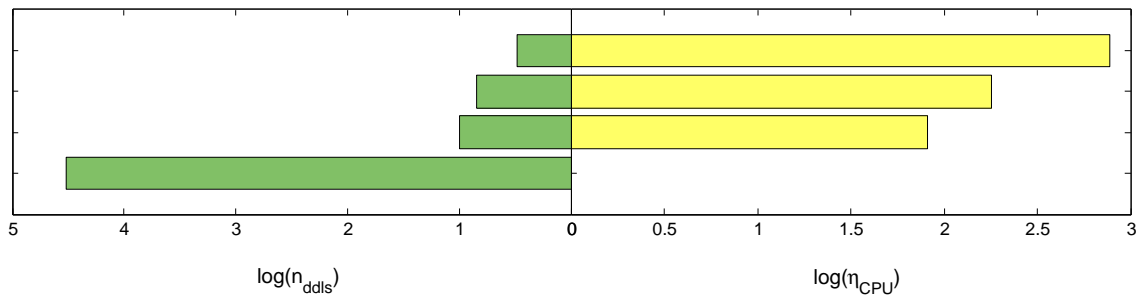
L'identification des coefficients du modèle d'ordre réduit semble donc être une méthode très prometteuse lorsque le calcul explicite des coefficients du modèle d'ordre réduit conduit à une réponse de mauvaise qualité ou devient trop coûteux. Elle permet en effet de reproduire très fidèlement la réponse originale et ce, pour un coût de calcul extrêmement faible. Toutefois il ne s'agit que d'une méthode d'identification qui repose entièrement sur l'expression formelle des équations du modèle d'ordre réduit issues de la projection de Galerkin. Ainsi, l'identification ne permet pas de modéliser correctement les effets qui sont justement à l'origine des défaillances du modèle d'ordre réduit POD-Galerkin, mais fournit simplement une correction plus efficace. Il est donc primordial d'améliorer l'expression de la projection de Galerkin en tenant compte par exemple des effets de la viscosité artificielle.



**Figure 6.17 :** Temps CPU relatif à la construction d'un modèle d'ordre réduit avec l'approche POD-Galerkin ou par identification des coefficients. Les différents cas représentés sont listés dans le tableau 6.9.

Cas No	Approche	$M$	$n_p$	$q$	Correction
1	PODG	64	1,5	10	KdL
2	PODG	64	1,5	10	KdL_E
3	PODG	120	2,5	10	KdL
4	IPOD	64	1,5	10	$p = 7$
5	IPOD	120	1,5	10	$p = 7$

**Tableau 6.9 :** Définition des paramètres des modèles d'ordre réduit construits avec l'approche POD-Galerkin (PODG) ou l'identification des coefficients (IPOD) pour l'évaluation des temps de calcul.



**Figure 6.18 :** Gain en temps CPU et en nombre de degrés de liberté avec différents modèles d'ordre réduit. L'ordonnée correspond au modèle non-réduit ainsi qu'à différents modèles d'ordre réduit construits par identification des coefficients avec  $q = 10, 7$  et  $3$  modes POD. Les abscisses du graphique de gauche représentent le logarithme du nombre de degrés de liberté avec chaque modèle, tandis que le graphique de droite indique le logarithme du gain en temps de calcul pour la simulation d'une période d'oscillation du système. Pour le modèle non réduit, le gain est fixé à 1.



---

# Conclusion

---

Le travail effectué au cours de cette thèse a contribué au développement d'une méthode de réduction efficace dans le but de diminuer les coûts de calcul qui deviennent considérables lorsqu'il s'agit de modéliser correctement l'écoulement d'un fluide dont le couplage avec le mouvement d'une structure est à l'origine de phénomènes aéroélastiques. En effet, la résolution des équations non-linéaires d'Euler ou de Navier-Stokes qui sont indispensables pour capter certains phénomènes implique un nombre de degrés de liberté si important que le coût de calcul devient rapidement prohibitif lorsqu'il s'agit de mener des études paramétriques du système.

La première partie de ce mémoire, dévolue à la présentation de la méthode de réduction et à son application à un système dynamique linéaire, a été l'occasion d'évaluer plusieurs formulations sur un problème relativement simple. Dans le chapitre 1, l'état de l'art des principales méthodes de réduction physiques a révélé que la décomposition orthogonale propre remplit les critères nécessaires pour le développement d'un modèle d'ordre réduit non-linéaire représentatif d'un fluide compressible visqueux. En effet, l'approche POD-Galerkin dans laquelle la décomposition orthogonale propre est combinée à une projection de Galerkin des équations de Navier-Stokes conduit à un modèle d'ordre réduit dans lequel les non-linéarités sont préservées. Le principe de la décomposition orthogonale propre rappelé au chapitre 2 a permis d'établir un certain nombre de propriétés employées par la suite pour la construction du modèle d'ordre réduit. L'application de la méthode au chapitre 3 à un système dynamique linéaire a démontré que l'approche POD-Galerkin possédait de bonnes propriétés en termes d'erreur de reconstruction et de convergence en comparaison avec une formulation analytique ou discrète. Par la suite, trois techniques ont été développées puis comparées pour tenir compte avec l'approche POD-Galerkin de conditions aux limites non-homogènes de type Dirichlet et Neumann. La réponse du système pour diverses fréquences et amplitudes d'oscillation est alors obtenue simplement en modifiant le terme source du modèle d'ordre réduit non-autonome qui apparaît afin de reproduire l'effet de la condition limite. Ce premier modèle d'ordre réduit constitue ainsi l'archétype du problème aéroélastique dans lequel le mouvement de la structure correspond à la condition limite non-homogène.

Dans la seconde partie de ce mémoire le travail a consisté à développer un modèle d'ordre réduit non-linéaire représentatif d'un fluide compressible visqueux. Dans le chapitre 4 plusieurs formulations des équations de Navier-Stokes résolues en pratique par les méthodes de discrétisation classiques — en Volumes Finis ou Éléments Finis par exemple — ont été rappelées. La solution de ces équations a en particulier servi à générer l'ensemble des clichés nécessaires au calcul des modes POD de la base de projection. Une discussion sur le choix des équations adéquates pour la construction du modèle d'ordre réduit a alors abouti à choisir les variables primitives modifiées au lieu des variables conservatives afin que les équations de Navier-Stokes prennent la forme d'un système dynamique quadratique qui se prête bien à la projection de Galerkin. La formulation de Vigo [240] établie avec une description eulérienne a alors été étendue à une description ALE formulée dans le repère absolu ou dans le repère entraîné. Dans ce dernier cas, l'emploi des composantes de la vitesse absolue ou relative a conduit à deux formulations différentes. L'hypothèse d'un mouvement de corps rigide qui a finalement été émise se justifie par la nécessité de conserver un domaine indéformable de sorte que la définition du produit scalaire intervenant dans le processus de la décomposition orthogonale propre reste cohérente. Au chapitre 5, le modèle d'ordre réduit POD-Galerkin des équations des Navier-Stokes a été développé à partir de la formulation dans le repère absolu. Celui-ci a permis de reproduire le phénomène de détachement de vortex dans le sillage du profil NACA0012 fortement incliné. Le manque de dissipation important constaté dans la réponse du modèle d'ordre réduit a pu être corrigé au moyen de diverses techniques reposant soit sur l'introduction d'amortissement dans le système, soit sur un calibrage des coefficients du modèle d'ordre réduit. Finalement, le travail mené au chapitre 6 a



consisté à développer le modèle d'ordre réduit à partir des équations de Navier-Stokes écrites dans le repère entraîné en faisant l'hypothèse d'un mouvement de corps rigide de manière à préserver la cohérence de la définition des modes POD. Le modèle d'ordre réduit a alors été mis en œuvre pour reproduire le phénomène d'oscillation d'un choc sur la paroi d'un profil NACA0064 oscillant dans un écoulement transsonique modélisé par les équations d'Euler. Dans ce cas, la réponse s'est avérée être très instable puisque aucune dissipation n'est introduite dans le modèle d'ordre réduit. Les méthodes de correction employées au chapitre 5 se sont alors révélées être insuffisantes pour corriger correctement la réponse à (très) long terme. Le recours à la méthode d'identification des coefficients a néanmoins permis de remédier à ce problème de stabilité et par ailleurs le coût de construction du modèle d'ordre réduit s'en est trouvé nettement amélioré.

Le potentiel de la méthode de réduction développée dans ce mémoire au moyen d'une approche POD-Galerkin a été démontré sur deux exemples représentatifs d'un écoulement non-linéaire, visqueux et compressible tel que celui rencontré dans les applications turbomachines. Des gains substantiels en temps de calcul ont ainsi pu être obtenus puisque l'écoulement est reproduit de manière très précise en un temps extrêmement court. Il faut cependant relativiser les gains présentés sur le graphique de la figure 6.17 puisqu'ils ne concernent que l'intégration en temps du modèle d'ordre réduit et ne tiennent pas compte de la construction du modèle d'ordre réduit qui comprend notamment la génération des clichés avec le modèle non-réduit. Le gain ne sera donc en réalité effectif que lorsque le *même* modèle d'ordre réduit sera employé pour reproduire la réponse sur une certaine gamme de paramètres.

Le principal apport de la thèse est donc le développement de la forme générique du modèle d'ordre réduit des équations de Navier-Stokes. En effet, la projection de Galerkin a ici été écrite en détail pour une base de modes POD, mais d'une manière plus générale elle reste valable pour n'importe quelle base orthonormale sur laquelle le champ aérodynamique est décomposé. Par conséquent le choix d'une autre base spatiale est envisageable et permettrait de s'affranchir de l'hypothèse d'un mouvement de corps rigide émise ici pour préserver la cohérence de la définition des modes POD.

Pour conclure, quelques pistes de développements à explorer sont proposées pour améliorer ou poursuivre la construction de modèles d'ordre réduit non-linéaires d'un fluide compressible visqueux :

- Étude et amélioration de la stabilité : les problèmes d'instabilité des modèles d'ordre réduit POD-Galerkin semblent devoir être attribués principalement à la modélisation de la viscosité. En effet, la loi de Sutherland a été négligée et l'introduction d'une viscosité artificielle analogue à celle employée en Volumes Finis [157] semble nécessaire. Il serait par ailleurs souhaitable de disposer de critères théoriques pour évaluer la stabilité à partir des opérateurs du modèle d'ordre réduit au lieu de la réponse intégrée sur le long terme. Un premier moyen de procéder consiste à linéariser le système et à analyser les valeurs propres à l'instar de la procédure proposée par Dowell *et al.* [60]. Une analyse plus fine devrait cependant être envisagée puisqu'il s'agit d'un système dynamique non-linéaire dont la stabilité peut être estimée d'une certaine manière au moyen des exposants de Lyapunov et des multiplicateurs de Floquet [98].
- Prise en compte d'une structure déformable : en l'état actuel des choses, la formulation développée n'est valable que pour une structure indéformable animée d'un mouvement de corps rigide. L'extension au cas d'une structure déformable est délicate et nécessite de reformuler la décomposition orthogonale propre dans le cadre d'un domaine spatial variable au cours du temps. Une alternative possible est de remplacer les modes POD par les PIPs [107, 137] qui ne correspondent pas à des corrélations spatiales et autoriseraient le travail sur un domaine déformable.
- Réponse paramétrique : l'enjeu final reste la capacité du modèle d'ordre réduit à reproduire la réponse sur une certaine gamme de paramètres. La formulation actuelle a été développée de telle sorte qu'il est possible de faire varier les vitesses de translation et de rotation du profil, si bien que l'amplitude et la fréquence du mouvement imposé sont déjà deux paramètres qui peuvent potentiellement être considérés pour la réponse paramétrique. En revanche, la réponse sur une certaine gamme de nombre de Mach nécessitera certainement le développement de méthodes d'interpolation entre plusieurs bases POD [10, 147].

Enfin, d'un point de vue purement informatique, il faut signaler que l'ensemble des développements a été effectué ici avec le langage de programmation Scilab. L'emploi d'un langage plus robuste tel que le C++ permettrait de tirer parti d'un certain nombre de fonctionnalités déjà développées au sein du logiciel *elsA*, et de bénéficier de l'approche orientée objet récemment développée pour la POD [14].

## **Troisième partie**

# **Annexes**



## Éléments d'algèbre linéaire et matriciel

---

### Sommaire

A.1 Matrice d'applications linéaires et changement de base . . . . .	211
A.2 Décomposition en valeurs singulières . . . . .	213
A.3 Résolution approchée de problèmes linéaires . . . . .	214

---

Dans cette annexe, on rappelle un certain nombre de définitions et de propriétés d'algèbre linéaire auxquelles il est souvent fait référence dans le manuscrit. La définition des matrices en tant que représentation d'une application linéaire est d'abord rappelée avant de présenter la procédure de décomposition en valeurs singulières d'une matrice quelconque puis de donner quelques techniques de résolution approchées d'un système linéaire. Un certain nombre de compléments pourront être trouvés dans les ouvrages de Strang [223] et Golub et Van Loan [93] par exemple.

### A.1 Matrice d'applications linéaires et changement de base

On se place dans le cadre général d'un espace vectoriel  $E$  défini sur un corps  $\mathbb{K}$ , de dimension finie  $N$  tel que  $\mathcal{B} = (e_1, \dots, e_N)$  soit une base de  $E$ . On désignera par  $\mathcal{M}(N, M)$  l'ensemble des matrices de taille  $N \times M$  dont les composantes sont des scalaires définis sur le corps  $\mathbb{K}$ . Tout vecteur  $x \in E$  se décompose sur la base  $\mathcal{B}$  en

$$x = \sum_{i=1}^N x_i e_i, \tag{A.1}$$

où les  $x_i \in \mathbb{K}$  sont les *coordonnées* de  $x$  dans la base  $\mathcal{B}$ . Le vecteur  $x$  est représenté dans la base  $\mathcal{B}$  par la matrice unicolonne

$$\text{Mat}_{\mathcal{B}}(x) = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \tag{A.2}$$

qui appartient à  $\mathbb{R}^N$  et dont les coefficients sont les coordonnées de  $x$ . On considère à présent une famille  $(x_1, \dots, x_P)$  de  $P$  vecteurs de  $E$ . Alors la matrice de cette famille dans la base  $\mathcal{B}$  est la matrice  $\text{Mat}_{\mathcal{B}}(x_1, \dots, x_P) \in \mathcal{M}(N, P)$  dont la  $j$ -ième colonne est  $\text{Mat}_{\mathcal{B}}(x_j)$  puisque pour tout  $j$ ,  $x_j = \sum_{i=1}^N x_{ij} e_i$  d'après la relation (A.1). La matrice de la famille  $(x_1, \dots, x_P)$  s'écrit alors

$$\text{Mat}_{\mathcal{B}}(x_1, \dots, x_P) = \begin{bmatrix} x_{11} & \cdots & x_{1j} & \cdots & x_{1P} \\ \vdots & & \vdots & & \vdots \\ x_{N1} & \cdots & x_{Nj} & \cdots & x_{NP} \end{bmatrix}. \tag{A.3}$$

Les matrices sont en particulier employées pour représenter des applications linéaires quelconques dans les bases de départ et d'arrivée de l'application. Soient  $E$  et  $F$  deux espaces vectoriels de dimensions finies  $P$  et  $N$  respectivement munis des bases  $\mathcal{B} = (e_1, \dots, e_P)$  et  $\mathcal{C} = (f_1, \dots, f_N)$ , et  $u : E \rightarrow F$  une application linéaire de  $E$  dans  $F$ . On appelle matrice de  $u$  relativement aux bases  $\mathcal{B}$  et  $\mathcal{C}$  la matrice  $\text{Mat}_{\mathcal{B}, \mathcal{C}}(u)$  de la famille  $(u(e_1), \dots, u(e_P))$  exprimée dans la base  $\mathcal{C}$  avec la relation  $u(e_j) = \sum_{i=1}^N a_{ij} f_i$ , ce qui conduit à

$$\text{Mat}_{\mathcal{B}, \mathcal{C}}(u) = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1P} \\ \vdots & & \vdots & & \vdots \\ a_{N1} & \cdots & a_{Nj} & \cdots & a_{NP} \end{bmatrix}. \quad (\text{A.4})$$

Chaque colonne correspond donc à l'image par  $u$  du vecteur de base  $e_j \in E$  exprimée dans la base  $\mathcal{C}$  sur les vecteurs de base  $f_i$ . La notion de *changement de base* apparaît naturellement lorsqu'on considère seulement l'espace vectoriel  $E$  et deux bases  $\mathcal{B} = (e_1, \dots, e_N)$  et  $\mathcal{B}' = (e'_1, \dots, e'_N)$ . On appelle alors *matrice de passage de  $\mathcal{B}'$  à  $\mathcal{B}$*  la matrice de la famille  $(e'_1, \dots, e'_N)$  exprimée dans  $\mathcal{B}$  telle que  $e'_j = \sum_{i=1}^N p_{ij} e_i$ . Ainsi la matrice de passage s'écrit

$$\mathbf{P}_{\mathcal{B}' \rightarrow \mathcal{B}} = \text{Mat}_{\mathcal{B}, \mathcal{B}'}(\text{Id}_E) = \begin{bmatrix} p_{11} & \cdots & p_{1j} & \cdots & p_{1N} \\ \vdots & & \vdots & & \vdots \\ p_{N1} & \cdots & p_{Nj} & \cdots & p_{NN} \end{bmatrix}. \quad (\text{A.5})$$

L'opération de changement de base peut être vue comme une application linéaire qui est simplement l'identité de  $E$ , mais où la base  $\mathcal{B}'$  est employée pour exprimer les vecteurs de l'ensemble de départ alors que les vecteurs de l'image sont exprimés dans  $\mathcal{B}$ . On a les propriétés suivantes :

- La matrice de passage d'une base dans elle-même est l'identité :

$$\mathbf{P}_{\mathcal{B} \rightarrow \mathcal{B}} = \mathbf{I}_N. \quad (\text{A.6})$$

- Soient  $\mathcal{B}, \mathcal{B}'$  et  $\mathcal{B}''$  trois bases de  $E$ , alors

$$\mathbf{P}_{\mathcal{B}'' \rightarrow \mathcal{B}} = \mathbf{P}_{\mathcal{B}'' \rightarrow \mathcal{B}'} \mathbf{P}_{\mathcal{B}' \rightarrow \mathcal{B}}. \quad (\text{A.7})$$

La relation précédente traduit la *transitivité* du changement de base, qui s'exprime simplement par un produit matriciel entre les matrices de passage.

- La matrice de passage d'une base  $\mathcal{B}$  à une base  $\mathcal{B}'$  est inversible, et son inverse est donné par la matrice de passage de  $\mathcal{B}'$  à  $\mathcal{B}$  ce qui s'écrit

$$(\mathbf{P}_{\mathcal{B}' \rightarrow \mathcal{B}})^{-1} = \mathbf{P}_{\mathcal{B} \rightarrow \mathcal{B}'}. \quad (\text{A.8})$$

On peut donc définir une formule de changement de base pour un vecteur et une matrice quelconques. Soient  $\mathcal{B}$  et  $\mathcal{B}'$  deux bases de l'espace vectoriel  $E$  de dimension finie  $N$ , et  $\mathbf{P}_{\mathcal{B}' \rightarrow \mathcal{B}}$  la matrice de passage entre les deux bases. Pour tout vecteur  $x \in E$ , on note  $\mathbf{X} = \text{Mat}_{\mathcal{B}}(x)$  la représentation de ce vecteur dans la base  $\mathcal{B}$  et  $\mathbf{X}' = \text{Mat}_{\mathcal{B}'}(x)$  sa représentation dans  $\mathcal{B}'$ . Alors la formule de passage pour le vecteur  $x$  s'écrit

$$\boxed{\mathbf{X} = \mathbf{P}_{\mathcal{B}' \rightarrow \mathcal{B}} \mathbf{X}'}. \quad (\text{A.9})$$

En effet, un vecteur  $x \in E$  se décompose sur les bases  $\mathcal{B}$  et  $\mathcal{B}'$  en  $x = \sum_{i=1}^N x_i e_i = \sum_{j=1}^N x'_j e'_j$  selon la relation (A.1). Or la définition (A.5) de la matrice de passage permet d'écrire  $x = \sum_{j=1}^N x'_j e'_j = \sum_{j=1}^N x'_j (\sum_{i=1}^N p_{ij} e_i) = \sum_{i=1}^N (\sum_{j=1}^N p_{ij} x'_j) e_i$ . Ainsi, la coordonnée  $x_i$  du vecteur  $x$  exprimé dans  $\mathcal{B}$  correspond à  $\sum_{j=1}^N p_{ij} x'_j$  d'où la relation (A.9) entre les composantes des coordonnées du vecteur  $x$  exprimé dans les deux bases. Inversement, on obtient  $\mathbf{X}' = (\mathbf{P}_{\mathcal{B}' \rightarrow \mathcal{B}})^{-1} \mathbf{X} = \mathbf{P}_{\mathcal{B} \rightarrow \mathcal{B}'} \mathbf{X}$  au moyen de la relation (A.8).

De la même manière on considère une application linéaire  $u : E \rightarrow F$  définie sur les espaces vectoriels  $E$  et  $F$  de dimension  $P$  et  $N$ . Soient  $\mathcal{B}$  et  $\mathcal{B}'$  deux bases de  $E$  telle que  $\mathbf{P} = \mathbf{P}_{\mathcal{B}' \rightarrow \mathcal{B}}$  soit la matrice passage de l'une à l'autre, et  $\mathcal{C}$  et  $\mathcal{C}'$  deux bases de  $F$  telle que  $\mathbf{Q} = \mathbf{Q}_{\mathcal{C}' \rightarrow \mathcal{C}}$  soit la matrice passage de l'une à l'autre. On note  $\mathbf{M} = \text{Mat}_{\mathcal{B}, \mathcal{C}}(u)$  et  $\mathbf{M}' = \text{Mat}_{\mathcal{B}', \mathcal{C}'}(u)$  les matrices de  $\mathcal{M}(N, P)$  représentant l'application linéaire  $u$ . On a alors la relation de passage

$$\boxed{\mathbf{M}' = \mathbf{Q}^{-1} \mathbf{M} \mathbf{P}}. \quad (\text{A.10})$$

Si en outre  $u$  est un automorphisme de  $E$  dans  $E$ , et si on note  $\mathbf{P} = \mathbf{P}_{\mathcal{B}' \rightarrow \mathcal{B}}$ ,  $\mathbf{M} = \text{Mat}_{\mathcal{B}, \mathcal{B}}(u)$  et  $\mathbf{M}' = \text{Mat}_{\mathcal{B}', \mathcal{B}'}(u)$ , alors la relation devient

$$\boxed{\mathbf{M}' = \mathbf{P}^{-1} \mathbf{M} \mathbf{P}}. \quad (\text{A.11})$$

Les relations données aux équations (A.9) et (A.11) seront particulièrement utiles pour transposer les équations de Navier-Stokes d'un repère absolu  $\mathcal{R}_A = (O, \mathcal{B}_A)$  muni de la base  $\mathcal{B}_A$  dans le repère entraîné  $\mathcal{R}_E = (M_0, \mathcal{B}_E)$  muni de la base mobile  $\mathcal{B}_E$ . Le lecteur pourra se reporter par exemple à [223, Annexe A] pour une analyse plus détaillée concernant les relations de changement de base.

## A.2 Décomposition en valeurs singulières

La décomposition en valeurs singulières d'une matrice  $\mathbf{A} \in \mathcal{M}(N, M)$  est un moyen efficace de factoriser et éventuellement d'approcher la matrice par une autre de rang inférieur. La décomposition est énoncée dans le théorème suivant.

### Théorème A.1 : Décomposition en Valeurs Singulières

Soit  $\mathbf{A} \in \mathcal{M}(N, M)$  une matrice quelconque de rang  $1 \leq r \leq \min(N, M)$ , alors il existe deux matrices  $\mathbf{Q}_1 \in \mathcal{M}(N, N)$  et  $\mathbf{Q}_2 \in \mathcal{M}(M, M)$  et une matrice diagonale  $\mathbf{\Sigma} \in \mathcal{M}(N, M)$  telles que

$$\mathbf{A} = \mathbf{Q}_1 \mathbf{\Sigma} \mathbf{Q}_2^{\text{T}}. \quad (\text{A.12})$$

La décomposition précédente est la décomposition en valeurs singulières de la matrice  $\mathbf{A}$ . Les colonnes des matrices  $\mathbf{Q}_1$  et  $\mathbf{Q}_2$  sont les vecteurs singuliers à gauche et à droite respectivement et vérifient les conditions d'orthogonalité

$$\mathbf{Q}_1^{\text{T}} \mathbf{Q}_1 = \mathbf{I}_N \quad \text{et} \quad \mathbf{Q}_2^{\text{T}} \mathbf{Q}_2 = \mathbf{I}_M. \quad (\text{A.13})$$

La matrice  $\mathbf{\Sigma}$  contient sur sa diagonale les  $r$  valeurs singulières  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$  non nulles, telle que

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_1 & & 0 & \dots & 0 \\ & \ddots & \vdots & & \vdots \\ & & \sigma_r & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix} = \begin{pmatrix} \mathbf{\Sigma}_r & \mathbf{0}_{r, M-r} \\ \mathbf{0}_{N-r, r} & \mathbf{0}_{N-r, M-r} \end{pmatrix}. \quad (\text{A.14})$$

La première preuve de la décomposition en valeurs singulières pour les matrices rectangulaires et complexes est attribuée à Eckart et Young [66] mais le calcul efficace n'a été rendu possible que récemment avec notamment l'algorithme de Golub/Kahan [91, 92].

Dans la pratique, on a souvent recours à la décomposition en valeurs singulières *tronquée* pour laquelle on ne considère que les  $r$  premiers vecteurs singuliers dans les matrices  $\mathbf{Q}_1$  et  $\mathbf{Q}_2$  puisque les  $N - r$  vecteurs restant dans  $\mathbf{Q}_1$  et les  $M - r$  vecteurs restant dans  $\mathbf{Q}_2$  appartiennent à l'espace nul. Cela conduit à travailler avec les matrices  $\mathbf{Q}_{1,r} \in \mathcal{M}(N, r)$  et  $\mathbf{Q}_{2,r} \in \mathcal{M}(M, r)$  telles que  $\mathbf{Q}_1 = [\mathbf{Q}_{1,r} \ \mathbf{Q}_{1, N-r}]$  et  $\mathbf{Q}_2 = [\mathbf{Q}_{2,r} \ \mathbf{Q}_{2, M-r}]$  respectivement. Dans ce cas la décomposition en valeurs singulières de la matrice  $\mathbf{A}$  s'écrit

$$\mathbf{A} = (\mathbf{Q}_{1,r} \ \mathbf{Q}_{1, N-r}) \begin{pmatrix} \mathbf{\Sigma}_r & \mathbf{0}_{r, M-r} \\ \mathbf{0}_{N-r, r} & \mathbf{0}_{N-r, M-r} \end{pmatrix} \begin{pmatrix} \mathbf{Q}_{2,r}^{\text{T}} \\ \mathbf{Q}_{2, M-r}^{\text{T}} \end{pmatrix} = \mathbf{Q}_{1,r} \mathbf{\Sigma}_r \mathbf{Q}_{2,r}^{\text{T}}. \quad (\text{A.15})$$

Les valeurs singulières associées à la décomposition décroissent en général de façon extrêmement rapide, ce qui signifie que le contenu de la matrice est principalement caractérisé par les premiers vecteurs singuliers qui sont associés aux plus grandes valeurs singulières. Il est alors courant d'approcher la matrice  $\mathbf{A}$  par une décomposition en valeurs singulières tronquée qui, à la différence de l'équation (A.15), est inexacte. Si les  $p < r$  premières valeurs singulières sont conservées, alors la matrice  $\mathbf{A}$  s'approche au moyen des matrices  $\mathbf{Q}_{1,p} \in \mathcal{M}(N, p)$ ,  $\mathbf{Q}_{2,p} \in \mathcal{M}(M, p)$  et  $\mathbf{\Sigma}_p \in \mathcal{M}(p, p)$  par

$$\mathbf{A} \approx \mathbf{A}_p = \mathbf{Q}_{1,p} \mathbf{\Sigma}_p \mathbf{Q}_{2,p}^{\text{T}}. \quad (\text{A.16})$$

La décomposition en valeurs singulières est un processus largement utilisé en algèbre linéaire qui permet notamment de calculer la *pseudo-inverse* d'une matrice. Celui-ci généralise la définition de l'inverse d'une matrice au cas d'une matrice rectangulaire quelconque. Par définition, le pseudo-inverse d'une matrice  $\mathbf{A} \in \mathcal{M}(N, M)$  est donné par l'expression

$$\mathbf{A}^+ = \mathbf{Q}_2 \Sigma^+ \mathbf{Q}_1^T. \quad (\text{A.17})$$

La matrice  $\Sigma^+$  correspond elle-même au pseudo-inverse de la matrice  $\Sigma$ , qui est simplement définie pour une matrice diagonale au moyen des inverses non nuls de la diagonale principale par

$$\Sigma^+ = \begin{pmatrix} \Sigma_r^+ & \mathbf{0}_{r, M-r} \\ \mathbf{0}_{N-r, r} & \mathbf{0}_{N-r, M-r} \end{pmatrix} \quad \text{avec} \quad \Sigma_r^+ = \text{diag}(\sigma_1^{-1}, \dots, \sigma_r^{-1}). \quad (\text{A.18})$$

Le pseudo-inverse ainsi défini par l'équation (A.17) est l'unique matrice de  $\mathcal{M}(N, M)$  qui vérifie les quatre conditions de Moore-Penrose suivantes :

$$\mathbf{A} \mathbf{A}^+ \mathbf{A} = \mathbf{A}, \quad (\text{A.19a})$$

$$\mathbf{A}^+ \mathbf{A} \mathbf{A}^+ = \mathbf{A}^+, \quad (\text{A.19b})$$

$$(\mathbf{A} \mathbf{A}^+)^T = \mathbf{A} \mathbf{A}^+, \quad (\text{A.19c})$$

$$(\mathbf{A}^+ \mathbf{A})^T = \mathbf{A}^+ \mathbf{A}. \quad (\text{A.19d})$$

Si les colonnes de  $\mathbf{A}$  sont linéairement indépendantes, alors  $r = M$  et  $\mathbf{A}^T \mathbf{A}$  est inversible ; il s'ensuit que le pseudo-inverse est explicitement défini par  $\mathbf{A}^+ = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$  avec les propriétés (A.19a) et (A.19c)<sup>1</sup>. Par conséquent on retrouve dans ce cas la définition du pseudo-inverse obtenue avec le problème moindres carrés (A.22). De la même façon, si les lignes de  $\mathbf{A}$  sont linéairement indépendantes, alors  $r = N$  et la matrice  $\mathbf{A} \mathbf{A}^T$  est inversible ; on montre alors que le pseudo-inverse est donné par  $\mathbf{A}^+ = \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1}$ . Enfin, si les lignes et les colonnes de  $\mathbf{A}$  sont linéairement indépendantes, alors la matrice  $\mathbf{A}$  est inversible et le pseudo-inverse coïncide avec l'inverse classique :  $\mathbf{A}^+ = \mathbf{A}^{-1}$ .

### A.3 Résolution approchée de problèmes linéaires

À plusieurs reprises dans ce document, il est question de résoudre un problème linéaire de la forme

$$\mathbf{A} \mathbf{X} = \mathbf{B}, \quad (\text{A.20})$$

où  $\mathbf{X} \in \mathbb{R}^M$  est le vecteur des inconnues qui doivent être déterminées,  $\mathbf{A} \in \mathcal{M}(N, M)$  est la matrice des données qui est non inversible dans le cas général et  $\mathbf{B} \in \mathbb{R}^N$  est le vecteur des observations.

Le problème (A.20) a soit une solution unique, soit une infinité de solutions, soit aucune solution en fonction de la valeur du rang  $r < \min(N, M)$  de la matrice  $\mathbf{A}$ . Pour plus de détails, le lecteur pourra consulter [223, Chap. 2] et [93, Chap. 5]. Lorsque  $N < M$ , alors le problème est *sous-déterminé* puisqu'il y a moins d'équations que d'inconnues à déterminer. Dans ce cas,  $r \leq N$  et on peut toujours supposer que les  $r$  premières lignes de la matrice  $\mathbf{A}$  sont indépendantes et que les  $N - r$  dernières sont des combinaisons linéaires des  $r$  premières. Alors si les  $N - r$  derniers termes du second membre  $\mathbf{B}$  vérifient les mêmes combinaisons linéaires dites de compatibilité, les  $N - r$  dernières équations sont superflues. Dans le cas contraire, le système n'a pas de solution puisque tout vecteur solution d'une des  $r$  premières lignes ne peut être solution d'une des  $N - r$  dernières lignes. Lorsque  $r = N$ , il existe seulement  $r < M$  inconnues jouant réellement un rôle dans le problème, et  $M - r$  variables dites libres. Il existe par conséquent une infinité de solutions qui diffèrent seulement d'un vecteur du sous-espace  $\text{Ker}(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R}^M, \mathbf{A} \mathbf{x} = \mathbf{0}\}$ . Au contraire, lorsque  $N > M$ , le problème est *sur-déterminé* puisqu'il y a trop d'équations par rapport au nombre d'inconnues et le problème ne possède en général pas de solution.

Lorsque le système n'a pas de solution exacte, il est toujours possible de chercher une solution approchée, telle qu'elle satisfasse au mieux le système d'équations. La solution classique consiste à minimiser

---

<sup>1</sup>En effet, en remplaçant la relation (A.19c) dans (A.19a) on a :  $\mathbf{A} = (\mathbf{A} \mathbf{A}^+)^T \mathbf{A} = (\mathbf{A}^+)^T \mathbf{A}^T \mathbf{A}$ . En prenant la transposée de la dernière égalité, on a  $\mathbf{A}^T = \mathbf{A}^T \mathbf{A} \mathbf{A}^+$ , et puisque la matrice  $\mathbf{A}^T \mathbf{A}$  est inversible lorsque  $r = M$ , on a bien  $\mathbf{A}^+ = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$ .

l'erreur de projection orthogonale par une approche moindres carrés (voir [223, Chap. 3] et [93, Chap. 5]). La solution moindres carrés  $\tilde{\mathbf{X}}$  du système indéterminé caractérisé par l'équation (A.20) satisfait alors les équations *normales*

$$(\mathbf{A}^\top \mathbf{A}) \tilde{\mathbf{X}} = \mathbf{A}^\top \mathbf{B}. \quad (\text{A.21})$$

Si les colonnes de la matrice  $\mathbf{A}$  sont linéairement indépendantes de sorte que  $r = M$ , alors la matrice  $\mathbf{A}^\top \mathbf{A}$  est carrée, symétrique et surtout inversible. Par conséquent l'unique solution du problème moindres carrés donné par l'équation (A.21) est définie par la relation

$$\tilde{\mathbf{X}} = (\mathbf{A}^\top \mathbf{A})^{-1} (\mathbf{A}^\top \mathbf{B}) = \mathbf{A}^+ \mathbf{B}, \quad (\text{A.22})$$

qui fournit la meilleure solution — dans le sens où les valeurs obtenues maximisent la projection.

Cependant, l'évaluation de la solution précédente est la plupart du temps délicate puisque la matrice  $\mathbf{A}$  est souvent très mal conditionnée. Ceci se caractérise par le nombre de conditionnement

$$\kappa(\mathbf{A}) = \sigma_1 / \sigma_r, \quad (\text{A.23})$$

qui correspond au rapport de la valeur singulière maximale sur la valeur singulière minimale introduites toutes deux à l'équation (A.14) dans le cadre de la décomposition en valeurs singulières. Le nombre de conditionnement donne une estimation de la sensibilité de la solution obtenue si une petite perturbation du second membre est pratiquée. Lorsque ce dernier prend des valeurs très élevées, le pseudo-inverse calculé par les relations (A.17) ou (A.22) est en général lui-même très mal conditionné puisqu'il fait intervenir les inverses des valeurs singulières.

Pour améliorer l'approximation de la solution, il est nécessaire de réduire le nombre de conditionnement. Les techniques les plus courantes consistent soit à pratiquer une régularisation de Tikhonov [229] du problème (A.20), soit à effectuer une décomposition en valeurs singulières de la matrice  $\mathbf{A}$  puis à éliminer les lignes et les colonnes des matrices de la décomposition qui sont associées aux valeurs singulières les plus petites. On se concentre ici sur la seconde solution à l'aide de laquelle la matrice  $\mathbf{A}$  peut être approchée par la décomposition en valeurs singulières (A.15). Il est alors possible de définir son pseudo-inverse par la relation (A.17). En outre, si le rang de la matrice  $\mathbf{A}$  vaut  $r = M$  alors la matrice  $\mathbf{A}^\top \mathbf{A}$  est inversible et on a indiqué au paragraphe A.2 que le pseudo-inverse est explicitement défini par  $\mathbf{A}^+ = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top$ . Par conséquent la définition du pseudo-inverse à partir de la décomposition en valeurs singulières coïncide avec celle obtenue avec la formulation du problème moindres carrés (A.22) si bien que le problème du mauvais conditionnement de la matrice  $\mathbf{A}$  n'est pas résolu. Pour y remédier, on peut tenter d'effectuer une approximation du pseudo-inverse dans laquelle n'interviennent que les  $p < r$  plus grandes valeurs singulières associées à la décomposition en valeurs singulières, de sorte que le nombre de conditionnement  $\kappa(\mathbf{A}_p) = \sigma_p / \sigma_1$  soit nettement plus petit. On considère donc la décomposition en valeurs singulières tronquée de la matrice  $\mathbf{A}$  définie à l'équation (A.16) et le pseudo-inverse associé à la matrice  $\mathbf{A}_p$  est alors approché par

$$\mathbf{A}_p^+ = \mathbf{Q}_{2,p} \Sigma_p^+ \mathbf{Q}_{1,p}^\top. \quad (\text{A.24})$$

Le système (A.20) est finalement résolu de façon approchée et la solution est définie par

$$\tilde{\mathbf{X}} = \mathbf{A}_p^+ \mathbf{B}. \quad (\text{A.25})$$

Le choix de l'indice de troncature  $p$  est toutefois délicat et plusieurs techniques ont été développées pour déterminer correctement le pseudo-inverse défini à l'équation (A.24). La solution la plus simple consiste à imposer un nombre de conditionnement  $\kappa_0(\mathbf{A}_p)$  que la matrice à inverser ne doit pas dépasser. Les valeurs singulières les plus petites sont alors éliminées successivement jusqu'à ce que le nombre de conditionnement soit inférieur à  $\kappa_0(\mathbf{A}_p)$ . Cependant, ce choix est arbitraire et surtout la troncature est pratiquée indépendamment du second membre  $\mathbf{B}$ . La méthode de la courbe en L proposée par Hansen [105, 106] offre un critère souvent efficace de sélection de l'indice de troncature  $p$  en fonction du second membre. Le principe s'appuie en partie sur une régularisation de Tikhonov du problème (A.20) pour lequel on cherche à minimiser  $\|\mathbf{A} \mathbf{X} - \mathbf{B}\|^2 + \gamma \|\mathbf{L} \mathbf{X}\|^2$ , où  $\mathbf{L}$  est un opérateur choisi de manière adéquate et  $\gamma$  est le paramètre de régularisation de Tikhonov. La solution du problème régularisé est alors  $\mathbf{X}_\gamma = (\mathbf{A}^\top \mathbf{A} + \gamma^2 \mathbf{L}^\top \mathbf{L})^{-1} (\mathbf{A}^\top \mathbf{B})$ . La difficulté dans la méthode de régularisation de Tikhonov est de déterminer le facteur  $\gamma$  approprié. Hansen [105, 106] a donc suggéré de déterminer la valeur du paramètre  $\gamma$  comme étant celle pour laquelle la



courbe définie par  $\delta = \|\mathbf{L} \mathbf{X}_\gamma\|$  en abscisse et  $\eta = \|\mathbf{A} \mathbf{X}_\gamma - \mathbf{B}\|$  en ordonnée a la plus grande courbure. En effet, cette courbe possède en général une partie relativement verticale et une autre relativement horizontale, d'où le nom de courbe en L. Le point de courbure maximale correspond alors au coin du L. La courbure est d'autant plus marquée lorsqu'on considère la courbe  $(\log(\delta), \log(\eta))$ . Le principe peut être appliqué pour déterminer efficacement l'indice de troncature de la décomposition en valeurs singulières. En effet, si on considère que  $\mathbf{L} = \mathbf{I}$  et qu'on dispose déjà de la décomposition en valeurs singulières de la matrice  $\mathbf{A}$ , alors on peut définir une courbe discrète de coordonnées  $(\delta_p, \eta_p)$  telles que  $\delta_p = \|\mathbf{X}_p\|$  et  $\eta_p = \|\mathbf{A} \mathbf{X}_p - \mathbf{B}\|$ . Le vecteur  $\mathbf{X}_p$  est dans ce cas construit en conservant seulement les  $p$  premiers vecteurs singuliers de la décomposition en valeurs singulières tronquée de sorte que  $\tilde{\mathbf{X}}_p = \mathbf{A}_p^+ \mathbf{B}$ . L'angle maximal sur la courbe discrète indique alors l'indice de troncature  $p$  à conserver. D'autres variantes pour la détermination de l'indice de troncature ont été développées et comparées à la méthode de la courbe en L par Xu [249] entre autres. Par ailleurs Cordier *et al.* [48] ont employé une régularisation de Tikhonov étendue à l'ordre un ou deux pour déterminer les coefficients de leur modèle d'ordre réduit. Le principe consiste simplement à choisir la matrice  $\mathbf{L}$  de sorte qu'elle représente un opérateur de dérivation d'ordre un ou deux. Toutefois les résultats obtenus dans [48] indiquent que la régularisation à l'ordre zéro ( $\mathbf{L} = \mathbf{I}$ ) fournit la meilleure approximation des coefficients du modèle d'ordre réduit.

# Compléments théoriques sur la décomposition orthogonale propre

## Sommaire

<b>B.1</b>	Expression de l'opérateur $R$ dans une base canonique . . . . .	217
<b>B.2</b>	Décomposition spectrale de l'opérateur $R$ . . . . .	218
B.2.1	Détermination du rang de l'opérateur $R$ . . . . .	218
B.2.1.1	Caractère auto-adjoint . . . . .	218
B.2.1.2	Détermination du rang de $R$ . . . . .	218
B.2.2	Décomposition spectrale des opérateurs de Hilbert-Schmidt . . . . .	220
B.2.2.1	Compacité pour les opérateurs de Hilbert-Schmidt . . . . .	220
B.2.2.2	Théorème de décomposition spectrale . . . . .	221
<b>B.3</b>	Équivalence entre la méthode directe et la méthode des clichés en dimension finie . . . . .	222

Dans ce paragraphe, on établit quelques propriétés fondamentales de l'opérateur  $R$  intervenant dans le problème aux valeurs propres (2.16), de manière à justifier l'existence de solutions et au passage celle de la décomposition orthogonale propre. Dans un premier temps, l'expression du terme général de l'opérateur dans une base canonique de l'espace de Hilbert est donnée, puis l'application du théorème spectral pour les opérateurs de Hilbert-Schmidt sera justifiée dans un second paragraphe. Enfin, pour terminer, l'équivalence entre la méthode directe et celle des clichés sera démontrée en dimension finie.

On considère pour la suite une base orthonormée  $\{e_i\}_{i \in I}$  de l'espace de Hilbert  $H$ , telle que l'ensemble  $I$  soit éventuellement infini. On introduit par ailleurs le sous-espace  $S_X = \text{Vect}\{x^{(m)}, m = 1, \dots, M\}$  correspondant à l'espace engendré par l'ensemble  $X$  des clichés. Le sous-espace  $S_X$  de dimension finie  $r = \dim S_X$  est alors entièrement caractérisé par une base orthonormée notée  $\{f_i\}_{i=1}^r$ .

## B.1 Expression de l'opérateur $R$ dans une base canonique

Les clichés  $x^{(m)}$  — et plus généralement tout vecteur  $y$  de  $H$  — se décomposent sur la base  $\{e_i\}_{i \in I}$  de l'espace  $H$  au moyen des relations

$$x^{(m)} = \sum_{i \in I} \eta_i^{(m)} e_i \quad \forall m \in \llbracket 1; M \rrbracket, \tag{B.1a}$$

$$y = \sum_{i \in I} \mu_i e_i \quad \forall y \in H. \tag{B.1b}$$

En utilisant les décompositions (B.1a) et (B.1b) sur la base  $\{e_i\}_{i \in I}$ , le produit scalaire entre un cliché  $x^{(m)}$  et un vecteur quelconque  $y \in H$  s'écrit  $\langle x^{(m)}, y \rangle = \sum_{i,j \in I} \eta_i^{(m)} \mu_j \langle e_i, e_j \rangle = \sum_{i \in I} \eta_i^{(m)} \mu_i$  puisque  $\langle e_i, e_j \rangle = \delta_{ij}$ .

L'image  $Ry$  d'un élément quelconque  $y \in H$  par l'opérateur  $R$  défini à l'équation (2.11) s'exprime alors sur la base canonique avec la relation

$$Ry = E \left[ \langle x^{(m)}, y \rangle x^{(m)} \right] = \sum_{m=1}^M \alpha_m \left( \sum_{i \in I} \eta_i^{(m)} \mu_i \right) \sum_{j \in I} \eta_j^{(m)} e_j = \sum_{i, j \in I} \left( \sum_{m=1}^M \alpha_m \eta_i^{(m)} \eta_j^{(m)} \right) \mu_i e_j. \quad (\text{B.2})$$

Pour mettre en évidence le terme général de l'opérateur  $R$ , on considère uniquement la  $j$ -ième composante du vecteur  $Ry$  qui correspond à  $\langle Ry, e_j \rangle$ . En réorganisant les termes on obtient l'expression  $\langle Ry, e_j \rangle = \sum_{i \in I} \sum_{m=1}^M \alpha_m \eta_i^{(m)} \eta_j^{(m)} \mu_i$  qui permet d'isoler le terme général de l'opérateur  $R$  de la partie relative au vecteur  $y$  lorsque les vecteurs sont décomposés sur la base  $\{e_i\}_{i \in I}$ . Le terme général qui s'écrit donc

$$R_{ij} = \sum_{m=1}^M \alpha_m \eta_i^{(m)} \eta_j^{(m)} = E \left[ \eta_i^{(m)} \eta_j^{(m)} \right] \quad (\text{B.3})$$

est défini par la moyenne du produit des composantes  $i$  et  $j$  des clichés exprimés sur la base  $\{e_i\}_{i \in I}$ . Lorsqu'une base spatiale est employée,  $R_{ij}$  correspond à la fonction de corrélation spatiale entre un point  $i$  et un point  $j$  de l'espace.

## B.2 Décomposition spectrale de l'opérateur $R$

L'existence de la décomposition orthogonale propre d'un ensemble  $X$  de clichés repose sur l'utilisation de la décomposition spectrale pour les opérateurs de Hilbert-Schmidt. Afin de justifier que l'opérateur de corrélation vérifie les propriétés de ce type d'opérateurs, il est nécessaire dans un premier temps d'établir le rang de l'opérateur de corrélation avant de pouvoir invoquer le théorème de décomposition spectrale.

### B.2.1 Détermination du rang de l'opérateur $R$

#### B.2.1.1 Caractère auto-adjoint

**Définition B.1 :** Soit  $E$  un espace de Hilbert et  $T \in \mathcal{L}(E)$  une application linéaire de  $E$  dans lui-même. On dit que l'opérateur  $T$  est normal si il commute avec son adjoint noté  $T^*$ , i. e.  $TT^* = T^*T$ . On dit par ailleurs que  $T$  est hermitien (ou auto-adjoint, ou simplement symétrique si  $\mathbb{K} = \mathbb{R}$ ) si il coïncide avec son adjoint :  $T = T^*$ . Cette propriété se traduit pour tout  $y, z \in E$  par

$$\langle Ty, z \rangle = \langle y, Tz \rangle, \quad (\text{B.4})$$

ce qui signifie par conséquent que tout opérateur hermitien est normal.

On vérifie alors que cette définition s'applique à l'opérateur  $R$  pour  $y, z \in H$  quelconques. En développant l'expression de  $R$  et en utilisant la propriété de commutation de l'opérateur de moyenne  $E[\cdot]$  avec le produit scalaire qui avait été donnée à l'équation (2.5b), on obtient

$$\begin{aligned} \langle Ry, z \rangle &= \langle E[\langle x^{(m)}, y \rangle x^{(m)}], z \rangle = E[\langle \langle x^{(m)}, y \rangle x^{(m)}, z \rangle] = E[\langle x^{(m)}, y \rangle \langle x^{(m)}, z \rangle] \\ &= E[\langle \langle x^{(m)}, z \rangle x^{(m)}, y \rangle] = \langle E[\langle x^{(m)}, z \rangle x^{(m)}], y \rangle = \langle y, Rz \rangle. \end{aligned} \quad (\text{B.5})$$

La dernière égalité provient du fait que le produit scalaire est une application symétrique et par conséquent l'opérateur  $R$  est bien auto-adjoint.

#### B.2.1.2 Détermination du rang de $R$

**Définition B.2 :** Soit  $E, F$  deux espaces vectoriels sur un corps  $\mathbb{K}$  et  $T \in \mathcal{L}(E, F)$  une application linéaire de  $E$  dans  $F$ . Le rang de l'opérateur  $T$  est égal à la dimension de l'image  $T(E)$  de  $E$  par  $T$ , définie par  $\text{Im}(T) = \{Tx, x \in E\}$ . On dit alors que  $T$  est de rang fini si et seulement si la dimension de l'image  $\text{Im}(T)$  est finie.

D'après la définition précédente, le rang de l'opérateur  $R \in \mathcal{L}(H)$  est la dimension de son image  $\text{Im}(R)$ . Or celle-ci est étroitement liée à la dimension de l'espace  $S_X$  engendré par l'ensemble  $X$  des clichés étant donné la définition (2.11) de  $R$ . Dans ce qui suit, on établit que le rang de  $R$  noté pour l'instant  $\text{rg}(R)$  est égal à la dimension  $r$  de l'espace  $S_X$ . À cet effet, on commence par démontrer le lemme suivant.

**Lemme B.1 :** Soit  $y \in S_X$ . Si  $\langle x^{(m)}, y \rangle = 0$  pour tout  $m = 1, \dots, M$  alors  $y = 0$ .

*Démonstration.* Puisque  $y \in S_X = \text{Vect} \{x^{(m)}, m = 1, \dots, M\}$ , il est possible de l'exprimer comme une combinaison linéaire des clichés qui engendrent  $S_X$  :  $y = \sum_{i=1}^M a_i x^{(i)}$ , où les  $a_i$  sont des scalaires appartenant à  $\mathbb{K}$ . La somme pondérée par les coefficients  $a_m$  de la relation  $\langle x^{(m)}, y \rangle = 0$  du lemme qui est supposée vraie pour tout  $m \in \llbracket 1; M \rrbracket$  donne alors  $\sum_{m=1}^M a_m \langle x^{(m)}, y \rangle = 0$ . Finalement en introduisant la somme dans le premier membre du produit scalaire on a  $\langle \sum_{m=1}^M a_m x^{(m)}, y \rangle = 0$ . Cette dernière relation correspond en fait au carré de la norme du vecteur  $y$  qui doit être nulle ( $\langle y, y \rangle = \|y\|^2$ ). Ceci implique que le vecteur  $y$  est nécessairement le vecteur nul puisque le produit scalaire est une application définie positive.  $\square$

**Propriété B.1 :** Soit  $r = \dim S_X$ . Alors  $\text{rg}(R) = r$  et  $\text{Im}(R) = S_X$ .

*Démonstration.* On peut montrer cette propriété en deux étapes au cours desquelles l'image  $\text{Im}(R) = \text{Vect}_{i \in I} \{Re_i\}$  est encadrée par deux ensembles de dimension identiquement égale à  $r$ .

$\text{Im}(R) \subset S_X$

En développant l'expression de l'image par  $R$  du vecteur de base  $e_i$  on a  $Re_i = E[\beta_{mi} x^{(m)}]$  avec  $\beta_{mi} = \langle x^{(m)}, e_i \rangle \in \mathbb{K}$ . Toute image par  $R$  d'un des vecteurs de la base  $\{e_i\}_{i \in I}$  s'exprime alors comme une combinaison linéaire des clichés lorsqu'on développe l'expression de l'opérateur de moyenne  $E[\cdot]$ . Ainsi, pour tout  $i \in I$

$$Re_i = \sum_{m=1}^M \alpha_m \beta_{mi} x^{(m)} = \sum_{m=1}^M \gamma_{mi} x^{(m)}, \quad (\text{B.6})$$

et par conséquent  $Re_i \in \text{Vect}_{m \in \llbracket 1; M \rrbracket} \{x^{(m)}\} = S_X$ , ce qui signifie que l'image de l'opérateur  $R$  est incluse dans l'espace engendré par les clichés :

$$\text{Im}(R) \subset S_X. \quad (\text{B.7})$$

On note alors

$$r = \dim S_X \quad (\text{B.8})$$

la dimension de l'espace engendré par l'ensemble des clichés qui est clairement finie puisque la famille des  $\{x^{(m)}\}_{m=1}^M$  est une collection finie. On impose alors qu'au moins un cliché soit non nul de sorte que  $r \geq 1$ . Par ailleurs, si la famille des clichés est libre, alors le rang atteint sa valeur maximale  $r = M$  et dans ce cas  $\{x^{(m)}\}_{m=1}^M$  est une base de  $S_X$ .

$\text{Vect}_{j \in \llbracket 1; r \rrbracket} (Rf_j) \subset \text{Im}(R)$

Soit une base orthonormée  $\{f_j\}_{j=1}^r$  de l'espace  $S_X$ . L'inclusion est triviale puisque  $\text{Im}(R) = \{Ry, \forall y \in H\}$  et  $f_j$  est un élément de  $S_X$  donc par conséquent de  $H$ . On a donc immédiatement

$$\text{Vect}_{j \in \llbracket 1; r \rrbracket} (Rf_j) \subset \text{Im}(R). \quad (\text{B.9})$$

On cherche à présent à montrer que l'espace engendré par l'image de la base  $\{f_j\}_{j=1}^r$  par  $R$  est aussi de dimension  $r$ . Pour cela, on montre que la famille  $\{Rf_j\}_{j=1}^r$  est une famille libre. Les clichés se décomposent sur la base  $\{f_j\}_{j=1}^r$  puisqu'ils appartiennent à  $S_X$ , ce qui conduit à  $x^{(m)} = \sum_{i=1}^r \eta_i^{(m)} f_i$ . Ainsi, le terme  $Rf_j$  se développe de la même façon que pour obtenir la relation (B.2) en :  $Rf_j = \sum_{m=1}^M \alpha_m (\sum_{i=1}^r \eta_i^{(m)} \eta_j^{(m)} f_i)$ . On considère alors la famille de scalaires  $(a_1, \dots, a_r)$  telle que  $\sum_{j=1}^r a_j Rf_j = 0$ . La famille  $\{Rf_j\}_{j=1}^r$  est donc libre à condition que tous les coefficients  $a_j$  soient nuls. L'intervention des symboles de sommation conduit alors à l'égalité

$$\sum_{i=1}^r \left( \sum_{m=1}^M \alpha_m \left( \sum_{j=1}^r a_j \eta_i^{(m)} \eta_j^{(m)} \right) \right) f_i = 0. \quad (\text{B.10})$$

Or les  $\{f_j\}_{j=1}^r$  forment une base de  $S_X$  et par conséquent une famille libre, ce qui permet d'écrire  $\sum_{m=1}^M \alpha_m (\sum_{j=1}^r a_j \eta_i^{(m)} \eta_j^{(m)}) = 0$  pour tout  $i = 1, \dots, r$ . Cette dernière relation est alors multipliée par la somme des coefficients  $a_i$  et la réorganisation des sommes donne

$$\sum_{m=1}^M \alpha_m \left( \sum_{i=1}^r a_i \eta_i^{(m)} \right) \left( \sum_{j=1}^r a_j \eta_j^{(m)} \right) = 0. \quad (\text{B.11})$$

Finalement on définit le vecteur  $a = \sum_{i=1}^r a_i f_i \in S_X$  comme la combinaison linéaire des coefficients introduits comme pondération de la famille  $Rf_j$  dans la relation (B.10). Les deux sommes apparaissant dans la dernière expression de l'équation (B.11) s'interprètent alors comme le produit scalaire entre le vecteur  $a$  et un cliché  $x^{(m)}$  puisque

$$\langle a, x^{(m)} \rangle = \left\langle \sum_{i=1}^r a_i f_i, \sum_{j=1}^r \eta_j^{(m)} f_j \right\rangle = \sum_{i,j=1}^r a_i \eta_j^{(m)} \underbrace{\langle f_i, f_j \rangle}_{=\delta_{ij}} = \sum_{i=1}^r a_i \eta_i^{(m)}. \quad (\text{B.12})$$

Par conséquent en remplaçant la relation (B.12) dans (B.11) on obtient :  $\sum_{m=1}^M \alpha_m \langle a, x^{(m)} \rangle^2 = 0$ . Or la famille des  $\alpha_m$  qui correspond aux poids de l'opérateur de moyenne  $E[\cdot]$  ne contient que des termes strictement positifs, si bien que la condition précédente se réduit à  $\langle a, x^{(m)} \rangle = 0$  pour tout  $m = 1, \dots, M$ . D'après le lemme B.1, cela signifie que  $a = 0$  puisque ce vecteur est contenu dans  $S_X$  par définition. Pour conclure, il suffit de remarquer que  $a = \sum_{i=1}^r a_i f_i$  et puisque les  $\{f_j\}_{j=1}^r$  forment une base de  $S_X$  et par conséquent une famille libre on a bien  $a_i = 0$  pour tout  $i = 1, \dots, r$ .

La famille des  $Rf_j$  est donc bien une famille libre, et puisque c'est trivialement une famille génératrice de  $\text{Vect}_{j \in \llbracket 1; r \rrbracket} \{Rf_j\}$ , il s'agit d'une base. L'espace  $\text{Vect}_{j \in \llbracket 1; r \rrbracket} \{Rf_j\}$  est donc de dimension égale au cardinal de la base, soit  $r$ . L'image de  $R$  est donc encadrée par les relations d'inclusion (B.7) et (B.9) telle que

$$\text{Vect}_{j \in \llbracket 1; r \rrbracket} (Rf_j) \subset \text{Im}(R) \subset S_X. \quad (\text{B.13})$$

Finalement, puisque les dimensions de  $\text{Vect}_{j \in \llbracket 1; r \rrbracket} \{Rf_j\}$  et  $S_X$  sont identiquement égales à  $r$ , la dimension de  $\text{Im}(R)$  est nécessairement égale aussi à  $r$ , ce qui signifie que  $\text{rg}(R) = r$ . Le rang de l'opérateur  $R$  est alors compris dans l'intervalle  $\llbracket 1; M \rrbracket$  qui dépend du nombre  $M$  de clichés mais aussi de leur dépendance linéaire. En outre, puisque  $\text{Im}(R) \subset S_X$  et que les dimensions sont égales, on a

$$\text{Im}(R) = S_X. \quad (\text{B.14})$$

□

### Remarque B.1 : Expression des modes POD en fonction des clichés

La relation (B.6) établie pour les vecteurs de base  $e_i \in I$  peut être généralisée à tout vecteur  $y \in H$ , de sorte que l'image  $Ry$  se décompose sur la famille des clichés de l'ensemble  $X$  qui ne forme pas nécessairement une base — sauf si les clichés sont linéairement indépendants. Cette décomposition s'applique en particulier aux modes POD  $\varphi^{(j)} \in S_X \subset H$ , ce qui permet d'écrire  $R\varphi^{(j)} = \sum_{m=1}^M \gamma_m^{(j)} x^{(m)}$  en notant  $\gamma_m^{(j)} = \alpha_m \langle x^{(m)}, \varphi^{(j)} \rangle$  les coefficients de cette décomposition. Par ailleurs, puisque les modes POD sont solutions du problème aux valeurs propres  $R\varphi^{(j)} = \lambda_j \varphi^{(j)}$ , la relation précédente conduit à

$$\varphi^{(j)} = \sum_{m=1}^M \frac{\gamma_m^{(j)}}{\lambda_j} x^{(m)}. \quad (\text{B.15})$$

On retrouve ainsi la décomposition (2.20) en notant  $c_m^{(j)} = \gamma_m^{(j)} / \lambda_j$  ce qui est possible pour tout  $j = 1, \dots, r$  puisque dans ce cas les modes POD sont associés à des valeurs propres  $\lambda_j$  non nulles. La relation précédente justifie donc la décomposition (2.20) qui a été supposée pour formuler le problème aux valeurs propres annexe de la méthode des clichés.

## B.2.2 Décomposition spectrale des opérateurs de Hilbert-Schmidt

### B.2.2.1 Compacité pour les opérateurs de Hilbert-Schmidt

Puisqu'on a établi auparavant que l'opérateur  $R$  était de rang fini, les propriétés suivantes tirées de [243, p. 437] s'appliquent à l'opérateur  $R$ .

**Propriété B.2 :** Tout opérateur  $T \in \mathcal{L}(E, F)$  de rang fini est de Hilbert-Schmidt.

**Propriété B.3 :** Les opérateurs de Hilbert-Schmidt sont compacts. En outre, le sous-espace des opérateurs de rang fini est dense dans l'ensemble  $\mathcal{H}(E, F)$  des opérateurs de Hilbert-Schmidt pour la norme de Hilbert-Schmidt.

### B.2.2.2 Théorème de décomposition spectrale

Le présentation qui suit reprend les définitions et démonstrations établies en particulier dans les ouvrages de Schwartz [205], Chap. XXIII, §7 et 8 et de Wagschal [243], Chap. 3, §G. Le lecteur est en particulier invité à s'y reporter pour les démonstrations des propriétés et du théorème de décomposition spectral qui suivent.

Soit  $E$  un espace de Hilbert sur un corps  $\mathbb{K}$  et  $T \in \mathcal{L}(E)$  une application linéaire de  $E$  dans lui-même. On considère alors l'opérateur  $T_\lambda = \lambda \text{Id}_E - T$ , où  $\text{Id}_E$  est l'identité de l'espace  $E$  et  $\lambda \in \mathbb{K}$  est une valeur dite *régulière* si  $T_\lambda$  est un isomorphisme de  $E$  dans  $E$ . L'ensemble  $\rho(T)$  des valeurs régulières est appelé *l'ensemble résolvant* de  $T$  et si  $\lambda$  est une valeur régulière, l'application  $(\lambda \text{Id}_E - T)^{-1}$  est la résolvante de  $T$  au point  $\lambda$ .

Le complémentaire de l'ensemble résolvant noté  $\sigma(T)$  porte le nom de *spectre* de  $T$ . Lorsque  $\lambda \in \sigma(T)$  on parle de *valeur spectrale* de l'opérateur  $T$ . Ceci signifie que  $T_\lambda$  n'est pas une bijection — *i. e.*  $\lambda \text{Id}_E - T$  n'est pas inversible — et en particulier si  $T_\lambda$  n'est pas injectif,  $\lambda$  est une *valeur propre*. On note alors  $E_\lambda \neq \emptyset$  le sous-espace propre associé à la valeur propre  $\lambda$  défini par

$$E_\lambda = \text{Ker } T_\lambda = \{x \in E, \lambda x - Tx = 0\}. \quad (\text{B.16})$$

Les éléments de  $E_\lambda$  correspondent alors aux vecteurs propres associés à  $\lambda$ . On notera par ailleurs  $\mathcal{P}_\lambda$  le projecteur orthogonal de  $E$  sur le sous-espace propre  $E_\lambda$ .

**Propriété B.4 :** Soit  $T \in \mathcal{L}(E)$  un opérateur hermitien, alors son spectre  $\sigma(T)$  est réel.

La propriété précédente tirée de [243, p.426] s'applique à l'opérateur  $R$  puisqu'on a démontré au paragraphe B.2.1 qu'il est auto-adjoint (*i. e.* hermitien). Ainsi, l'opérateur  $R$  possède un spectre réel. Par ailleurs, on peut montrer que l'opérateur  $R$  est positif. Il suffit pour cela de montrer que pour tout  $y \in H$ ,  $\langle Ry, y \rangle \geq 0$ . Le développement du produit scalaire donne

$$\begin{aligned} \langle Ry, y \rangle &= E [\langle x^{(m)}, y \rangle x^{(m)}, y \rangle] = E [\langle x^{(m)}, y \rangle \langle x^{(m)}, y \rangle] \\ &= E [\langle x^{(m)}, y \rangle^2] = \sum_{m=1}^M \alpha_m \underbrace{\langle x^{(m)}, y \rangle^2}_{\geq 0}, \end{aligned} \quad (\text{B.17})$$

et l'opérateur  $R$  est donc positif si et seulement si les poids  $\alpha_m$  associés à l'opérateur de moyenne sont tous positifs, ce qui est en pratique le cas. On ne dispose pas de résultats sur le caractère défini de l'opérateur dans le cas général. Tout au plus, on se contentera de dire que  $R$  est défini non négatif — *i. e.* positif mais pas forcément défini. On a alors l'équivalence suivante [243, p.430-31] :

**Propriété B.5 :** Soit  $T \in \mathcal{L}(E)$  un opérateur hermitien compact. Alors :

$$T \text{ positif} \iff \sigma(T) \in [0; +\infty[.$$

L'opérateur  $R$  intervenant dans le problème aux valeurs propres (2.16) est donc finalement hermitien (ou symétrique puisque  $\mathbb{K} = \mathbb{R}$ ), compact et positif. On en déduit donc que son spectre est contenu dans la droite des réels positifs  $\mathbb{R}_+$ . Finalement, le théorème de décomposition spectrale [205, p.427] fournit des résultats importants concernant l'orthogonalité et la quantité de valeurs propres associées à l'opérateur  $R$  impliqué dans la définition de la décomposition orthogonale propre.

### Théorème B.1 : Théorème de décomposition spectrale

Soient  $E$  un espace de Hilbert et  $T \in \mathcal{L}(E)$  un opérateur compact symétrique si  $\mathbb{K} = \mathbb{R}$  ou hermitien si  $\mathbb{K} = \mathbb{C}$ . Alors :

- (i) l'ensemble des valeurs propres  $\sigma(T)$  est fini ou peut être rangé en une suite convergant vers 0, 0 appartenant ou non au spectre ;
- (ii) pour chaque valeur propre  $\lambda \in \sigma(T)$ , le sous-espace propre associé  $E_\lambda$  est de dimension finie. De plus, les sous-espaces propres sont orthogonaux deux à deux ;
- (iii)  $E$  est la somme hilbertienne des sous-espaces propres  $E_\lambda$  ; la décomposition porte alors le nom de décomposition spectrale de  $E$  relativement à  $T$  et s'écrit

$$E = \bigoplus_{\lambda \in \sigma(T)} E_\lambda ; \quad (\text{B.18})$$

- (iv) les familles  $(\lambda \mathcal{P}_\lambda)_{\lambda \in \sigma(T)}$  et  $(\bar{\lambda} \mathcal{P}_\lambda)_{\lambda \in \sigma(T)}$  sont sommables dans  $\mathcal{L}(E)$  et de somme  $T$  et  $T^*$  telles que

$$T = \sum_{\lambda \in \sigma(T)} \lambda \mathcal{P}_\lambda \quad \text{et} \quad T^* = \sum_{\lambda \in \sigma(T)} \bar{\lambda} \mathcal{P}_\lambda. \quad (\text{B.19})$$

Le théorème spectral fournit un certain nombre d'informations intéressantes, à commencer par le fait qu'il existe au plus une infinité dénombrable de valeurs propres pouvant être organisées en une série convergant vers 0. Toutefois, on a montré auparavant que  $R$  était un opérateur de Hilbert-Schmidt de rang fini. Par conséquent, le nombre de valeurs propres non nulles est dans ce cas fini est égal à  $r$ , *i. e.* la dimension de l'espace engendré par l'ensemble  $X$  des clichés. Ainsi, l'opérateur  $R$  est caractérisé par  $r$  valeurs propres réelles positives.

Le second point indique que les sous-espaces propres sont orthogonaux entre eux, si bien que les vecteurs propres associés à deux valeurs propres distinctes sont orthogonaux entre eux. Enfin, la décomposition spectrale (B.18) signifie que n'importe quel vecteur de l'espace  $E$  s'écrit de façon unique comme combinaison linéaire des vecteurs propres du spectre  $\sigma(T)$ . Cependant, les vecteurs propres associés aux valeurs propres non nulles de l'opérateur  $R$  ne permettent de représenter que l'ensemble des éléments du sous-espace engendré par les clichés<sup>1</sup>. En particulier les clichés se décomposent bien sur la base de ces vecteurs propres et la décomposition (2.17) existe bien.

### B.3 Équivalence entre la méthode directe et la méthode des clichés en dimension finie

Dans ce paragraphe, on cherche à établir le lien d'équivalence entre la méthode directe et celle des clichés sans recourir à l'hypothèse d'indépendance linéaire des clichés qui a été émise pour établir la méthode des clichés au paragraphe 2.2.2. Pour simplifier, on se place dans un espace de Hilbert de dimension finie  $N = \dim(H) < \infty$  de manière à pouvoir faire appel à des concepts d'algèbre linéaire matricielle. Des résultats d'équivalence entre la décomposition orthogonale propre et décomposition en valeurs singulières pourront aussi être trouvés dans l'article de Liang *et al.* [144] et le rapport de Volkwein [241].

Soit  $\{e_i\}_{i \in [1;N]}$  une base orthonormée de l'espace  $H$  telle que les clichés se décomposent en  $x^{(m)} = \sum_{i=1}^N \eta_i^{(m)} e_i$ . En suivant la définition (A.3), on définit la matrice des clichés  $\mathbf{X} = \text{Mat}_{e_i}(X)$  comportant les coordonnées des clichés de l'ensemble  $X$  exprimés dans la base  $\{e_i\}_{i \in [1;N]}$ . On obtient ainsi la matrice

$$\mathbf{X} = \begin{bmatrix} \eta_1^{(1)} & \cdots & \eta_1^{(M)} \\ \vdots & & \vdots \\ \eta_N^{(1)} & \cdots & \eta_N^{(M)} \end{bmatrix}, \quad (\text{B.20})$$

dont le rang correspond à la dimension de l'espace engendré par ses colonnes, *i. e.* à la dimension de  $S_X$ . D'après la notation introduite en (B.8), on a  $\text{rg}(\mathbf{X}) = r$  avec  $1 \leq r \leq M$ .

On introduit pour la suite les clichés  $\tilde{x}^{(m)} = \sqrt{\alpha_m} x^{(m)} = \sum_{i=1}^N \sqrt{\alpha_m} \eta_i^{(m)} e_i$ . Les poids de pondération associés à l'opérateur de moyenne sont regroupés dans la matrice diagonale  $\alpha = \text{diag}(\alpha_1, \dots, \alpha_M)$  et on

<sup>1</sup>Holmes *et al.* [112] montrent dans le cas de  $H = L^2(\Omega)$  que la base des vecteurs propres permet de représenter toutes les fonctions contenues dans le sous-espace engendré par les vecteurs propres, et même n'importe quelle fonction indistinguable au sens de la norme de  $L^2(\Omega)$  d'un cliché  $x^{(m)}$ .

définit la matrice  $\tilde{\alpha} = \text{diag}(\sqrt{\alpha_1}, \dots, \sqrt{\alpha_M})$  comme au paragraphe 2.2.2. Les clichés  $\tilde{x}^{(m)}$  se déduisent donc des  $x^{(m)}$  par la relation matricielle  $\tilde{\mathbf{X}} = \mathbf{X} \tilde{\alpha}$ . La matrice  $\tilde{\mathbf{X}}$  est de même rang que la matrice des clichés  $\mathbf{X}$  puisque tous les poids de pondération associés à l'opérateur de moyenne sont strictement positifs.

D'après l'expression (B.3) établie au paragraphe B.1, l'opérateur linéaire  $R$  peut s'écrire — en dimension finie — comme une matrice de taille  $N \times N$  définie par le produit matriciel

$$\mathbf{R} = \mathbf{X} \boldsymbol{\alpha} \mathbf{X}^\top = \tilde{\mathbf{X}} \tilde{\mathbf{X}}^\top. \quad (\text{B.21})$$

On considère par ailleurs l'opérateur

$$\mathbf{R}^* = \tilde{\alpha}^\top \mathbf{X}^\top \mathbf{X} \tilde{\alpha} = \tilde{\mathbf{X}}^\top \tilde{\mathbf{X}} \quad (\text{B.22})$$

qui correspond à l'opérateur de la méthode des clichés établi au paragraphe 2.2.2 lorsque le produit scalaire est celui de la base canonique. Dans ce qui suit, on cherche alors à montrer que les valeurs et les vecteurs propres associés à  $\mathbf{R}^*$  permettent de retrouver de manière équivalente les valeurs propres et les modes POD correspondant à l'opérateur  $\mathbf{R}$  de la méthode directe. L'objectif est donc de montrer le théorème suivant.

**Théorème B.2 : Équivalence entre la méthode directe et la méthode des clichés**

Soit  $\mathbf{X}$  la matrice représentative des clichés dans un espace de Hilbert de dimension finie  $N$  et  $\varphi^{(j)}$ ,  $j = 1, \dots, r$  les modes POD associés aux clichés. Alors

$$\begin{aligned} \varphi^{(j)} \text{ solution de } \mathbf{R} \varphi^{(j)} = \lambda_j \varphi^{(j)} \text{ pour tout } j \in \llbracket 1; r \rrbracket \text{ avec } \mathbf{R} = \tilde{\mathbf{X}} \tilde{\mathbf{X}}^\top \\ \Downarrow \\ \varphi^{(j)} = 1/\sqrt{\lambda_j} \tilde{\mathbf{X}} \mathbf{d}^{(j)} \text{ avec } \mathbf{d}^{(j)} \text{ solution de } \mathbf{R}^* \mathbf{d}^{(j)} = \lambda_j \mathbf{d}^{(j)} \text{ pour tout } j \in \llbracket 1; r \rrbracket \text{ et } \mathbf{R}^* = \tilde{\mathbf{X}}^\top \tilde{\mathbf{X}}. \end{aligned}$$

Pour établir l'équivalence des deux méthodes, une décomposition en valeurs singulières de la matrice  $\tilde{\mathbf{X}}$  est pratiquée. Seuls les vecteurs singuliers associés aux valeurs propres non nulles sont conservés, de sorte que la décomposition s'écrit (voir l'équation (A.15) au paragraphe A.2)

$$\tilde{\mathbf{X}} = \mathbf{Q}_{1,r} \boldsymbol{\Sigma}_r \mathbf{Q}_{2,r}^\top. \quad (\text{B.23})$$

La matrice orthogonale  $\mathbf{Q}_{1,r} \in \mathcal{M}(N, r)$  (resp.  $\mathbf{Q}_{2,r} \in \mathcal{M}(M, r)$ ) contient les vecteurs singuliers à gauche (resp. à droite) de  $\tilde{\mathbf{X}}$  et la matrice diagonale  $\boldsymbol{\Sigma}_r \in \mathcal{M}(r, r)$  regroupe sur sa diagonale principale les valeurs singulières non nulles  $\sigma_1 \geq \dots \geq \sigma_r > 0$ .

L'introduction de la décomposition en valeurs singulières (B.23) de la matrice  $\tilde{\mathbf{X}}$  dans l'expression (B.21) de  $\mathbf{R}$  conduit à  $\mathbf{R} = \mathbf{Q}_{1,r} \boldsymbol{\Sigma}_r \mathbf{Q}_{2,r}^\top \mathbf{Q}_{2,r} \boldsymbol{\Sigma}_r \mathbf{Q}_{1,r}^\top = \mathbf{Q}_{1,r} \boldsymbol{\Sigma}_r^2 \mathbf{Q}_{1,r}^\top$  puisque  $\mathbf{Q}_{2,r}$  est une matrice orthogonale. On réécrit de préférence la relation précédente sous la forme d'un problème aux valeurs propres en multipliant à droite par  $\mathbf{Q}_{1,r}$  et en utilisant l'orthogonalité des vecteurs singuliers à gauche ce qui donne

$$\mathbf{R} \mathbf{Q}_{1,r} = \mathbf{Q}_{1,r} \boldsymbol{\Sigma}_r^2. \quad (\text{B.24})$$

Par conséquent le problème (B.24) est identique à celui de la méthode directe qui s'écrit en notation matricielle  $\mathbf{R} \boldsymbol{\Phi} = \boldsymbol{\Phi} \boldsymbol{\Lambda}$  avec  $\boldsymbol{\Phi} = [\varphi^{(1)} \dots \varphi^{(r)}]$  et où  $\boldsymbol{\Lambda} \in \mathcal{M}(r, r)$  est la matrice carrée diagonale contenant les  $r$  valeurs propres associées aux modes POD. Par identification des termes, on obtient donc

$$\boldsymbol{\Phi} = \mathbf{Q}_{1,r} \quad \text{et} \quad \boldsymbol{\Lambda} = \boldsymbol{\Sigma}_r^2. \quad (\text{B.25})$$

Ainsi, les  $r$  premiers vecteurs singuliers à gauche  $\mathbf{Q}_{1,r}$  de la décomposition en valeurs singulières de  $\tilde{\mathbf{X}}$  correspondent aux modes POD  $\boldsymbol{\Phi}$  et le carré des valeurs singulières de  $\tilde{\mathbf{X}}$  s'identifie aux valeurs propres du problème POD formulé avec la méthode directe.

L'opérateur  $\mathbf{R}^*$  se transforme de manière similaire en utilisant la décomposition en valeurs singulières (B.23) de  $\tilde{\mathbf{X}}$  qui, remplacée dans (B.22), donne  $\mathbf{R}^* = \tilde{\mathbf{X}}^\top \tilde{\mathbf{X}} = \mathbf{Q}_{2,r} \boldsymbol{\Sigma}_r \mathbf{Q}_{1,r}^\top \mathbf{Q}_{1,r} \boldsymbol{\Sigma}_r \mathbf{Q}_{2,r}^\top = \mathbf{Q}_{2,r} \boldsymbol{\Sigma}_r^2 \mathbf{Q}_{2,r}^\top$ . Encore une fois, on ramène l'expression précédente à celle d'un problème aux valeurs propres en multipliant à droite par  $\mathbf{Q}_{2,r}$ , ce qui conduit à

$$\mathbf{R}^* \mathbf{Q}_{2,r} = \mathbf{Q}_{2,r} \boldsymbol{\Sigma}_r^2. \quad (\text{B.26})$$



Il s'ensuit que les  $r$  premiers vecteurs singuliers à droite  $\mathbf{Q}_{2,r}$  de la décomposition en valeurs singulières de  $\tilde{\mathbf{X}}$  correspondent aux vecteurs propres  $\mathbf{d}^{(j)}$  du problème aux valeurs propres  $\mathbf{R}^* \mathbf{D} = \mathbf{D} \mathbf{\Lambda}$ , où  $\mathbf{D} \in \mathcal{M}(M, r)$  est la matrice  $\mathbf{D} = [\mathbf{d}^{(1)} \dots \mathbf{d}^{(r)}]$  qui regroupe dans ses colonnes les vecteurs propres. Par ailleurs, le carré des valeurs singulières de  $\tilde{\mathbf{X}}$  s'identifie aux valeurs propres du problème POD formulé avec la méthode des clichés et on a donc les relations

$$\mathbf{D} = \mathbf{Q}_{2,r} \quad \text{et} \quad \mathbf{\Lambda} = \Sigma_r^2. \quad (\text{B.27})$$

Les opérateurs  $\mathbf{R}$  et  $\mathbf{R}^*$  possèdent donc les mêmes valeurs propres  $\sigma_i^2 = \lambda_i$ , ce qui démontre l'équivalence des méthodes directe et des clichés. Les vecteurs propres correspondent alors aux modes POD avec la méthode directe tandis qu'avec la méthode des clichés les vecteurs propres correspondent aux coefficients de pondération de la décomposition des modes POD sur les clichés.

Finalement, à partir de la matrice des clichés pondérés  $\tilde{\mathbf{X}} = \mathbf{X} \tilde{\boldsymbol{\alpha}}$ , les modes POD et les valeurs propres associées peuvent être obtenus de façon équivalente en résolvant soit le problème de la méthode directe, soit celui de la méthode des clichés, soit en recourant à une décomposition en valeurs singulières. Les relations entre les modes POD  $\mathbf{\Phi}$ , les valeurs propres  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_r)$  et les coefficients  $\mathbf{D}$  permettant de décomposer les modes POD sur les clichés sont résumées dans le tableau ci-dessous lorsque ces quantités sont calculées par les trois méthodes possibles.

	Méthode directe	Méthode des clichés	SVD
Opérateur de corrélation	$\mathbf{R} = \tilde{\mathbf{X}} \tilde{\mathbf{X}}^\top$	$\mathbf{R}^* = \tilde{\mathbf{X}}^\top \tilde{\mathbf{X}}$	—
Décomposition	$\mathbf{R} \mathbf{\Phi} = \mathbf{\Phi} \mathbf{\Lambda}$	$\mathbf{R}^* \mathbf{D} = \mathbf{D} \mathbf{\Lambda}$	$\tilde{\mathbf{X}} = \mathbf{Q}_{1,r} \Sigma_r \mathbf{Q}_{2,r}^\top$
Modes POD	$\mathbf{\Phi}$	$\mathbf{\Phi} = \tilde{\mathbf{X}} \mathbf{D} \mathbf{\Lambda}^{-1/2}$	$\mathbf{\Phi} = \mathbf{Q}_{1,r}$
Valeurs propres	$\mathbf{\Lambda}$	$\mathbf{\Lambda}$	$\mathbf{\Lambda} = \Sigma_r^2$
Pondération des clichés	$\mathbf{D} = \tilde{\mathbf{X}}^\top \mathbf{\Phi} \mathbf{\Lambda}^{-1/2}$	$\mathbf{D}$	$\mathbf{D} = \mathbf{Q}_{2,r}$

**Tableau B.1 :** Relations entre les modes POD  $\mathbf{\Phi}$ , les valeurs propres  $\mathbf{\Lambda}$  et les coefficients  $\mathbf{D}$  lorsque ces quantités sont calculées par la méthode directe, la méthode des clichés ou une décomposition en valeurs singulières. Les opérateurs de corrélation ainsi que les problèmes aux valeurs propres sont précisés dans les deux premières lignes, de même que l'expression de la SVD.

Dans le paragraphe 2.2.2, on a introduit la décomposition (2.20) qui a conduit à formuler le problème aux valeurs propres symétrisé (2.25). Les coefficients  $\mathbf{c}^{(j)}$  sont reliés aux vecteurs propres  $\mathbf{d}^{(j)}$  par la relation  $\mathbf{c}^{(j)} = \tilde{\boldsymbol{\alpha}} \mathbf{d}^{(j)}$  et les modes POD normalisés sont donc donnés par la relation  $\mathbf{\Phi} = \tilde{\mathbf{X}} \mathbf{D} \mathbf{\Lambda}^{-1/2}$  fournie dans le tableau B.1. De façon équivalente, on peut en déduire une expression des coefficients de la matrice  $\mathbf{D}$  en fonction des modes POD  $\mathbf{\Phi}$  pour compléter les quantités obtenues par la méthode directe. Pour cela, il suffit « d'inverser » la relation en multipliant à droite par  $\mathbf{\Lambda}^{1/2}$  et à gauche par  $\tilde{\mathbf{X}}^\top$ , ce qui donne :  $\tilde{\mathbf{X}}^\top \mathbf{\Phi} \mathbf{\Lambda}^{1/2} = (\tilde{\mathbf{X}}^\top \tilde{\mathbf{X}}) \mathbf{D}$ . On reconnaît à droite l'opérateur de corrélation  $\mathbf{R}^*$  si bien que le produit  $\mathbf{R}^* \mathbf{D}$  est remplacé par  $\mathbf{D} \mathbf{\Lambda}$ . Finalement, en remplaçant dans l'expression précédente et en multipliant à droite par  $\mathbf{\Lambda}^{-1}$ , on parvient à la relation  $\mathbf{D} = \tilde{\mathbf{X}}^\top \mathbf{\Phi} \mathbf{\Lambda}^{-1/2}$  qui fournit les coefficients  $\mathbf{D}$  en fonction des quantités  $\mathbf{\Phi}$  et  $\mathbf{\Lambda}$  produites par la résolution du problème aux valeurs propres de la méthode directe.

Pour terminer, il est utile de disposer d'une expression matricielle des amplitudes modales  $a_j^{(m)}$  regroupées dans la matrice  $\mathbf{A} \in \mathcal{M}(r, M)$  de sorte que la décomposition orthogonale propre des clichés donnée à l'équation (2.17) s'écrive en notation matricielle

$$\mathbf{X} = \mathbf{\Phi} \mathbf{A}. \quad (\text{B.28})$$

Les amplitudes modales sont alors simplement données par la relation  $\mathbf{A} = \mathbf{\Phi}^\top \mathbf{X}$  puisque les modes POD sont orthonormés. Par ailleurs il est parfois nécessaire de disposer d'une relation directe en fonction des coefficients  $\mathbf{D}$ . Pour cela il suffit de remplacer l'expression de  $\mathbf{\Phi}$  par celle donnée dans le tableau en fonction de  $\mathbf{D}$ , ce qui conduit à  $\mathbf{A} = (\tilde{\mathbf{X}} \mathbf{D} \mathbf{\Lambda}^{-1/2})^\top \mathbf{X} = \mathbf{\Lambda}^{-1/2} \mathbf{D}^\top (\tilde{\mathbf{X}}^\top \mathbf{X})$ . La relation précédente est alors multipliée à droite par  $\tilde{\boldsymbol{\alpha}} \mathbf{D}$ , de sorte que dans le membre de droite apparaisse le terme  $(\tilde{\mathbf{X}}^\top \tilde{\mathbf{X}}) \mathbf{D} = \mathbf{R}^* \mathbf{D} = \mathbf{D} \mathbf{\Lambda}$ . Par conséquent, la relation se réduit à  $\mathbf{A} \tilde{\boldsymbol{\alpha}} \mathbf{D} = \mathbf{\Lambda}^{-1/2} \mathbf{D}^\top \mathbf{D} \mathbf{\Lambda} = \mathbf{\Lambda}^{1/2}$  puisque les vecteurs propres  $\mathbf{D}$  sont orthonormaux. Finalement, les amplitudes se déduisent par la relation suivante :  $\mathbf{A} = \mathbf{\Lambda}^{1/2} \mathbf{D}^\top \tilde{\boldsymbol{\alpha}}^{-1}$ . En fonction du problème résolu, la matrice des amplitudes modales  $\mathbf{A}$  s'écrit donc

$$\mathbf{A} = \mathbf{\Phi}^\top \mathbf{X} = \mathbf{Q}_{1,r}^\top \mathbf{X} \quad \text{ou} \quad \mathbf{A} = \mathbf{\Lambda}^{1/2} \mathbf{D}^\top \tilde{\boldsymbol{\alpha}}^{-1} = \Sigma_r \mathbf{Q}_{2,r}^\top \tilde{\boldsymbol{\alpha}}^{-1}. \quad (\text{B.29})$$

En conclusion, on a montré que les opérateurs  $\mathbf{R}$  et  $\mathbf{R}^*$  permettaient tous deux de calculer les modes POD associés aux clichés représentés par la matrice  $\mathbf{X}$ . La formulation POD directe fournit directement les modes POD comme solution du problème aux valeurs propres alors que la méthode des clichés nécessite de résoudre un problème aux valeurs propres annexe de taille  $M \times M$  à l'issue duquel les modes POD doivent être reconstruits par combinaison linéaire des clichés suivant l'expression donnée dans le théorème B.2. Il est important de noter que la condition d'indépendance linéaire des clichés n'est intervenue à aucun moment alors qu'elle était nécessaire au paragraphe 2.2.2. Par conséquent, en dimension finie, les méthodes directe et des clichés sont équivalentes, même si le rang  $r$  de la matrice des clichés est strictement inférieur à  $M$ .



# Compléments d'analyse du modèle d'ordre réduit du système linéaire autonome

---

## Sommaire

C.1 Influence de l'échantillonnage des clichés d'un système périodique . . . . .	227
C.2 Convergence des modes POD vers les modes propres . . . . .	232
C.2.1 Un résultat préliminaire . . . . .	232
C.2.2 Convergence des vecteurs propres pour la méthode directe . . . . .	233
C.2.3 Convergence des vecteurs propres pour la méthode des clichés . . . . .	235

---

Le système linéaire et autonome considéré dans cette annexe correspond à l'exemple de la barre encastrée-libre traité au paragraphe 3.2 dont on a étudié la réponse à une déformation initiale appliquée à l'extrémité libre. La formulation analytique du modèle d'ordre réduit est tout d'abord employée pour mettre en évidence l'importance de l'échantillonnage dans le cas d'un système périodique, puis, dans un second temps, des éléments de démonstration de la convergence des modes POD vers les modes propres seront fournis pour ce type particulier de système dynamique.

## C.1 Influence de l'échantillonnage des clichés d'un système périodique

Parmi les paramètres importants entrant en ligne de compte dans la construction du modèle d'ordre réduit, la durée sur laquelle les clichés sont prélevés requiert une certaine attention. Les résultats présentés ici complètent ceux du paragraphe 3.3.2 où l'influence du nombre  $M$  de clichés et du nombre  $q$  de modes POD conservés dans la base de projection a été étudiée.

La solution analytique (3.4) du système en réponse à une déformée initiale avait permis de mettre en évidence l'expression (3.6) de la fréquence fondamentale  $T_0$ . Les clichés sont prélevés en  $M$  instants  $t_m$  qu'il est nécessaire de choisir judicieusement afin que l'opérateur de corrélation contienne un maximum d'informations pour calculer les modes POD. Un échantillonnage uniforme des clichés est pratiqué puisque la réponse du système est périodique et ne présente pas de phénomène localisé dans le temps qui justifierait l'emploi d'un échantillonnage variable. Ce processus de sélection est toutefois dangereux lorsque la réponse du système est périodique car un certain nombre de clichés peuvent devenir redondants. L'objectif est donc de déterminer sous quelles conditions l'échantillonnage uniforme fournit le maximum d'information.

La résolution du modèle non réduit est conduite sur un intervalle de temps  $I_e = [t_0; t_0 + T_e]$  au cours duquel les clichés sont prélevés. La durée  $T_e$  correspond à la *durée d'échantillonnage* qui s'exprime en fonction de la période fondamentale  $T_0$  par la relation

$$T_e = n_p T_0, \tag{C.1}$$

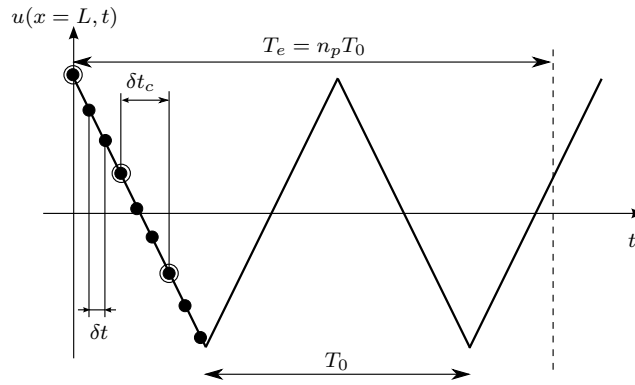
où  $n_p \in \mathbb{R}$  désigne le nombre de périodes sur lesquelles les clichés seront extraits — ce nombre étant en général non entier. Au cours de l'intervalle de temps  $I_e$ , des clichés du modèle non réduit sont prélevés en  $M$  instants pour calculer les modes POD. Lorsque l'échantillonnage est uniforme, ceci conduit à définir le pas de temps d'échantillonnage des clichés

$$\delta t_c = \frac{T_e}{M}. \quad (\text{C.2})$$

En remplaçant la durée  $T_e$  par son expression (C.1), on obtient la relation

$$\frac{M}{n_p} = \frac{T_0}{\delta t_c} \quad (\text{C.3})$$

qui lie le rapport du nombre de clichés sur le nombre de périodes au rapport de la période fondamentale sur le pas d'échantillonnage des clichés. Les paramètres introduits ci-dessus sont résumés pour plus de clarté sur la figure C.1.



**Figure C.1 :** Représentation sur la réponse schématique du système des paramètres temporels employés pour l'analyse de l'influence de l'échantillonnage des clichés. Les points  $\bullet$  correspondent à la discrétisation temporelle du modèle non réduit avec le pas de temps  $\delta t$ , alors que les cercles  $\circ$  représentent les clichés prélevés de manière uniforme avec le pas de temps  $\delta t_c$ .

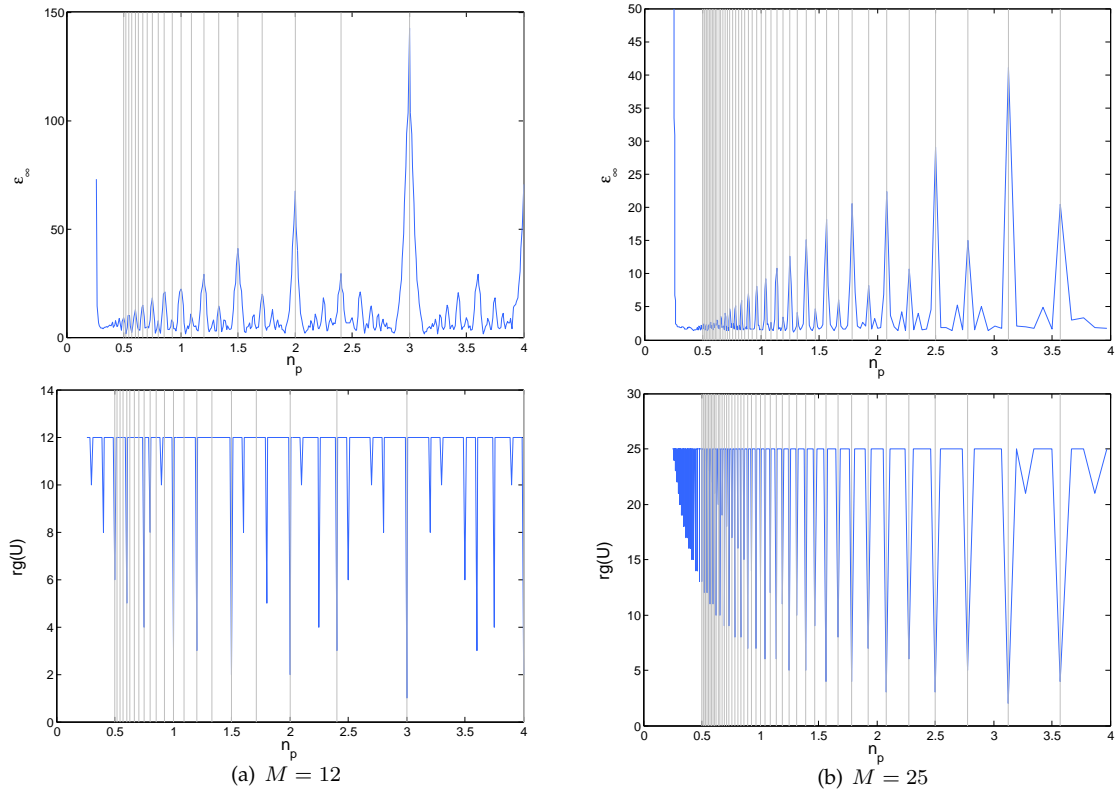
Pour certaines combinaisons de  $M$  et  $n_p$ , le rapport  $M/n_p$  est entier de sorte que le pas d'échantillonnage des clichés devient un sous-multiple de la fréquence fondamentale du système. Ainsi, pour tout  $k \in \mathbb{N}^*$

$$\frac{M}{n_p} = k \iff T_0 = k \delta t_c. \quad (\text{C.4})$$

Les combinaisons pour lesquelles le rapport  $M/n_p$  est entier sont problématiques puisque certains clichés sélectionnés deviennent redondants, ce qui entraîne une diminution de la quantité d'information contenue dans la base des clichés. Ce phénomène découle simplement du fait que les clichés ne sont plus linéairement indépendants et se traduit donc par la chute du rang de la matrice des clichés et par voie de conséquence de celui de l'opérateur de corrélation. Le nombre de valeurs propres non nulles diminue, de même que le nombre de modes POD disponibles.

Afin de vérifier à quelles conditions le rang de l'opérateur de corrélation peut chuter, on trace sur la figure C.2 l'évolution de l'erreur infinie  $\varepsilon_\infty$  commise dans la réponse du modèle d'ordre réduit ainsi que la valeur du rang  $r$  de l'opérateur de corrélation en fonction du nombre de périodes  $n_p \in [0, 25; 4, 0]$  sur lesquelles les clichés sont échantillonnés. Le résultat est fourni pour deux valeurs du nombre total de clichés  $M = 12$  et  $M = 25$ . Plusieurs remarques intéressantes peuvent être formulées à partir de l'analyse des graphiques :

- l'erreur de reconstruction explose lorsque  $n_p \leq 0, 25$  ;
- il existe des pics d'erreur localisés ;
- l'amplitude de ces pics augmente globalement avec  $n_p$ .



**Figure C.2 :** Évolution de l'erreur de reconstruction infinie  $\varepsilon_\infty$  et du rang  $r$  de l'opérateur de corrélation en fonction du nombre de périodes  $n_p$  sur lesquelles les clichés sont échantillonnés. Les résultats sont présentés pour deux valeurs du nombre  $M$  de clichés. Sur chaque graphique, les droites verticales représentent les valeurs critiques  $M/n_p = k$ , pour  $k \in \mathbb{N}^*$ .

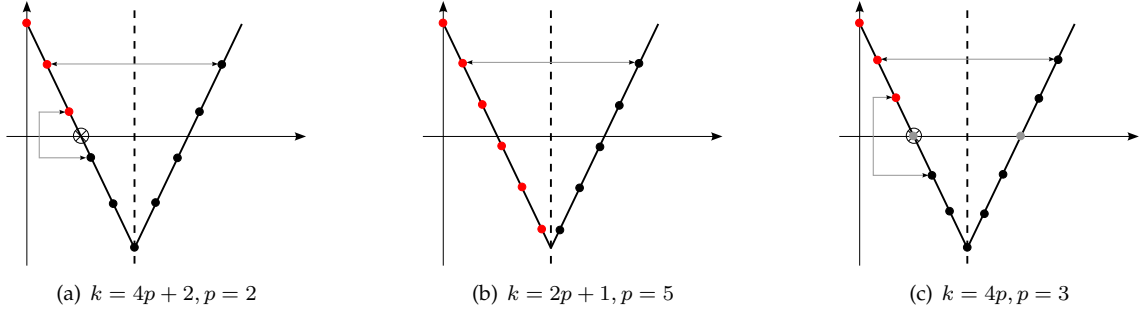
La première remarque signifie simplement que les informations contenues dans un ensemble de clichés échantillonnés sur une durée inférieure à un quart de période sont insuffisantes pour caractériser le système. En effet, seules les valeurs positives des déplacements sont rencontrées sur cet intervalle, et la solution reconstruite par le modèle d'ordre réduit correspond à la valeur absolue de la solution de référence.

Les pics d'erreur localisés apparaissent pour certaines combinaisons de  $M$  et  $n_p$ , lorsque certains clichés deviennent redondants en raison de la périodicité de la réponse. Ces valeurs critiques représentées par les droites grisées sur les graphiques de la figure C.2 correspondent aux valeurs entières de  $M/n_p$  pour lesquelles le pas de temps d'échantillonnage devient un sous-multiple de la période fondamentale. La redondance des clichés est confirmée par les graphiques de la seconde ligne de la figure C.2 qui indiquent que le rang évolue d'une façon aussi irrégulière que l'erreur infinie. La diminution du rang s'explique par le fait que suivant les combinaisons choisies, certains clichés ne sont plus linéairement indépendants.

La courbe pour  $M = 25$  fait en outre clairement ressortir le fait que l'amplitude de l'erreur  $\varepsilon_\infty$  entre deux pics successifs n'augmente pas de manière régulière : les pics associés à des valeurs critiques pour lesquelles  $k$  est pair sont plus intenses que lorsque  $k$  est impair. Par exemple, pour  $M = 25$ , les pics à  $n_p = 2, 5$  et  $n_p = 3, 125$  correspondent à  $k = 10$  et  $k = 8$  ; ils encadrent le pic d'amplitude inférieure à  $n_p = 2, 78$  qui correspond à  $k = 9$ . On reviendra plus tard sur l'interprétation de cette observation.

La chute du rang de l'opérateur de corrélation est dans la plupart des cas associée au fait que  $M/n_p$  prend une valeur entière  $k \in \mathbb{N}^*$ . Ainsi, prendre  $M$  clichés sur  $n_p$  périodes revient en réalité à en prendre  $k$  sur une période complète ( $n_p = 1, 0$ ). On peut donc raisonner sur une période d'oscillation du système et observer ce qui se passe suivant les valeurs de  $k$ . Le fait que  $k$  soit entier implique que le pas de temps d'échantillonnage des clichés est un sous-multiple de la période d'oscillation du système d'après la relation (C.4). Or la réponse du système est loin d'être quelconque puisqu'elle possède des propriétés importantes de périodicité et de parité. Il est clair que la réponse (3.4) est périodique de période  $T_0$  et est en outre

symétrique par rapport aux axes  $t_n = nT_0/2$  pour  $n \in \mathbb{N}^*$  et aux points  $t_n = nT_0/4$  avec  $n$  impair. Sur une période d'oscillation, l'axe central  $t = T_0/2$  constitue donc un axe de symétrie, tandis que les points  $T_0/4$  et  $3T_0/4$  sont des points de symétrie. Suivant la parité de  $k$ , le pas de temps d'échantillonnage des clichés  $\delta t_c$  prend des valeurs telles que certains clichés se répètent. Les schémas de la figure C.3 traduisent ce phénomène pour trois valeurs de  $k$  : une impaire, une paire et multiple de 2 et une paire multiple de 4.



**Figure C.3 :** Schématisation de la redondance des clichés sur une période d'oscillation du système pour trois valeurs de  $k$  telles que  $M/n_p = k$ . Les droites verticales en pointillés (—) indiquent l'axe de symétrie et le symbole  $\otimes$  désigne le point de symétrie centrale. Les points en rouge (•) correspondent aux clichés linéairement indépendants subsistants après élimination des redondances par symétrie.

Lorsque  $k$  est impair, seul l'axe  $t = T_0/2$  joue un rôle : la moitié des clichés est identique par symétrie axiale, si bien que le rang de la matrice des clichés chute de  $(k + 1)/2$ . La situation se dégrade lorsque  $k$  est pair et multiple de 2 seulement. Dans ce cas, en plus de la symétrie axiale qui élimine la moitié des clichés, vient s'ajouter la symétrie centrale en  $t = T_0/4$  qui agit sur les  $(k + 2)/2$  clichés restant. La symétrie centrale est liée au caractère impair de la solution et exprime la redondance de certains clichés qui sont simplement opposés les uns des autres. Au final, en supprimant ces clichés, il ne reste plus que  $(k + 2)/4$  clichés linéairement indépendants. Enfin le cas le plus alarmant correspond aux valeurs de  $k$  multiples de 4 : pour ces valeurs, on retrouve évidemment les symétries axiales et centrales qui éliminent une grande partie des clichés, mais il se trouve que dans ce dernier cas, un des clichés restant est nul et doit donc être supprimé aussi. On peut alors établir les relations suivantes exprimant le rang de la matrice des clichés  $\mathbf{U}$  en fonction du nombre  $k$  correspondant au rapport  $M/n_p$ . Pour tout  $p \in \mathbb{N}$ ,

$$\begin{aligned}
 \text{si } k = 2p + 1 &\implies \text{rg}(\mathbf{U}) = \frac{k + 1}{2}, \\
 \text{si } k = 4p + 2 &\implies \text{rg}(\mathbf{U}) = \frac{k + 2}{4}, \\
 \text{si } k = 4p &\implies \text{rg}(\mathbf{U}) = \frac{k}{4}.
 \end{aligned} \tag{C.5}$$

Les relations précédentes sont valables lorsque  $n_p \geq 0,5$ . En effet, pour  $0,25 \leq n_p < 0,5$ , seule la symétrie centrale en  $T_0/4$  est encore susceptible d'agir et d'éliminer ainsi quelques clichés, ce qui peut encore se produire si  $k$  est pair. Enfin, lorsque  $n_p < 0,25$ , aucune symétrie ne provoque de redondance de sorte que le rang de la matrice des clichés reste maximal et égal au nombre de clichés. Cependant, pour  $n_p < 0,25$ , les informations contenues dans les clichés ne sont pas assez représentatives de la réponse du système.

L'étude de l'évolution de l'erreur de reconstruction en combinaison avec celle du rang de la matrice des clichés a permis de déterminer que la borne inférieure pour  $n_p$  se situe aux alentours de 0,25. Dans l'intervalle  $[0,25; 0,5]$ , les valeurs critiques  $M/n_p \in \mathbb{N}^*$  sont très proches et il devient donc délicat de trouver a priori des valeurs de  $n_p$  qui conviennent. Le choix  $n_p \approx 1$  avec  $n_p \neq 1$  semble donc être un bon compromis puisque les risques de tomber sur une valeur critique dans un voisinage de 1 s'amenuisent et l'erreur de reconstruction est encore acceptable. On prendra souvent dans les applications  $n_p \approx 1,3$ .

Pour conclure cette analyse, rappelons qu'à l'origine la décomposition orthogonale propre est une méthode d'analyse de données et que les modes POD mis en évidence ne sont représentatifs de la base de

données que si les clichés contiennent le maximum d'informations sur la réponse. L'échantillonnage est donc une partie cruciale du processus qui détermine en grande partie la qualité du modèle d'ordre réduit et le choix de prélever uniformément les clichés n'est pas forcément judicieux dans le cas d'une réponse périodique. Une extraction aléatoire des clichés sur l'intervalle  $I_e$  serait potentiellement plus adéquate, et renouerait sur le principe avec les concepts probabilistes de la décomposition de Karhunen-Loève.

**Remarque C.1 : Diminution du rang pour  $M/n_p = k, k \notin \mathbb{N}$**

Si on observe en détail les graphiques de la figure C.2, on constate qu'il existe d'autres valeurs de  $n_p$  pour lesquelles le rang de la matrice des clichés diminue, sans que cela ne se traduise nécessairement par un accroissement important de l'erreur. En réalité, la diminution du rang à cause des symétries axiales et centrales ne nécessite pas que  $k$  soit entier. De manière générale, si  $M/n_p \in \mathbb{Q}$ , le rapport se réduit à  $M'/n'_p$  avec  $(M', n'_p) \in \mathbb{N}^* \times \mathbb{N}^*$  de sorte que

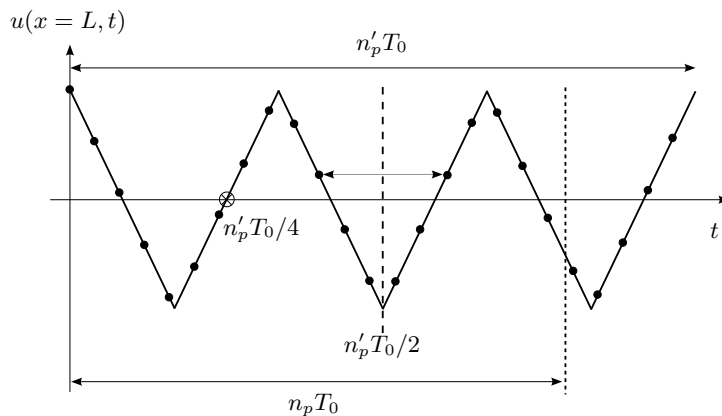
$$\frac{M}{n_p} = \frac{M'}{n'_p}.$$

Ainsi, prendre  $M$  clichés sur  $n_p$  périodes revient à en prendre  $M'$  sur un nombre entier  $n'_p$  de périodes. Les valeurs critiques évoquées précédemment correspondent donc au cas particulier où la fraction rationnelle se réduit avec  $n'_p = 1$ . Toutefois dans le cas général  $n'_p$  est différent de 1 mais reste toujours entier donc le risque de rencontrer des symétries sur l'intervalle  $[0; n'_p T_0]$  persiste. La symétrie centrale se situe à présent en  $t = n'_p T_0 / 2$  et le point de symétrie centrale se trouve quant à lui en  $t = n'_p T_0 / 4$ . Pour être sûr que les clichés sont indépendants, il est donc nécessaire que l'axe et le point de symétrie de l'intervalle de temps équivalent  $[0; n'_p T_0]$  soient situés en dehors de l'intervalle de temps  $[0; n_p T_0]$ . Le principe est résumé sur la figure C.4. Intuitivement, cela signifie que  $n'_p \gg n_p$ , ou encore que la fraction rationnelle  $M/n_p$  ne peut pas être « beaucoup » réduite.

Plus précisément, cela se traduit par  $n'_p T_0 / 2 > n_p T_0$  pour l'axe de symétrie et  $n'_p T_0 / 4 > n_p T_0$  pour le point de symétrie centrale. L'indépendance des clichés est alors assurée sous les deux conditions

$$n'_p > 2 n_p \quad \text{et} \quad n'_p > 4 n_p.$$

Lorsque ces conditions sont remplies, le rang de la matrice des clichés est maximal et on peut espérer que l'erreur de reconstruction sera faible. Au contraire, le non-respect de ces conditions conduit à la répétition de quelques clichés, dont le nombre est en général moins important que lorsque  $k$  est entier. En conclusion il semble donc que le choix idéal consiste à prendre le nombre de période  $n_p$  tel que le rapport  $M/n_p$  soit irrationnel. En effet lorsque  $M/n_p \in \mathbb{R} \setminus \mathbb{Q}$  il n'est pas possible d'interpréter l'échantillonnage au moyen d'un nombre de période équivalent  $n'_p$  entier et par conséquent les risques de redondance des clichés par symétrie sont évités.



**Figure C.4 :** Schématisation des symétries s'exerçant sur l'intervalle de temps équivalent. Prendre  $M = 20$  clichés sur une durée  $n_p = 2, 4$  revient en fait à considérer  $M' = 25$  clichés sur  $n'_p = 3$  périodes complètes. Il existe donc un axe de symétrie en  $n'_p T_0 / 2$  qui élimine dans le cas présent les 7 clichés dans l'intervalle  $[n'_p T_0 / 2; n_p T_0]$ ; le point de symétrie centrale en  $n'_p T_0 / 4$  ne joue ici aucun rôle.



## C.2 Convergence des modes POD vers les modes propres dans le cas d'un système linéaire autonome

Feeny et Kappagantu [74] et Kerschen et Golinval [129] ont montré que sous certaines conditions, les modes POD convergent vers les modes propres lorsque le nombre de clichés  $M$  utilisés pour construire l'opérateur de corrélation augmente. Ce résultat a en particulier été démontré lorsque le système considéré est (i) linéaire, (ii) non amorti, (iii) autonome et (iv) possède une matrice de masse proportionnelle à l'identité. Cette convergence se généralise par ailleurs à des systèmes pour lesquels la matrice de masse n'est pas proportionnelle à l'identité par un simple changement de variables. En outre, la convergence persiste pour des systèmes faiblement amortis. Dans le cas d'un système non-autonome où l'excitation est harmonique, Kerschen et Golinval [129] ont montré que le mode POD dominant est lié à la réponse harmonique forcée du système, et si la matrice de masse est proportionnelle à l'identité, alors les modes POD restants sont les modes propres du système linéaire. Enfin dans le cas d'un système non-linéaire, les modes POD représentent la meilleure approximation linéaire des modes propres non-linéaires. Cependant, les conditions établies dans [74, 129] sont incomplètes puisqu'il n'y a convergence en fonction du nombre de clichés que si ces derniers sont linéairement indépendants.

Pour établir la convergence des modes POD vers les modes propres, on considère la formulation analytique du modèle d'ordre réduit développée dans le paragraphe 3.2.1, à l'aide de laquelle les modes POD ont été décomposés sur la base des modes propres par la relation (3.12). Il suffit alors de montrer que les coefficients de la décomposition (3.12) s'annulent tous sauf un à la limite de la convergence :  $\mu_l^{(j)} \rightarrow \delta_{lj}$  lorsque  $M \rightarrow \infty$ . Pour cela, on considère ici des clichés non centrés ( $\bar{\zeta} = 0$ ) et on établit dans un premier temps une conjecture sur la convergence des séries de la forme  $\sum_{m=1}^M \alpha_m \cos(\kappa_n t_m)$  qui dépend en particulier du fait que les clichés sont indépendants ou non. À l'aide de cette relation, la forme asymptotique de l'opérateur aux valeurs propres ainsi que celle des vecteurs propres sera déterminée aussi bien pour la méthode directe que pour la méthode des clichés.

### C.2.1 Un résultat préliminaire

Pour commencer, on cherche à caractériser la convergence des séries de la forme

$$S_n(M) = \sum_{m=1}^M \alpha_m \cos(\kappa_n t_m), \quad (\text{C.6})$$

où  $M$  désigne le nombre de clichés, les  $\alpha_m$  sont les coefficients de pondération associés à l'opérateur de moyenne  $E[\cdot]$ , les  $t_m$  sont les instants où les clichés sont prélevés et la fréquence  $\kappa$  est un multiple de  $\pi$  de la forme  $\kappa_n = (n+1)\pi c/L$  avec  $n \in \mathbb{N}^*$ . On considère ici que les clichés sont prélevés sur l'intervalle de temps  $I_e$  si bien que faire tendre  $M$  vers l'infini revient à choisir des pas de temps d'échantillonnage  $\delta t_c = T_e/M$  de plus en plus petits. On conjecture alors que les séries de type  $S_n(M)$  convergent vers zéro lorsque le nombre de clichés  $M$  augmente et que le nombre de période  $n_p$  est correctement choisi :

$$S_n(M) \xrightarrow{M \rightarrow \infty} 0. \quad (\text{C.7})$$

Le résultat précédent est difficile à démontrer rigoureusement dans le cas général car il dépend fortement de la façon dont sont répartis les clichés, et en particulier du nombre de périodes  $n_p$  sur lequel ils sont prélevés. L'annulation des séries — observée en pratique pour la plupart des valeurs de  $n_p$  — peut être justifiée en partie par l'existence d'axes et de points de symétrie pour la fonction cosinus qui provoquent la répétition des valeurs  $\cos(\kappa_n t_m)$  avec des signes opposés (voir à titre indicatif les schémas de la figure C.3). Lorsque les poids de pondération  $\alpha_m$  de l'opérateur de moyenne  $E[\cdot]$  sont tous identiques, la plupart des valeurs des cosinus s'annulent et il subsiste quelques termes résiduels. Ces termes se compensent seulement si le rapport  $M/n_p$  est bien choisi, *i. e.* si les termes restant sont indépendants et de signe opposé. Lorsque  $M$  augmente, la somme des termes résiduels devient alors quasiment nulle et les séries convergent. On ne détaille pas plus les raisons pour lesquelles les séries convergent puisque l'objectif n'est pas de montrer rigoureusement la relation C.7 mais simplement de se servir de ce résultat observé numériquement pour justifier pourquoi les modes POD convergent vers les modes propres du système.

## C.2.2 Convergence des vecteurs propres pour la méthode directe

Les clichés sont décomposés sur la base des modes propres selon la relation (3.7) et les modes POD s'expriment de la même manière au moyen de la relation (3.12). Seuls  $p$  modes propres sont conservés pour l'approximation de la solution analytique. L'opérateur de corrélation prend donc la forme (B.3) avec  $\eta_i^{(m)} = \zeta_i^{(m)}$  et le problème aux valeurs propres de la méthode directe s'écrit  $\mathbf{R}\boldsymbol{\mu}^{(j)} = \lambda_j \boldsymbol{\mu}^{(j)}$  avec la formulation analytique. Pour montrer la convergence des modes POD vers les modes propres, il suffit de montrer que  $\mu_l^{(j)} \rightarrow \delta_{lj}$  quand  $M \rightarrow \infty$ , ce qui signifie que les vecteurs propres  $\boldsymbol{\mu}^{(j)}$  de la formulation analytique convergent vers les vecteurs de la base canonique de  $\mathbb{R}^p$ . Ces vecteurs notés ici  $\mathbf{e}_j$  contiennent une seule composante non nulle sur le  $j$ -ième terme.

On détaille donc pour commencer l'expression de  $\mathbf{R}\mathbf{e}_j$  dans le but de montrer que lorsque  $M$  tend vers l'infini, les vecteurs de la base canonique sont des vecteurs propres de  $\mathbf{R}$ . Le produit  $\mathbf{R}\mathbf{e}_j$  donne trivialement la  $j$ -ième colonne de  $\mathbf{R}$  qui s'écrit  $[E[\zeta_1^{(m)}\zeta_j^{(m)}], \dots, E[\zeta_p^{(m)}\zeta_j^{(m)}]]^T$  d'après l'expression (B.3) de l'opérateur aux valeurs propres de la formulation analytique. Il faut donc montrer que lorsque  $M$  tend vers l'infini, seul le  $j$ -ième terme du vecteur  $\mathbf{R}\mathbf{e}_j$  subsiste. Le terme général se décompose en

$$\begin{aligned} E[\zeta_i^{(m)}\zeta_j^{(m)}] &= \sum_{m=1}^M \alpha_m \zeta_i \zeta_j \cos(k_i t_m) \cos(k_j t_m) \\ &= \frac{\zeta_i \zeta_j}{2} \sum_{m=1}^M \alpha_m [\cos((k_i + k_j)t_m) + \cos((k_i - k_j)t_m)]. \end{aligned} \quad (\text{C.8})$$

Les sommes apparaissant dans l'expression précédente sont pour la plupart du même type que les séries  $S_n(M)$  pour lesquelles on a expliqué qu'elles convergeaient vers 0 lorsque  $M$  était grand pour tout  $n \geq 0$ . En effet, lorsque  $k_i \neq k_j$ , la somme et la différence des nombres d'ondes s'écrivent respectivement  $(i + j - 1)\pi c/L$  et  $(i - j - 1)\pi c/L$  et il est donc possible d'identifier l'entier  $n$  correspondant à la fréquence  $\kappa_n$ . Ainsi les termes de la somme s'annulent tous si le nombre de clichés est assez grand et que  $i \neq j$ . Dans le cas contraire, la somme s'écrit

$$E[\zeta_i^{(m)}\zeta_i^{(m)}] = \sum_{m=1}^M \alpha_m \zeta_i^2 \cos^2(k_i t_m) = \frac{\zeta_i^2}{2} \sum_{m=1}^M \alpha_m [1 + \cos((2k_i)t_m)]. \quad (\text{C.9})$$

La somme des cosinus s'annule ici encore puisqu'il s'agit des séries  $S_{2i-2}(M)$ . Par contre, le premier terme subsiste et est égal à 1 pour tout  $M$  puisqu'il s'agit de la somme des poids de pondération associés à la moyenne. Par conséquent, en notant  $\lambda_j = \zeta_j^2/2$  on peut écrire

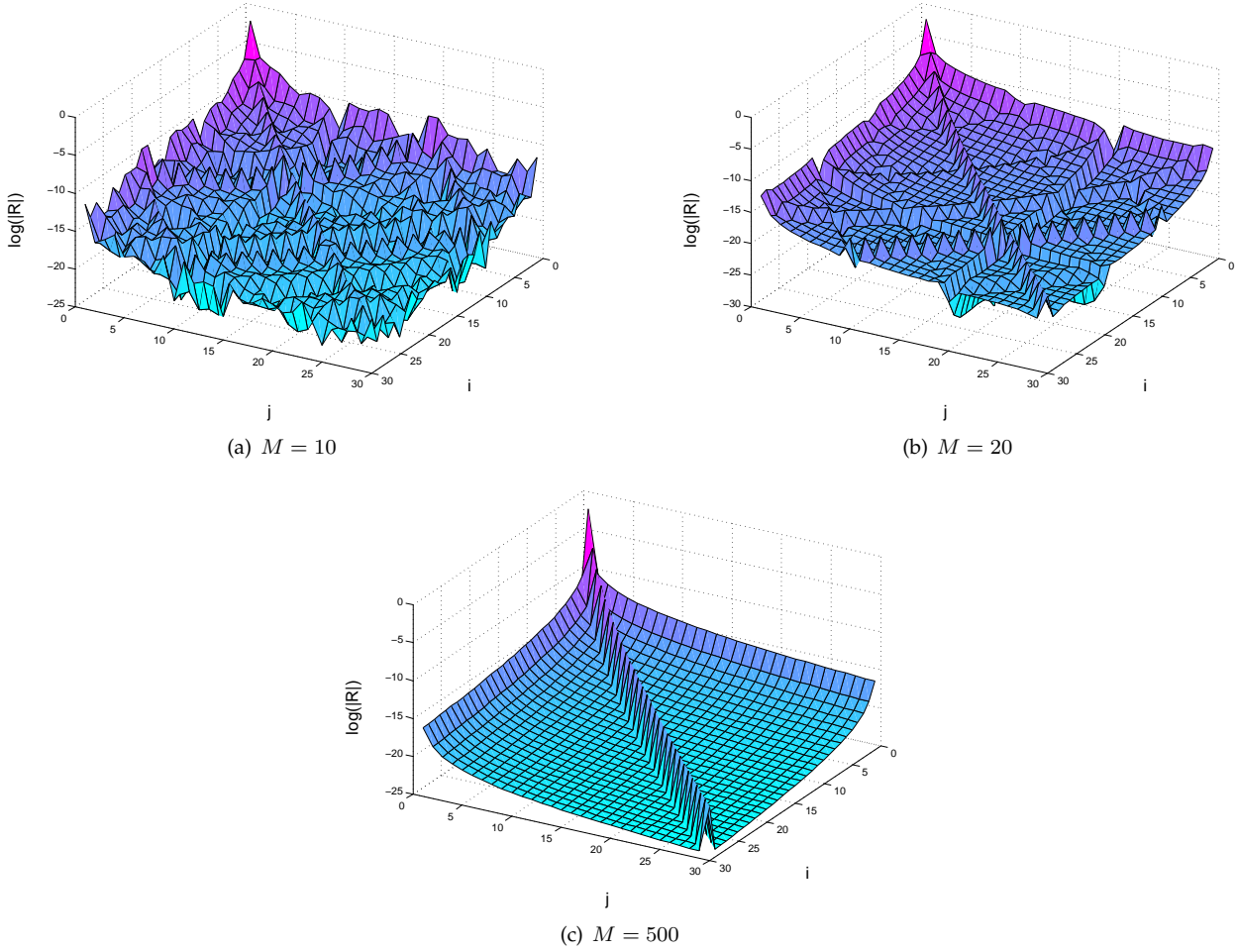
$$\lim_{M \rightarrow \infty} E[\zeta_i^{(m)}\zeta_j^{(m)}] = \lambda_j \delta_{ij}. \quad (\text{C.10})$$

Ainsi, lorsque  $M$  tend vers l'infini, les vecteurs propres de l'opérateur  $\mathbf{R}$  tendent bien vers les vecteurs  $\mathbf{e}_j$  de la base canonique de  $\mathbb{R}^p$  et on a  $\lim_{M \rightarrow \infty} \mathbf{R}\mathbf{e}_j = \lambda_j \mathbf{e}_j$ . Par ailleurs, il est encore plus intéressant de remarquer que la relation (C.10) exprime le fait que lorsque le nombre de clichés tend vers l'infini, l'opérateur  $\mathbf{R}$  tend vers la matrice des valeurs propres. On obtient donc

$$\lim_{M \rightarrow \infty} \mathbf{R} = \text{diag}(\lambda_j) \quad (\text{C.11})$$

puisque les  $E[\zeta_i^{(m)}\zeta_j^{(m)}]$  correspondent au terme général de l'opérateur  $\mathbf{R}$ . Cela corrobore le fait que les clichés choisis sont bien décorrélés et cela est d'autant plus prononcé que le nombre de clichés augmente.

Les graphiques de la figure C.5 illustrent bien cette convergence lorsque les clichés ont été prélevés uniformément sur  $n_p \approx 1$  période d'oscillation. L'augmentation du nombre de clichés provoque alors un « lissage » de l'opérateur aux valeurs propres qui converge bien vers une matrice pratiquement diagonale dans laquelle les valeurs des pics de la diagonale principale correspondent aux valeurs propres de l'opérateur. Enfin, le tableau C.1 regroupe les 10 premières valeurs propres obtenues pour plusieurs valeurs du nombre  $M$  de clichés et met en évidence la convergence des valeurs propres  $\lambda_j$  du problème aux valeurs propres vers les coefficients  $\zeta_j^2/2$  qui avaient été identifiés aux valeurs propres pour établir la relation (C.10).



**Figure C.5 :** Convergence de l'opérateur aux valeurs propres de la méthode directe pour la formulation POD analytique en fonction du nombre de clichés pour  $n_p \approx 1, 0$ .

$j$	$M = 5$	$M = 10$	$M = 20$	$M = 50$	$M = 100$	$M = 500$	$\zeta_j^2/2$
1	2.170e-01	2.082e-01	2.036e-01	2.007e-01	1.997e-01	1.989e-01	1.987e-01
2	2.327e-03	2.557e-03	2.509e-03	2.477e-03	2.466e-03	2.456e-03	2.454e-03
3	3.756e-04	3.104e-04	3.246e-04	3.209e-04	3.195e-04	3.183e-04	3.180e-04
4	7.573e-05	8.440e-05	8.442e-05	8.352e-05	8.317e-05	8.286e-05	8.278e-05
5	-	3.012e-05	3.097e-05	3.056e-05	3.043e-05	3.032e-05	3.029e-05
6	-	1.449e-05	1.372e-05	1.369e-05	1.364e-05	1.359e-05	1.357e-05
7	-	7.874e-06	7.040e-06	7.016e-06	6.990e-06	6.966e-06	6.959e-06
8	-	4.177e-06	3.968e-06	3.957e-06	3.943e-06	3.930e-06	3.926e-06
9	-	3.249e-06	2.400e-06	2.398e-06	2.390e-06	2.382e-06	2.380e-06
10	-	-	1.534e-06	1.537e-06	1.532e-06	1.527e-06	1.525e-06

**Tableau C.1 :** Convergence des 10 premières valeurs propres de l'opérateur de corrélation en fonction du nombre de clichés. Les valeurs propres  $\lambda_j$  convergent vers les coefficients  $\zeta_j^2/2$  de la décomposition (3.4) des modes POD sur les modes propres.

### C.2.3 Convergence des vecteurs propres pour la méthode des clichés

Il s'agit ici encore de démontrer que  $\mu_l^{(j)} \rightarrow \delta_{lj}$  quand  $M \rightarrow \infty$  mais en utilisant cette fois le problème aux valeurs propres de la méthode des clichés  $\mathbf{R}^* \mathbf{d}^{(j)} = \lambda_j \mathbf{d}^{(j)}$  avec  $\mathbf{c}^{(j)} = \tilde{\alpha} \mathbf{d}^{(j)}$ . En observant à la suite de tests numériques la forme des vecteurs propres du problème non symétrique  $\tilde{\mathbf{R}}^* \mathbf{c}^{(j)} = \lambda_j \mathbf{c}^{(j)}$ , on peut conjecturer que la valeur asymptotique de leur composante est

$$c_i^{(j)} = \alpha_i \zeta_j^{(i)}. \quad (\text{C.12})$$

En supposant l'expression précédente vraie, les composantes du vecteur  $\mathbf{d}^{(j)}$  sont données par  $d_i^{(j)} = \sqrt{\alpha_i} \zeta_j^{(i)}$ . Par conséquent, l'expression du produit  $\mathbf{R}^* \mathbf{d}^{(j)}$  se développe avec la relation (3.9) en

$$\mathbf{R}^* \mathbf{d}^{(j)} = \sum_{m=1}^M \begin{bmatrix} \sqrt{\alpha_1} \alpha_m \left( \sum_{n=1}^p \zeta_n^{(1)} \zeta_n^{(m)} \right) \zeta_j^{(m)} \\ \vdots \\ \sqrt{\alpha_M} \alpha_m \left( \sum_{n=1}^p \zeta_n^{(M)} \zeta_n^{(m)} \right) \zeta_j^{(m)} \end{bmatrix}, \quad (\text{C.13})$$

et en réorganisant les termes, la  $i$ -ème composante du vecteur  $\mathbf{R}^* \mathbf{d}^{(j)}$  s'écrit

$$\left[ \mathbf{R}^* \mathbf{d}^{(j)} \right]_i = \sqrt{\alpha_i} \sum_{n=1}^p \zeta_n^{(i)} \left( \sum_{m=1}^M \alpha_m \zeta_n^{(m)} \zeta_j^{(m)} \right) = \sqrt{\alpha_i} \sum_{n=1}^p \zeta_n^{(i)} E \left[ \zeta_n^{(m)} \zeta_j^{(m)} \right]. \quad (\text{C.14})$$

Or d'après la relation (C.10) on peut réduire la somme à un seul terme et justifier ainsi que l'expression (C.12) conjecturée pour les  $\mathbf{c}^{(j)}$  correspond bien à des vecteurs propres pour  $\mathbf{R}^*$  puisque

$$\left[ \mathbf{R}^* \mathbf{d}^{(j)} \right]_i \xrightarrow{M \rightarrow \infty} \sqrt{\alpha_i} \sum_{n=1}^p \zeta_n^{(i)} \lambda_j \delta_{nj} = \lambda_j \sqrt{\alpha_i} \zeta_j^{(i)} = d_i^{(j)}. \quad (\text{C.15})$$

En utilisant la relation (3.12) donnant les coefficients  $\mu_i^{(j)}$  en fonction des  $c_l^{(j)}$  et en introduisant l'expression asymptotique (C.12) des coefficients  $c_l^{(j)}$  on a alors

$$\mu_i^{(j)} = \sum_{l=1}^M c_l^{(j)} \zeta_n^{(l)} = \sum_{l=1}^M \alpha_l \zeta_j^{(l)} \zeta_n^{(l)} = E \left[ \zeta_j^{(l)} \zeta_n^{(l)} \right] \xrightarrow{M \rightarrow \infty} \lambda_j \delta_{nj}, \quad (\text{C.16})$$

et on retrouve bien avec la méthode des clichés le fait que les modes POD convergent vers les modes propres de la structure.

Ce paragraphe a permis de mettre en évidence le fait que les modes POD convergent bien vers les modes propres de la barre encastée-libre lorsque le nombre de clichés tend vers l'infini. Le résultat est vrai lorsque l'opérateur de corrélation est construit avec des clichés non centrés par la méthode directe ou des clichés. En réalité, la convergence n'est pas simple à démontrer rigoureusement et notre propos s'est appuyé ici sur des résultats numériques afin de confirmer la conjecture établie dans le premier paragraphe. En particulier, la convergence a bien lieu sous les conditions explicitées dans [74, 129] et rappelées au début de ce paragraphe, mais l'indépendance des clichés joue un rôle crucial.



## Formulaire de calcul tensoriel

### Sommaire

<b>D.1 Tenseurs et produits</b> . . . . .	<b>237</b>
D.1.1 Notations tensorielle et indicielle . . . . .	237
D.1.2 Tenseurs particuliers . . . . .	237
D.1.3 Produits de tenseurs . . . . .	238
<b>D.2 Opérateurs différentiels</b> . . . . .	<b>239</b>
<b>D.3 Relations usuelles entre les opérateurs</b> . . . . .	<b>241</b>
<b>D.4 Théorèmes de Green-Ostrogradski et d'intégration par parties</b> . . . . .	<b>242</b>

## D.1 Tenseurs et produits

### D.1.1 Notations tensorielle et indicielle

Dans le formulaire, on considère des opérateurs tensoriels définis dans l'espace  $\mathbb{R}^3$  et représentés dans une base cartésienne  $\mathcal{B} = \{\mathbf{e}_i\}_{i=1,2,3}$  orthonormée directe. Les scalaires correspondent à des tenseurs d'ordre 0, les vecteurs à des tenseurs d'ordre 1, et les matrices à des tenseurs d'ordre 2.

Les scalaires seront notés par des lettres minuscules. Les vecteurs et les tenseurs d'ordre supérieurs seront quant à eux indiqués en lettres grasses, minuscules pour les vecteurs, et majuscules pour les tenseurs d'ordre supérieurs. On notera  $\mathbf{u} = \text{Mat}_{\mathcal{B}}(\mathbf{u})$  la représentation dans la base  $\mathcal{B}$  du vecteur  $\mathbf{u}$  de sorte que  $\mathbf{u} = [u_1, u_2, u_3]^T$ . De la même manière,  $\mathbf{D}$  est la représentation dans la base  $\mathcal{B}$  de la matrice  $D$  dont le terme général est noté  $d_{ij}$ , pour  $i, j = 1, \dots, 3$ .

On utilisera, sauf mention contraire, la convention d'Einstein sur les indices répétés : la répétition d'un indice dans une même expression indique une sommation implicite sur cet indice.

### D.1.2 Tenseurs particuliers

*Symbole de Kronecker :* La forme tensorielle du symbole de Kronecker  $\delta_i$  sert de base pour définir ensuite des tenseurs d'ordre quelconque. Le vecteur  $\delta_i$  est défini comme un vecteur dont toutes les composantes sont nulles, sauf celle située en  $i$ -ième position qui vaut 1. Le produit scalaire de deux vecteurs  $\delta_i$  et  $\delta_j$  donne naissance au symbole de Kronecker classique  $\delta_{ij}$  tel que  $\delta_{ij} = 1$  si  $i = j$  et  $\delta_{ij} = 0$  si  $i \neq j$ . Le symbole de Kronecker classique  $\delta_{ij}$  correspond alors au terme général de la matrice identité notée

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{avec} \quad I_{ij} = \delta_{ij}. \tag{D.1}$$

Tenseur d'orientation : Le tenseur d'orientation  $\varepsilon$  est un tenseur d'ordre 3 défini par

$$\varepsilon_{ijk} = \begin{cases} 1 & \text{si } (i, j, k) \text{ est une permutation circulaire de } (1, 2, 3), \\ -1 & \text{si } (i, j, k) \text{ est une permutation non circulaire de } (1, 2, 3), \\ 0 & \text{si } (i, j, k) \text{ n'est pas une permutation de } (1, 2, 3). \end{cases} \quad (\text{D.2})$$

### D.1.3 Produits de tenseurs

Produit matrice - vecteur : Le produit d'une matrice de  $\mathcal{M}(3, 3)$  par un vecteur de  $\mathbb{R}^3$  donne un vecteur de  $\mathbb{R}^3$  tel que

$$\mathbf{v} = \mathbf{D} \mathbf{u} = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \sum_{k=1}^3 \begin{bmatrix} d_{1k} u_k \\ d_{2k} u_k \\ d_{3k} u_k \end{bmatrix} \quad \text{avec } v_i = d_{ik} u_k. \quad (\text{D.3})$$

Produit vecteur - matrice : Le produit d'un vecteur par une matrice n'est défini que lorsque la transposée du vecteur est considérée, de sorte que les dimensions soient compatibles. Ainsi le produit de la transposée d'un vecteur de  $\mathbb{R}^3$  par une matrice de  $\mathcal{M}(3, 3)$  donne le vecteur de  $\mathbb{R}^3$

$$\mathbf{v} = \mathbf{u}^T \mathbf{D} = [u_1 \quad u_2 \quad u_3] \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} = \sum_{k=1}^3 [d_{k1} u_k \quad d_{k2} u_k \quad d_{k3} u_k] \quad \text{avec } v_i = d_{ki} u_k. \quad (\text{D.4})$$

Produit matrice - matrice : Le produit de deux matrices de  $\mathcal{M}(3, 3)$  donne la matrice de  $\mathcal{M}(3, 3)$

$$\mathbf{F} = \mathbf{D} \mathbf{E} = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} = \sum_{k=1}^3 \begin{bmatrix} d_{1k} e_{k1} & d_{1k} e_{k2} & d_{1k} e_{k3} \\ d_{2k} e_{k1} & d_{2k} e_{k2} & d_{2k} e_{k3} \\ d_{3k} e_{k1} & d_{3k} e_{k2} & d_{3k} e_{k3} \end{bmatrix} \quad \text{avec } F_{ij} = d_{ik} e_{kj}. \quad (\text{D.5})$$

Produit scalaire : Le produit scalaire de deux vecteurs représenté par le symbole  $\cdot$  donne un scalaire tel que

$$p = \mathbf{u} \cdot \mathbf{v} = \sum_{k=1}^3 u_k v_k \quad \text{soit } p = u_k v_k. \quad (\text{D.6})$$

Produit tensoriel : Le produit tensoriel de deux vecteurs représenté par le symbole  $\otimes$  donne la matrice

$$\mathbf{D} = \mathbf{u} \otimes \mathbf{v} = \begin{bmatrix} u_1 v_1 & u_1 v_2 & u_1 v_3 \\ u_2 v_1 & u_2 v_2 & u_2 v_3 \\ u_3 v_1 & u_3 v_2 & u_3 v_3 \end{bmatrix} \quad \text{avec } D_{ij} = u_i v_j. \quad (\text{D.7})$$

Le produit tensoriel de deux vecteurs correspond donc au produit matriciel de  $\mathbf{u}$  avec  $\mathbf{v}^T$ .

Double produit de contraction : Le double produit de contraction de deux matrices représenté par le symbole  $:$  donne un scalaire tel que

$$p = \mathbf{D} : \mathbf{E} = \sum_{i=1}^3 \sum_{j=1}^3 d_{ij} e_{ij} \quad \text{soit } p = d_{ij} e_{ij}. \quad (\text{D.8})$$

Produit vectoriel : Le produit vectoriel de deux vecteurs représenté par le symbole  $\wedge$  fournit un vecteur orthogonal aux deux autres dont les composantes s'expriment à l'aide du tenseur d'orientation  $\varepsilon$  :

$$\mathbf{w} = \mathbf{u} \wedge \mathbf{v} = \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix} \quad \text{soit } w_i = \varepsilon_{ijk} u_j v_k. \quad (\text{D.9})$$

Produit tensoriel généralisé : Le produit tensoriel — éventuellement contracté — généralise la plupart des produits précédents et s'applique à deux tenseurs d'ordre non nécessairement égaux. Ainsi le produit entre

un tenseur  $\mathbf{T}$  d'ordre  $p$  de terme général  $T_{i_1 \dots i_p}$  et un tenseur  $\mathbf{G}$  d'ordre  $q$  de terme général  $G_{j_1 \dots j_q}$  s'écrit au moyen des symboles de Kronecker  $\delta_i$  et  $\delta_j$

$$\begin{aligned} \mathbf{A} = \mathbf{T} \otimes \mathbf{G} &= T_{i_1 \dots i_p} (\delta_{i_1} \otimes \dots \otimes \delta_{i_p}) \otimes G_{j_1 \dots j_q} (\delta_{j_1} \otimes \dots \otimes \delta_{j_q}) \\ &= T_{i_1 \dots i_p} G_{j_1 \dots j_q} (\delta_{i_1} \otimes \dots \otimes \delta_{i_p} \otimes \delta_{j_1} \otimes \dots \otimes \delta_{j_q}). \end{aligned} \quad (\text{D.10})$$

Le résultat est un tenseur  $\mathbf{A}$  d'ordre  $p + q$  dont le terme général est donné par  $A_{i_1 \dots i_p, j_1 \dots j_q} = T_{i_1 \dots i_p} G_{j_1 \dots j_q}$ . Par ailleurs, le symbole  $\overline{\otimes}$  désignera le produit tensoriel contracté une fois. Il suffit de le définir pour le symbole de Kronecker par la relation

$$\delta_i \overline{\otimes} \delta_j = \delta_{ij}. \quad (\text{D.11})$$

L'opération de contraction revient alors à contracter un indice puisque tous les termes d'indice  $j$  d'une expression seront nuls si  $j \neq i$ . Elle peut être effectuée 2 fois, 3 fois, ...  $n$  fois, le nombre de contractions étant symbolisé par autant de barres au-dessus du symbole de produit tensoriel. On retrouve alors un certain nombre de définitions données auparavant avec les notations usuelles pour des tenseurs d'ordre 1 ou 2 :

$$\delta_{ij} = \delta_i \overline{\otimes} \delta_j, \quad (\text{D.12a})$$

$$\mathbf{I} = \delta_{ij} \delta_i \otimes \delta_j, \quad (\text{D.12b})$$

$$\mathbf{D} \mathbf{u} = \mathbf{D} \overline{\otimes} \mathbf{u}, \quad (\text{D.12c})$$

$$\mathbf{D} \mathbf{E} = \mathbf{D} \overline{\otimes} \mathbf{E}, \quad (\text{D.12d})$$

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \overline{\otimes} \mathbf{v}, \quad (\text{D.12e})$$

$$\mathbf{D} : \mathbf{E} = \mathbf{D} \overline{\otimes} \mathbf{E}. \quad (\text{D.12f})$$

La généralisation du produit tensoriel sera particulièrement utile pour travailler avec des tenseurs d'ordre supérieur à 2 où les produits ne sont pas spécifiés par des symboles particuliers.

## D.2 Opérateurs différentiels

On précise ici les expressions des opérateurs de dérivées spatiales communément employés en mécanique des fluides pour un domaine de l'espace compris dans  $\mathbb{R}^3$ . Seules les expressions en coordonnées cartésiennes sont présentées. On notera les directions de l'espace  $(x_1, x_2, x_3) = (x, y, z)$ , la première forme étant celle adoptée pour les écritures indicielles.

Opérateur nabla : Il s'agit d'un vecteur contenant les opérateurs de dérivées partielles spatiales dans les trois directions de l'espace, tel que

$$\nabla = \begin{bmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{bmatrix}. \quad (\text{D.13})$$

Il n'est jamais utilisé seul mais est destiné à s'appliquer à des tenseurs d'ordre quelconque. En particulier, il permet de définir un certain nombre d'opérateurs de dérivation couramment rencontrés en physique.

Opérateur gradient : Le gradient d'un tenseur fournit les dérivées dans les trois directions de l'espace des différentes composantes constituant le tenseur. Il augmente donc d'un ordre le tenseur sur lequel il agit. On le note **grad** ou encore à l'aide de l'opérateur nabla  $\nabla$ . Le gradient d'un scalaire est le vecteur

$$\mathbf{grad} p = \nabla p = \begin{bmatrix} \partial p / \partial x \\ \partial p / \partial y \\ \partial p / \partial z \end{bmatrix} \quad \text{tel que} \quad (\nabla p)_i = \frac{\partial p}{\partial x_i}. \quad (\text{D.14})$$

Le gradient d'un vecteur donne la matrice

$$\mathbf{grad} \mathbf{u} = \nabla \mathbf{u} = \begin{bmatrix} \nabla^T u_1 \\ \nabla^T u_2 \\ \nabla^T u_3 \end{bmatrix} = \begin{bmatrix} \partial u_1 / \partial x & \partial u_1 / \partial y & \partial u_1 / \partial z \\ \partial u_2 / \partial x & \partial u_2 / \partial y & \partial u_2 / \partial z \\ \partial u_3 / \partial x & \partial u_3 / \partial y & \partial u_3 / \partial z \end{bmatrix} \quad \text{telle que} \quad (\nabla \mathbf{u})_{ij} = \frac{\partial u_i}{\partial x_j}. \quad (\text{D.15})$$



Chaque ligne de la matrice obtenue correspond au gradient de la composante scalaire  $u_i$ . On peut donc interpréter l'action du gradient sur un vecteur comme le produit tensoriel  $\mathbf{u} \otimes \nabla$ . On introduit par ailleurs l'opérateur  $\nabla^T \mathbf{u}$  qui correspond à la transposée  $(\nabla \mathbf{u})^T$  pour alléger les écritures.

Opérateur divergence : L'opérateur de divergence réduit d'un ordre les tenseurs sur lesquels il agit. On le note  $\text{div}$  ou encore à l'aide de l'opérateur  $\nabla$ . La divergence d'un vecteur donne donc le scalaire

$$\text{div } \mathbf{u} = \nabla \cdot \mathbf{u} = \frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} + \frac{\partial u_3}{\partial z} \quad \text{tel que} \quad \text{div } \mathbf{u} = \frac{\partial u_i}{\partial x_i}. \quad (\text{D.16})$$

La divergence d'un vecteur correspond en fait au produit scalaire de l'opérateur nabla avec le vecteur  $\mathbf{u}$ . Lorsque la divergence s'applique à une matrice, on obtient le vecteur

$$\text{div } \mathbf{D} = \begin{bmatrix} \partial d_{11}/\partial x + \partial d_{12}/\partial y + \partial d_{13}/\partial z \\ \partial d_{21}/\partial x + \partial d_{22}/\partial y + \partial d_{23}/\partial z \\ \partial d_{31}/\partial x + \partial d_{32}/\partial y + \partial d_{33}/\partial z \end{bmatrix} \quad \text{tel que} \quad (\text{div } \mathbf{D})_i = \frac{\partial d_{ij}}{\partial x_j} \quad (\text{D.17})$$

dans lequel chaque composante correspond à la divergence des lignes de la matrice.

Opérateur Laplacien : L'opérateur Laplacien conserve l'ordre des tenseurs sur lesquels il agit. On le note  $\Delta$  ou encore  $\nabla^2$ . Le Laplacien d'un scalaire correspond donc au scalaire

$$\Delta p = \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} \quad \text{tel que} \quad \Delta p = \frac{\partial^2 p}{\partial x_i \partial x_i}, \quad (\text{D.18})$$

qui est défini par la somme des dérivées secondes dans les trois directions. Le Laplacien d'un vecteur correspond quant à lui au Laplacien des composantes du vecteur :

$$\Delta \mathbf{u} = \begin{bmatrix} \partial^2 u_1/\partial x^2 + \partial^2 u_1/\partial y^2 + \partial^2 u_1/\partial z^2 \\ \partial^2 u_2/\partial x^2 + \partial^2 u_2/\partial y^2 + \partial^2 u_2/\partial z^2 \\ \partial^2 u_3/\partial x^2 + \partial^2 u_3/\partial y^2 + \partial^2 u_3/\partial z^2 \end{bmatrix} \quad \text{tel que} \quad (\Delta \mathbf{u})_i = \frac{\partial^2 u_i}{\partial x_j \partial x_j}. \quad (\text{D.19})$$

Enfin, par analogie, le Laplacien d'une matrice  $\mathbf{D}$  est une matrice dont le terme général  $\partial^2 d_{ij}/\partial x_k \partial x_k$  est le Laplacien des composantes de la matrice sur laquelle le Laplacien agit.

Opérateur rotationnel : Le rotationnel est noté  $\text{rot}$  ou encore à l'aide de l'opérateur nabla par l'intermédiaire du produit vectoriel tel que

$$\text{rot } \mathbf{u} = \nabla \wedge \mathbf{u} = \begin{bmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{bmatrix} \wedge \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} \partial u_3/\partial y - \partial u_2/\partial z \\ \partial u_1/\partial z - \partial u_3/\partial x \\ \partial u_2/\partial x - \partial u_1/\partial y \end{bmatrix} \quad \text{avec} \quad (\text{rot } \mathbf{u})_i = \varepsilon_{ijk} \frac{\partial u_k}{\partial x_j}. \quad (\text{D.20})$$

En mécanique des fluides on définit couramment le rotationnel de la vitesse  $\boldsymbol{\zeta} = \text{rot } \mathbf{u} = \nabla \wedge \mathbf{u}$  qui représente la vortacité.

Opérateur d'advection  $\mathbf{u} \cdot \nabla$  : L'opérateur  $\mathbf{u} \cdot \nabla$  est couramment employé pour écrire les termes de convection en mécanique des fluides. Il est défini comme le produit scalaire entre un vecteur  $\mathbf{u}$  désignant en pratique la vitesse de convection et l'opérateur  $\nabla$  tel que

$$\mathbf{u} \cdot \nabla = \begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix} \cdot \begin{bmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{bmatrix} = u_x \frac{\partial}{\partial x} + u_y \frac{\partial}{\partial y} + u_z \frac{\partial}{\partial z}. \quad (\text{D.21})$$

Il agit en général sur un scalaire ou un vecteur, auquel cas il prend respectivement la forme

$$\mathbf{u} \cdot \nabla p = u_x \frac{\partial p}{\partial x} + u_y \frac{\partial p}{\partial y} + u_z \frac{\partial p}{\partial z}, \quad (\text{D.22})$$

ou encore

$$\mathbf{u} \cdot \nabla \mathbf{v} = \begin{bmatrix} \mathbf{u} \cdot \nabla v_x \\ \mathbf{u} \cdot \nabla v_y \\ \mathbf{u} \cdot \nabla v_z \end{bmatrix} = \begin{bmatrix} u_x \partial v_x / \partial x + u_y \partial v_x / \partial y + u_z \partial v_x / \partial z \\ u_x \partial v_y / \partial x + u_y \partial v_y / \partial y + u_z \partial v_y / \partial z \\ u_x \partial v_z / \partial x + u_y \partial v_z / \partial y + u_z \partial v_z / \partial z \end{bmatrix} \quad \text{tel que} \quad (\mathbf{u} \cdot \nabla \mathbf{v})_i = u_j \frac{\partial v_i}{\partial x_j}. \quad (\text{D.23})$$

L'opérateur  $\mathbf{u} \cdot \nabla$  s'applique donc à chaque composante du vecteur sur lequel il s'applique. L'opérateur intervient notamment dans la définition de la dérivée matérielle donnée à l'équation (4.4).

### D.3 Relations usuelles entre les opérateurs

Gradient :

$$\mathbf{grad} (\Delta p) = \Delta (\mathbf{grad} p) \quad (\text{D.24a})$$

$$\mathbf{grad} (pq) = p \mathbf{grad} q + q \mathbf{grad} p \quad (\text{D.24b})$$

$$\mathbf{grad} (p \mathbf{u}) = p \mathbf{grad} \mathbf{u} + \mathbf{u} \otimes \mathbf{grad} p \quad (\text{D.24c})$$

$$\mathbf{grad} (\mathbf{u} \cdot \mathbf{v}) = (\mathbf{grad}^T \mathbf{u}) \mathbf{v} + (\mathbf{grad}^T \mathbf{v}) \mathbf{u} \quad (\text{D.24d})$$

$$(\mathbf{grad} \mathbf{u}) \mathbf{u} = \mathbf{grad} \left( \frac{\mathbf{u} \cdot \mathbf{u}}{2} \right) + (\mathbf{rot} \mathbf{u}) \wedge \mathbf{u} \quad (\text{D.24e})$$

Divergence :

$$\text{div} \mathbf{u} = \text{Tr}(\mathbf{grad} \mathbf{u}) \quad \text{où } \text{Tr}(\cdot) \text{ est l'opérateur de trace} \quad (\text{D.25a})$$

$$\text{div} (\Delta \mathbf{u}) = \Delta (\text{div} \mathbf{u}) \quad (\text{D.25b})$$

$$\text{div} (\mathbf{rot} \mathbf{u}) = 0 \quad (\text{D.25c})$$

$$\text{div} (p \mathbf{u}) = p \text{div} \mathbf{u} + \mathbf{u} \cdot \mathbf{grad} p \quad (\text{D.25d})$$

$$\text{div} (p \mathbf{D}) = p \text{div} \mathbf{D} + \mathbf{D} \mathbf{grad} p \quad (\text{D.25e})$$

$$\text{div} (\mathbf{u}^T \mathbf{D}) = \mathbf{u} \cdot \text{div} \mathbf{D} + \mathbf{grad} \mathbf{u} : \mathbf{D} \quad (\text{D.25f})$$

$$\text{div} (p \mathbf{I}) = \mathbf{grad} p \quad (\text{D.25g})$$

$$\text{div} (\mathbf{u} \wedge \mathbf{v}) = \mathbf{v} \cdot \mathbf{rot} \mathbf{u} - \mathbf{u} \cdot \mathbf{rot} \mathbf{v} \quad (\text{D.25h})$$

$$\text{div} (\mathbf{u} \otimes \mathbf{v}) = \mathbf{u} \text{div} \mathbf{v} + (\mathbf{grad} \mathbf{u}) \mathbf{v} \quad (\text{D.25i})$$

Laplacien :

$$\Delta p = \text{div} (\mathbf{grad} p) \quad (\text{D.26a})$$

$$\Delta \mathbf{u} = \text{div} (\mathbf{grad} \mathbf{u}) = \mathbf{grad} (\text{div} \mathbf{u}) - \mathbf{rot} (\mathbf{rot} \mathbf{u}) \quad (\text{D.26b})$$

$$\Delta (pq) = p \Delta q + q \Delta p + 2 \mathbf{grad} p \cdot \mathbf{grad} q \quad (\text{D.26c})$$

$$\Delta (p \mathbf{u}) = p \Delta \mathbf{u} + \mathbf{u} \Delta p + 2 (\mathbf{grad} \mathbf{u}) \mathbf{grad} p \quad (\text{D.26d})$$

Rotationnel :

$$\mathbf{rot} (\mathbf{grad} p) = \mathbf{0} \quad (\text{D.27a})$$

$$\mathbf{rot} (\Delta \mathbf{u}) = \Delta (\mathbf{rot} \mathbf{u}) \quad (\text{D.27b})$$

$$\mathbf{rot} (p \mathbf{u}) = p \mathbf{rot} \mathbf{u} + \mathbf{grad} p \wedge \mathbf{u} \quad (\text{D.27c})$$

$$\mathbf{rot} (\mathbf{u} \wedge \mathbf{v}) = \mathbf{u} \text{div} \mathbf{v} - \mathbf{v} \text{div} \mathbf{u} + (\mathbf{grad} \mathbf{u}) \mathbf{v} - (\mathbf{grad} \mathbf{v}) \mathbf{u} \quad (\text{D.27d})$$

## D.4 Théorèmes de Green-Ostrogradski et d'intégration par parties

Théorème de Green-Ostrogradski : Le théorème de Green-Ostrogradski, aussi appelé théorème de flux-divergence, est un théorème reliant la divergence d'un champ vectoriel au flux défini par ce champ. Il découle du théorème de Stokes qui est un résultat central sur l'intégration de formes différentielles en analyse vectorielle. Il s'énonce pour un domaine de l'espace  $\Omega \subset \mathbb{R}^3$  de frontière  $\partial\Omega$  suffisamment régulière ( $C^1$  par morceaux), sur laquelle on peut définir en tout point une normale extérieure  $\mathbf{n}$ . Le théorème de Green-Ostrogradski stipule alors que le flux du champ vectoriel  $\mathbf{u}$  à travers la surface fermée  $\partial\Omega$  est égal à la divergence de ce même champ vectoriel dans le volume  $\Omega$  délimité par cette surface. Cela se traduit par la relation

$$\int_{\Omega} \operatorname{div} \mathbf{u} \, d\Omega = \oint_{\partial\Omega} \mathbf{u} \cdot \mathbf{n} \, d\partial\Omega. \quad (\text{D.28})$$

L'expression précédente s'applique aussi à un scalaire  $p$  ou une matrice  $\mathbf{D}$  afin d'obtenir des formules permettant de passer d'intégrales volumiques sur  $\Omega$  à des intégrales sur la surface fermée  $\partial\Omega$  à l'aide des opérateurs différentiels introduits ci-dessus. On peut ainsi montrer que pour un scalaire  $p$ ,

$$\int_{\Omega} \mathbf{grad} p \, d\Omega = \oint_{\partial\Omega} p \mathbf{n} \, d\partial\Omega \quad (\text{D.29})$$

et pour une matrice  $\mathbf{D}$ ,

$$\int_{\Omega} \operatorname{div} \mathbf{D} \, d\Omega = \oint_{\partial\Omega} \mathbf{D} \mathbf{n} \, d\partial\Omega. \quad (\text{D.30})$$

Intégration par parties : Le théorème de Green-Ostrogradski permet de déduire une généralisation de l'intégration par parties lorsque l'intégrale porte sur un produit de champs vectoriels faisant intervenir une divergence. En particulier, il nous sera utile de disposer de formules d'intégrations par parties pour transformer des intégrales portant sur des champs vectoriels définis par  $p \operatorname{div} \mathbf{D} \cdot \mathbf{u}$  ou encore  $p \operatorname{div} [\nabla(qr)]$ , où  $p, q$  et  $r$  sont des scalaires,  $\mathbf{u}$  est un vecteur et  $\mathbf{D}$  une matrice. Pour pouvoir utiliser le théorème de Green-Ostrogradski, il faut établir une relation entre le champ vectoriel qu'on cherche à transformer et une divergence portant sur un champ associé. Dans le premier cas, il suffit de considérer la divergence du vecteur  $\mathbf{v} = p \mathbf{u}^T \mathbf{D}$  puisque les relations usuelles (D.25f) et (D.24c), conduisent à

$$p \operatorname{div} \mathbf{D} \cdot \mathbf{u} = \operatorname{div} (p \mathbf{u}^T \mathbf{D}) - p \nabla \mathbf{u} : \mathbf{D} - (\mathbf{u} \otimes \nabla p) : \mathbf{D}. \quad (\text{D.31})$$

Par conséquent, en intégrant sur le domaine  $\Omega$ , l'expression précédente se transforme en utilisant le théorème de Green-Ostrogradski pour le premier terme du membre de droite, ce qui donne

$$\boxed{\int_{\Omega} p \operatorname{div} \mathbf{D} \cdot \mathbf{u} \, d\Omega = \oint_{\partial\Omega} (p \mathbf{u}^T \mathbf{D}) \mathbf{n} \, d\partial\Omega - \int_{\Omega} p \nabla \mathbf{u} : \mathbf{D} \, d\Omega - \int_{\Omega} (\mathbf{u} \otimes \nabla p) : \mathbf{D} \, d\Omega}. \quad (\text{D.32})$$

De la même manière, il suffit de considérer la divergence du vecteur  $\mathbf{v} = p \nabla(qr)$  pour le second champ vectoriel à transformer. En effet, les relations usuelles (D.25d) et (D.24b) conduisent à

$$p \operatorname{div} [\nabla(qr)] = \operatorname{div} [p \nabla(qr)] - q \nabla r \cdot \nabla p - r \nabla q \cdot \nabla p. \quad (\text{D.33})$$

La formule d'intégration par parties s'obtient alors comme précédemment en intégrant sur le domaine  $\Omega$ , et en utilisant le théorème de Green-Ostrogradski pour le premier terme du membre de droite ce qui donne

$$\boxed{\int_{\Omega} p \operatorname{div} [\nabla(qr)] \, d\Omega = \oint_{\partial\Omega} p \nabla(qr) \cdot \mathbf{n} \, d\partial\Omega - \int_{\Omega} q \nabla r \cdot \nabla p \, d\Omega - \int_{\Omega} r \nabla q \cdot \nabla p \, d\Omega}. \quad (\text{D.34})$$

## Changements de repères et de variables et discrétisation des équations de Navier-Stokes

---

### Sommaire

---

<b>E.1</b>	<b>Écriture des équations de Navier-Stokes dans un repère non galiléen . . . . .</b>	<b>244</b>
E.1.1	Changement de base et repérage d'un point dans un repère entraîné . . . . .	244
E.1.1.1	Repérage d'un point dans un repère entraîné . . . . .	244
E.1.1.2	Loi de composition des vitesses et dérivation temporelle dans le repère entraîné . . . . .	245
E.1.1.3	Règles de changement de base pour les opérateurs différentiels . . . . .	247
E.1.2	Équations de Navier-Stokes en repère entraîné avec les composantes de la vitesse absolue . . . . .	249
E.1.2.1	Équation de conservation de la masse en repère entraîné . . . . .	249
E.1.2.2	Équation de conservation de la quantité de mouvement en repère entraîné . . . . .	249
E.1.2.3	Équation de conservation de l'énergie en repère entraîné . . . . .	250
E.1.3	Équations de Navier-Stokes en repère entraîné avec les composantes de la vitesse relative . . . . .	251
E.1.3.1	Équation de conservation de la masse en repère entraîné . . . . .	251
E.1.3.2	Équation de conservation de la quantité de mouvement en repère entraîné . . . . .	251
E.1.3.3	Équation de conservation de l'énergie en repère entraîné . . . . .	254
<b>E.2</b>	<b>Impact du jeu de variables sur la forme des équations de Navier-Stokes . . . . .</b>	<b>258</b>
E.2.1	Écriture générique des équations du fluide . . . . .	258
E.2.2	Variables quasi-conservatives $\mathbf{q} = [\rho, \mathbf{u}, e_T]$ et $\mathbf{q} = [\rho, \mathbf{u}, e_T]$ . . . . .	259
E.2.3	Variables thermiques $\mathbf{q} = [\rho, \mathbf{u}, \theta]$ . . . . .	260
E.2.4	Variables primitives $\mathbf{q} = [\rho, \mathbf{u}, p]$ . . . . .	261
E.2.5	Variables primitives modifiées $\mathbf{q} = [\vartheta, \mathbf{u}, p]$ . . . . .	261
<b>E.3</b>	<b>Discrétisation par la méthode des Volumes Finis . . . . .</b>	<b>262</b>
E.3.1	Discrétisation des équations de Navier-Stokes . . . . .	262
E.3.1.1	Discrétisation en temps . . . . .	264
E.3.1.2	Discrétisation des flux convectifs et diffusifs . . . . .	264
E.3.2	Approximation du gradient en Volumes Finis . . . . .	265

---

Dans cette annexe, on détaille dans un premier temps la manière dont les équations de Navier-Stokes se transposent dans un domaine non galiléen pour établir les équations (4.31) et (4.33). Ensuite, on décrit la forme des équations de Navier-Stokes écrites avec différents jeux de variables pour tenter d'établir la meilleure forme possible pour construire ultérieurement le modèle d'ordre réduit POD-Galerkin. Enfin, quelques détails relatifs à la discrétisation des équations de Navier-Stokes seront fournis pour terminer.

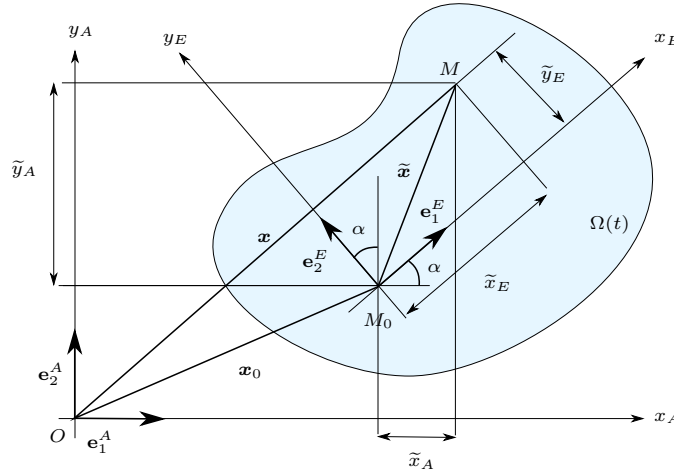
## E.1 Écriture des équations de Navier-Stokes dans un repère non galiléen avec les variables conservatives

### E.1.1 Changement de base et repérage d'un point dans un repère entraîné

La transposition des équations dans le repère entraîné repose tout d'abord sur les relations de changement de base rappelées dans l'annexe A.1 pour un vecteur et une matrice quelconques. Celles-ci permettent en effet d'établir la relation entre les positions puis les vitesses exprimées dans le repère absolu et entraîné. Enfin, on termine en donnant les règles de changement de base pour les opérateurs différentiels qui interviennent dans les équations de Navier-Stokes.

#### E.1.1.1 Repérage d'un point dans un repère entraîné

On considère un domaine de l'espace  $\Omega(t) \subset \mathbb{R}^3$  qui se déplace par rapport à un repère absolu  $\mathcal{B}_A$  défini par le point d'origine  $O$  et la base  $\mathcal{B}_A = (\mathbf{e}_1^A, \mathbf{e}_2^A, \mathbf{e}_3^A)$ . On associe au domaine  $\Omega(t)$  un repère local  $\mathcal{B}_E$  défini par le point d'origine  $M_0$  et la base  $\mathcal{B}_E = (\mathbf{e}_1^E, \mathbf{e}_2^E, \mathbf{e}_3^E)$ . On supposera pour la suite que les bases  $\mathcal{B}_A$  et  $\mathcal{B}_E$  sont orthonormées directes. On note enfin  $x_{\bullet,i}$  les directions spatiales relatives à la base  $\mathcal{B}_{\bullet}$ , telles que  $x_{\bullet,1} = x_{\bullet}, x_{\bullet,2} = y_{\bullet}, x_{\bullet,3} = z_{\bullet}$ , où le symbole  $\bullet$  désigne l'indice  $A$  ou  $E$ . La direction  $x_{\bullet,i}$  est celle du vecteur de base  $\mathbf{e}_i^{\bullet}$ . Le schéma de la figure E.1 résume les notations introduites ci-dessus et dans ce qui suit.



**Figure E.1 :** Notations pour un domaine  $\Omega(t) \subset \mathbb{R}^3$  relatives au repère absolu  $\mathcal{B}_A = (O, \mathcal{B}_A)$  muni de la base  $\mathcal{B}_A = (\mathbf{e}_1^A, \mathbf{e}_2^A)$  et au repère local  $\mathcal{B}_E = (M_0, \mathcal{B}_E)$  muni de la base  $\mathcal{B}_E = (\mathbf{e}_1^E, \mathbf{e}_2^E)$ . Le point  $M$  est repéré par sa position absolue  $\mathbf{x}$  ou sa position relative  $\tilde{\mathbf{x}}$  dans le repère local incliné d'un angle  $\alpha$  par rapport au repère absolu.

La position absolue d'un point  $M$  du domaine  $\Omega(t)$  est donnée par le vecteur distance  $\overrightarrow{OM}$  par rapport à l'origine du repère absolu  $O$ . La position relative  $\overrightarrow{M_0M}$  apparaît avec la relation de Chasles :  $\overrightarrow{OM} = \overrightarrow{OM_0} + \overrightarrow{M_0M}$ . On notera  $\mathbf{x} = \overrightarrow{OM}$  la position absolue du point  $M$ ,  $\mathbf{x}_0 = \overrightarrow{OM_0}$  la position absolue de l'origine  $M_0$  du repère local et  $\tilde{\mathbf{x}} = \overrightarrow{M_0M}$  la position relative du point  $M$  par rapport à l'origine  $M_0$  du repère local, si bien que la relation de Chasles s'écrit dans n'importe quelle base de l'espace

$$\mathbf{x} = \mathbf{x}_0 + \tilde{\mathbf{x}}. \quad (\text{E.1})$$

Pour préciser la relation précédente dans une certaine base de l'espace, on introduit les représentations d'un vecteur quelconque  $\mathbf{r} \in \mathbb{R}^3$  dans la base absolue  $\mathcal{B}_A$  et dans la base relative  $\mathcal{B}_E$  de sorte que

$$\mathbf{r}_A = \text{Mat}_{\mathcal{B}_A}(\mathbf{r}) = [r_{A,1}, r_{A,2}, r_{A,3}]^T \quad \text{et} \quad \mathbf{r}_E = \text{Mat}_{\mathcal{B}_E}(\mathbf{r}) = [r_{E,1}, r_{E,2}, r_{E,3}]^T. \quad (\text{E.2})$$

Par conséquent, les composantes des vecteurs dans les deux bases sont liées par

$$\mathbf{r}_A = \mathbf{R} \mathbf{r}_E \quad (\text{E.3})$$

selon la relation (A.9) appliquée ici avec la matrice de changement de base  $\mathbf{R} = \mathbf{P}_{\mathcal{B}_E \rightarrow \mathcal{B}_A}$ .

Dans la relation (E.1), on exprime alors les vecteurs  $\mathbf{x}$  et  $\mathbf{x}_0$  dans le repère absolu, mais il est plus judicieux d'employer la représentation du vecteur position relative dans la base locale  $\mathcal{B}_E$ , puisque si le domaine  $\Omega(t)$  représente un corps rigide, alors la position du point  $M$  sera indépendante du temps dans ce repère. Par conséquent la relation E.1 se réécrit

$$\mathbf{x}_A = \mathbf{x}_{0,A} + \tilde{\mathbf{x}}_A = \mathbf{x}_{0,A} + \mathbf{R} \tilde{\mathbf{x}}_E. \quad (\text{E.4})$$

Le changement de base est ici uniquement dû à une rotation de la base  $\mathcal{B}_E$  par rapport à la base  $\mathcal{B}_A$ . En effet une translation n'a aucune influence sur les vecteurs de base qui restent colinéaires. D'après la définition de l'angle  $\alpha(t)$  donnée sur la figure E.1, les vecteurs de la base  $\mathcal{B}_E$  s'expriment en fonction de ceux de la base  $\mathcal{B}_A$  et réciproquement. Pour une rotation dans le plan, les relations s'écrivent  $\mathbf{e}_1^E = \cos \alpha(t) \mathbf{e}_1^A + \sin \alpha(t) \mathbf{e}_2^A$  et  $\mathbf{e}_2^E = -\sin \alpha(t) \mathbf{e}_1^A + \cos \alpha(t) \mathbf{e}_2^A$ . Ainsi, la matrice de changement de base est simplement donnée par l'expression

$$\mathbf{R}_{2D}(\alpha) = \mathbf{P}_{\mathcal{B}_E \rightarrow \mathcal{B}_A} = \begin{bmatrix} \cos \alpha(t) & -\sin \alpha(t) \\ \sin \alpha(t) & \cos \alpha(t) \end{bmatrix}, \quad (\text{E.5})$$

où chaque colonne correspond au  $j$ -ième vecteur de la base  $\mathcal{B}_E$  exprimé sur les vecteurs de la base  $\mathcal{B}_A$ . De façon plus générale, une rotation de la base  $\mathcal{B}_E$  par rapport à la base  $\mathcal{B}_A$  dans  $\mathbb{R}^3$  s'obtient par une combinaison d'au plus trois rotations autour des différents axes. Les matrices de rotation autour des axes  $x$ ,  $y$  et  $z$  sont données en trois dimensions pour un angle  $\alpha(t)$  par

$$\mathbf{R}_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha(t) & \sin \alpha(t) \\ 0 & -\sin \alpha(t) & \cos \alpha(t) \end{bmatrix}, \quad \mathbf{R}_y(\alpha) = \begin{bmatrix} \cos \alpha(t) & 0 & -\sin \alpha(t) \\ 0 & 1 & 0 \\ \sin \alpha(t) & 0 & \cos \alpha(t) \end{bmatrix} \quad \text{et} \quad (\text{E.6})$$

$$\mathbf{R}_z(\alpha) = \begin{bmatrix} \cos \alpha(t) & \sin \alpha(t) & 0 \\ -\sin \alpha(t) & \cos \alpha(t) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

La rotation générale s'obtient simplement par multiplication des différentes matrices entre elles. Ces matrices de rotation appartiennent au groupe spécial orthogonal  $SO_3(\mathbb{R})$  dans lequel les matrices sont orthogonales et possèdent un déterminant égal à 1. Ces matrices vérifient donc les propriétés

$$\mathbf{R}^T \mathbf{R} = \mathbf{I} = \mathbf{R} \mathbf{R}^T, \quad (\text{E.7a})$$

$$\mathbf{R}^{-1} = \mathbf{R}^T. \quad (\text{E.7b})$$

Enfin, puisque la matrice de changement de base est orthogonale, la réciproque de la relation de changement de base donnée à l'équation (E.3) s'écrit

$$\mathbf{r}_E = \mathbf{R}^{-1} \mathbf{r}_A = \mathbf{R}^T \mathbf{r}_A. \quad (\text{E.8})$$

### E.1.1.2 Loi de composition des vitesses et dérivation temporelle dans le repère entraîné

Il est dans un premier temps nécessaire d'établir des relations semblables à (E.4) pour la vitesse des nœuds du maillage qui apparaît dans les équations de Navier-Stokes écrites en formulation ALE. Le vecteur vitesse absolue d'un point  $M$  dans le repère absolu est défini classiquement par la dérivée temporelle du vecteur position absolue exprimé dans la base  $\mathcal{B}_A$  tel que

$$\mathbf{s}_{a,A}(M) = \frac{d\mathbf{x}_A}{dt} = \dot{\mathbf{x}}_A. \quad (\text{E.9})$$

Connaissant les composantes du vecteur vitesse absolue  $\mathbf{s}_{a,A}(M)$  dans la base  $\mathcal{B}_A$ , l'expression du vecteur vitesse absolue  $\mathbf{s}_{a,E}(M)$  exprimé dans la base  $\mathcal{B}_E$  se déduit trivialement de la relation (E.8) puisque celle-ci est valable pour n'importe quel vecteur. On souhaite cependant obtenir la loi de composition des vitesses reliant la vitesse absolue définie à l'équation (E.9) à une vitesse relative qui reste à définir. La dérivation en temps de la relation (E.4) donne ainsi la relation

$$\mathbf{s}_{a,A}(M) = \frac{d\mathbf{x}_A}{dt} = \frac{d}{dt} (\mathbf{x}_{0,A} + \mathbf{R} \tilde{\mathbf{x}}_E) = \frac{d\mathbf{x}_{0,A}}{dt} + \frac{d\mathbf{R}}{dt} \tilde{\mathbf{x}}_E + \mathbf{R} \frac{d\tilde{\mathbf{x}}_E}{dt}, \quad (\text{E.10})$$

où le premier terme correspond à la *vitesse d'entraînement en translation* du repère mobile, exprimée dans le repère absolu  $\mathcal{R}_A$ . Elle ne dépend en fait que de la position de l'origine  $M_0$  du repère local  $\mathcal{R}_E$  et sera notée par la suite

$$\mathbf{s}_{0,A}(M_0) = \frac{d\mathbf{x}_{0,A}}{dt} = \dot{\mathbf{x}}_{0,A}. \quad (\text{E.11})$$

Les deux derniers termes de l'équation (E.10) font intervenir les composantes du vecteur position relative  $\tilde{\mathbf{x}}_E$  exprimé dans  $\mathcal{R}_E$ . Pour conserver une expression dans laquelle seules les composantes des vecteurs position exprimés dans  $\mathcal{R}_A$  apparaissent, il faut remplacer la position relative exprimée dans  $\mathcal{R}_E$  par son homologue dans  $\mathcal{R}_A$ . Ainsi, dans le second terme, on remplace la position relative par  $\tilde{\mathbf{x}}_E = \mathbf{R}^\top (\mathbf{x}_A - \mathbf{x}_{0,A}) = \mathbf{R}^\top \tilde{\mathbf{x}}_A$  en utilisant l'expression (E.4) ou simplement (E.8). On introduit ensuite la matrice

$$\boldsymbol{\Omega}_A = \frac{d\mathbf{R}}{dt} \mathbf{R}^\top, \quad (\text{E.12})$$

de sorte que le second terme de (E.10) s'écrive

$$\frac{d\mathbf{R}}{dt} \tilde{\mathbf{x}}_E = \frac{d\mathbf{R}}{dt} \mathbf{R}^\top \tilde{\mathbf{x}}_A = \boldsymbol{\Omega}_A \tilde{\mathbf{x}}_A. \quad (\text{E.13})$$

La matrice  $\boldsymbol{\Omega}_A$  est appelée *matrice ou tenseur de la vitesse angulaire*. C'est un tenseur anti-symétrique puisque la matrice de changement de base est orthogonale. En effet avec la propriété (E.7a) on a

$$\frac{d\mathbf{R}\mathbf{R}^\top}{dt} = \frac{d\mathbf{I}}{dt} = 0 \Rightarrow \dot{\mathbf{R}}\mathbf{R}^\top + \mathbf{R}\dot{\mathbf{R}}^\top = 0 \Rightarrow \boldsymbol{\Omega}_A = -\boldsymbol{\Omega}_A^\top. \quad (\text{E.14})$$

En outre, toute matrice anti-symétrique réelle  $\boldsymbol{\Omega} \in \mathcal{M}(3, 3)$  comporte seulement trois composantes indépendantes. On a donc coutume d'associer au tenseur  $\boldsymbol{\Omega}$  un pseudo-vecteur  $\boldsymbol{\omega}$  contenant les trois composantes indépendantes de  $\boldsymbol{\Omega}$  tel que pour tout vecteur  $\mathbf{r} \in \mathbb{R}^3$

$$\boldsymbol{\Omega} \mathbf{r} = \boldsymbol{\omega} \wedge \mathbf{r}. \quad (\text{E.15})$$

Le vecteur  $\boldsymbol{\omega}$  ainsi introduit est le *vecteur vitesse angulaire* qui est défini à partir du tenseur de permutation  $\boldsymbol{\varepsilon}$  introduit à l'équation (D.2) par la relation

$$\omega_i = -\frac{1}{2} \varepsilon_{ijk} \Omega_{jk}. \quad (\text{E.16})$$

Le second terme de l'équation (E.10) transformé à l'expression (E.13) devient alors

$$\frac{d\mathbf{R}}{dt} \tilde{\mathbf{x}}_E = \boldsymbol{\Omega}_A \tilde{\mathbf{x}}_A = \boldsymbol{\omega}_A \wedge \tilde{\mathbf{x}}_A. \quad (\text{E.17})$$

Il est donc lié à la vitesse angulaire du repère entraîné, par le biais de l'opérateur  $\boldsymbol{\Omega}_A$  et constitue la *vitesse d'entraînement en rotation*. Quant au dernier terme de l'équation (E.10), il est aussi modifié en remplaçant la position relative  $\tilde{\mathbf{x}}_E$  exprimée dans  $\mathcal{R}_E$  par  $\tilde{\mathbf{x}}_A$  au moyen de la relation (E.8). On introduit auparavant l'opérateur de dérivation temporelle dans le repère entraîné

$$\frac{\delta \bullet}{\delta t} = \mathbf{R} \frac{d}{dt} (\mathbf{R}^\top \bullet). \quad (\text{E.18})$$

Ainsi, le dernier terme dans (E.10) représente la *vitesse relative* du point  $M$  dans le repère entraîné

$$\mathbf{R} \frac{d\tilde{\mathbf{x}}_E}{dt} = \mathbf{R} \frac{d}{dt} (\mathbf{R}^\top \tilde{\mathbf{x}}_A) = \frac{\delta \tilde{\mathbf{x}}_A}{\delta t}, \quad (\text{E.19})$$

et finalement, la loi de composition des vitesses s'écrit

$$\boxed{\mathbf{s}_{a,A}(M) = \frac{d\mathbf{x}_A}{dt} = \mathbf{s}_{0,A}(M_0) + \boldsymbol{\omega}_A \wedge \tilde{\mathbf{x}}_A + \frac{\delta \tilde{\mathbf{x}}_A}{\delta t}} \quad (\text{E.20})$$

en remplaçant dans l'équation (E.10) les expressions (E.11), (E.17) et (E.19) des différentes vitesses. La vitesse d'un point  $M$  se décompose donc d'une part en une vitesse d'entraînement  $\mathbf{s}_{e,A}$  associée uniquement au mouvement de corps rigide du domaine  $\Omega(t)$ , et d'autre part en une vitesse relative  $\mathbf{s}_{d,A}$  due à la déformation éventuelle du domaine. Ces deux vitesses sont données respectivement par

$$\begin{aligned}\mathbf{s}_{e,A}(M) &= \mathbf{s}_{0,A}(M_0) + \boldsymbol{\omega}_A \wedge \tilde{\mathbf{x}}_A \quad \text{et} \\ \mathbf{s}_{d,A}(M) &= \frac{\delta \tilde{\mathbf{x}}_A}{\delta t}.\end{aligned}\tag{E.21}$$

**Remarque E.1 : Formule générale de dérivation en repère mobile**

L'équation (E.20) peut être écrite uniquement en fonction de la position relative  $\tilde{\mathbf{x}}_A$ . En effet,  $\mathbf{s}_{e,A}(M) - \mathbf{s}_{0,A}(M_0) = d/dt(\mathbf{x}_A - \mathbf{x}_{0,A}) = d\tilde{\mathbf{x}}_A/dt$ . Par conséquent on obtient en posant pour tout vecteur  $\mathbf{r}_A$  exprimé dans le repère  $\mathcal{B}_A$  la formule générale de dérivation en repère mobile issue de l'équation (E.20) en posant  $\mathbf{r}_A = \tilde{\mathbf{x}}_A$  :

$$\frac{d\mathbf{r}_A}{dt} = \frac{\delta \mathbf{r}_A}{\delta t} + \boldsymbol{\omega}_A(\mathcal{R}_E/\mathcal{B}_A) \wedge \mathbf{r}_A.\tag{E.22}$$

**Remarque E.2 : Loi de composition des accélérations**

En dérivant une seconde fois la loi de composition des vitesses, on peut établir une loi de composition des accélérations dans laquelle l'accélération absolue  $\mathbf{a}_{a,A}(M)$  est définie par la dérivée seconde de la position absolue. Après un certain nombre de calculs, on parvient à l'expression

$$\boxed{\mathbf{a}_{a,A}(M) = \frac{d^2 \mathbf{x}_A}{dt^2} = \underbrace{\frac{d^2 \mathbf{x}_{0,A}}{dt^2} + \frac{\delta \boldsymbol{\omega}_A}{\delta t} \wedge \tilde{\mathbf{x}}_A + \boldsymbol{\omega}_A \wedge (\boldsymbol{\omega}_A \wedge \tilde{\mathbf{x}}_A)}_{\mathbf{a}_{e,A}} + \underbrace{2 \boldsymbol{\omega}_A \frac{\delta \tilde{\mathbf{x}}_A}{\delta t}}_{\mathbf{a}_{C,A}} + \underbrace{\frac{\delta^2 \tilde{\mathbf{x}}_A}{\delta t^2}}_{\mathbf{a}_{r,A}}},\tag{E.23}$$

dans laquelle apparaissent l'accélération d'entraînement  $\mathbf{a}_{e,A}$  liée au mouvement de corps rigide, l'accélération de Coriolis  $\mathbf{a}_{C,A}$  et l'accélération relative  $\mathbf{a}_{r,A}$ . Les deux premières accélérations sont « fictives » et proviennent simplement du fait que le système est étudié depuis un référentiel non galiléen. Pour un système de masse  $m$ , elles donnent naissance à la force de Coriolis  $\mathbf{F}_C = -m \mathbf{a}_{C,A}$  et à la force d'inertie d'entraînement  $\mathbf{F}_{ie} = -m \mathbf{a}_{e,A}$ , le signe moins provenant du fait que dans le principe fondamental de la dynamique  $m \mathbf{a} = \mathbf{F}$ , les accélérations sont dans le membre de gauche alors que les forces apparaissent dans le membre de droite. La force centrifuge est quant à elle un cas particulier de la force d'inertie d'entraînement lorsque le repère entraîné est en rotation uniforme, de sorte  $\mathbf{a}_{e,A}$  se réduit à  $\boldsymbol{\omega}_A \wedge (\boldsymbol{\omega}_A \wedge \tilde{\mathbf{x}}_A)$ ; dans ce cas, l'accélération est dirigée perpendiculairement à l'axe de rotation, ce qui justifie le nom donné à cette force. On parlera dans la suite d'effets gyroscopiques pour se référer aux effets induits par les accélérations de Coriolis et d'inertie d'entraînement.

**E.1.1.3 Règles de changement de base pour les opérateurs différentiels**

Avant de pouvoir écrire les équations de Navier-Stokes dans le repère local, il est nécessaire d'établir des relations entre les opérateurs aux dérivées partielles qui interviennent dans les équations de Navier-Stokes, notamment le gradient et la divergence. On considère un vecteur quelconque  $\mathbf{r} \in \mathbb{R}^3$  pour lequel on dispose des représentations (E.2) dans les bases  $\mathcal{B}_A$  et  $\mathcal{B}_E$ , ainsi que des relations de passage (E.3) et (E.8) entre les deux bases.

La dérivée spatiale dans la direction  $x_{\bullet,i}$  du vecteur  $\mathbf{r}$  exprimé dans la base  $\mathcal{B}_\bullet$  s'écrit avec la convention d'Einstein sur les indices répétés qui sera utilisée dans la suite

$$\frac{\partial \mathbf{r}}{\partial x_{\bullet,i}} = \frac{\partial}{\partial x_{\bullet,i}} (r_{\bullet,j} \mathbf{e}_j^\bullet) = \frac{\partial r_{\bullet,j}}{\partial x_{\bullet,i}} \mathbf{e}_j^\bullet + r_{\bullet,j} \underbrace{\frac{\partial \mathbf{e}_j^\bullet}{\partial x_{\bullet,i}}}_{=0}.\tag{E.24}$$

Le dernier terme est nul si on suppose comme ici que les vecteurs de base  $\mathbf{e}_j^\bullet$  sont indépendants de la position spatiale. Pour établir le lien entre les dérivées par rapport aux directions relatives aux deux bases, il suffit de mettre en évidence la relation entre la dérivée des vecteurs de base. Pour cela, on utilise l'expression



des vecteurs de la base  $\mathcal{B}_E$  dans  $\mathcal{B}_A$  telle que  $\mathbf{e}_j^E = R_{ij} \mathbf{e}_i^A$ , où les  $R_{ij}$  sont les composantes de la matrice de passage (E.3). On obtient alors la relation

$$\frac{\partial x_{E,j}}{\partial x_{A,i}} \equiv \frac{\partial \mathbf{e}_j^E}{\partial \mathbf{e}_i^A} = R_{kj} \underbrace{\frac{\partial \mathbf{e}_k^A}{\partial \mathbf{e}_i^A}}_{=\delta_{ik}} = R_{ij}, \quad (\text{E.25})$$

à partir de laquelle on déduit simplement les relations entre les opérateurs différentiels. Le gradient d'un scalaire  $s$  nécessite ainsi d'évaluer la dérivée  $\partial s / \partial x_{A,i}$ , ce qui s'écrit

$$\frac{\partial s}{\partial x_{A,i}} = \frac{\partial s}{\partial x_{E,k}} \frac{\partial x_{E,k}}{\partial x_{A,i}} = \frac{\partial s}{\partial x_{E,k}} \frac{\partial R_{mk} x_{A,m}}{\partial x_{A,i}} = R_{ik} \frac{\partial s}{\partial x_{E,k}}. \quad (\text{E.26})$$

Par conséquent, si on note  $\nabla_{\bullet}$  le gradient d'un scalaire  $s$  par rapport à la base  $\mathcal{B}_{\bullet}$ , on obtient la relation de changement de base

$$\nabla_A s = \mathbf{R} \nabla_E s, \quad (\text{E.27})$$

qui correspond à la relation classique de changement de base (A.9) en considérant le vecteur  $\mathbf{X}' = \nabla_E s$ .

La relation équivalente pour le gradient d'un vecteur s'obtient en exprimant les composantes  $r_{A,j}$  du vecteur  $\mathbf{r}$  dans la base  $\mathcal{B}_E$  utilisée pour la dérivation. On a alors

$$\frac{\partial r_{A,j}}{\partial x_{A,i}} = \frac{\partial r_{A,j}}{\partial x_{E,k}} \frac{\partial x_{E,k}}{\partial x_{A,i}} = \frac{\partial R_{jl} r_{E,l}}{\partial x_{E,k}} \frac{\partial R_{mk} x_{A,m}}{\partial x_{A,i}} = R_{jl} R_{mk} \frac{\partial r_{E,l}}{\partial x_{E,k}} \delta_{mi} = R_{jl} \frac{\partial r_{E,l}}{\partial x_{E,k}} R_{ik}. \quad (\text{E.28})$$

Ainsi, en notant  $\nabla_{\bullet} \mathbf{r}$  le gradient du vecteur  $\mathbf{r}$  exprimé dans la base  $\mathcal{B}_{\bullet}$  et dérivé par rapport aux directions de cette même base, et en remplaçant  $R_{ik}$  par  $R_{ki}^{\top}$  qui est le terme général de la matrice  $\mathbf{R}^{\top}$ , on obtient la relation

$$\nabla_A \mathbf{r}_A = \mathbf{R} (\nabla_E \mathbf{r}_E) \mathbf{R}^{\top}. \quad (\text{E.29})$$

Ici encore, on retrouve la relation de changement de base (A.11) avec la matrice  $\mathbf{M}' = \nabla_E \mathbf{r}_E$ . En effet, puisque le changement de base considéré ici est une rotation, la relation (E.7b) donne  $\mathbf{P}^{-1} = \mathbf{R}^{-1} = \mathbf{R}^{\top}$ .

La divergence d'un vecteur  $\mathbf{r}$  est quant à elle donnée par la relation

$$\text{div}_A \mathbf{r}_A = \frac{\partial r_{A,i}}{\partial x_{A,i}} = \frac{\partial r_{A,i}}{\partial x_{E,j}} \frac{\partial x_{E,j}}{\partial x_{A,i}} = \frac{\partial R_{ik} r_{E,k}}{\partial x_{E,j}} \frac{\partial R_{lj} x_{A,l}}{\partial x_{A,i}} = R_{ik} R_{lj} \frac{\partial r_{E,k}}{\partial x_{E,j}} \delta_{li} = R_{ik} R_{ij} \frac{\partial r_{E,k}}{\partial x_{E,j}}. \quad (\text{E.30})$$

En remplaçant  $R_{ik}$  par  $R_{ki}^{\top}$ , la relation entre la divergence calculée dans les bases  $\mathcal{B}_A$  et  $\mathcal{B}_E$  s'écrit en notation tensorielle

$$\text{div}_A \mathbf{r}_A = (\mathbf{R}^{\top} \mathbf{R}) : (\nabla_E \mathbf{r}_E). \quad (\text{E.31})$$

Encore une fois, on peut utiliser le fait que le changement de base est une rotation. Ainsi, la matrice  $\mathbf{R}$  est orthogonale et la propriété (E.7a) permet de récrire la relation précédente en  $\text{div}_A \mathbf{r}_A = \mathbf{I} : (\nabla_E \mathbf{r}_E)$  ce qui se simplifie en

$$\text{div}_A \mathbf{r}_A = \text{div}_E \mathbf{r}_E. \quad (\text{E.32})$$

La divergence d'un vecteur est donc préservée lorsque le changement de base est une rotation.

Enfin, la divergence d'une matrice s'écrit  $\text{div}_A \mathbf{M}_A = \text{div}_A (\mathbf{R} \mathbf{M}_E \mathbf{R}^{\top})$ . En notation indicielle, la composante  $i$  de  $\text{div}_A \mathbf{M}_A$  s'écrit

$$\begin{aligned} \text{div}_A \mathbf{M}_A &= \frac{\partial}{\partial x_{A,j}} (R_{ik} M_{kl} R_{jl}) = \frac{\partial}{\partial x_{E,m}} (R_{ik} M_{kl} R_{jl}) \frac{\partial x_{E,m}}{\partial x_{A,j}} = R_{ik} R_{jl} \frac{\partial M_{kl}}{\partial x_{E,m}} \frac{\partial R_{nm} x_{A,n}}{\partial x_{A,j}} \\ &= R_{ik} R_{jl} R_{nm} \delta_{nj} \frac{\partial M_{kl}}{\partial x_{E,m}} = R_{ik} (R_{lj}^{\top} R_{jm}) \frac{\partial M_{kl}}{\partial x_{E,m}} = R_{ik} \delta_{lm} \frac{\partial M_{kl}}{\partial x_{E,m}} = R_{ik} \frac{\partial M_{km}}{\partial x_{E,m}}. \end{aligned} \quad (\text{E.33})$$

La dernière égalité correspond au produit  $\mathbf{R} \text{div}_E \mathbf{M}$ , d'où la relation vectorielle

$$\text{div}_A \mathbf{M}_A = \mathbf{R} \text{div}_E \mathbf{M}_E. \quad (\text{E.34})$$

Intuitivement les relations (E.32) et (E.34) pouvaient être établies à partir du fait que la divergence d'un vecteur est un scalaire dont la valeur n'est pas affectée lors d'un changement de base, et que la divergence d'une matrice est un vecteur dont l'expression dans la base  $\mathcal{B}_E$  est donnée par la relation (E.3).

## E.1.2 Formulation intégrale des équations de Navier-Stokes en repère entraîné avec les composantes de la vitesse absolue

À partir des résultats précédents, on peut établir la formulation FALE-REUE des équations (4.31) de Navier-Stokes dans le repère entraîné  $\mathcal{R}_E$  avec les composantes de la vitesse absolue. Il suffit ainsi de remplacer dans les équations scalaires — de conservation de la masse et de l'énergie — les expressions dans le repère absolu de la vitesse absolue et des vecteurs normaux par leurs expressions dans le repère local au moyen de la relation E.3 ; l'équation vectorielle de conservation de la quantité de mouvement doit quant à elle être projetée auparavant dans le repère local. Les détails sont développés ci-dessous.

### E.1.2.1 Équation de conservation de la masse en repère entraîné

Le bilan de conservation de la masse écrit en formulation ALE avait été donné à l'équation (4.27) dans le repère absolu avec les composantes de la vitesse absolue. Le terme sous l'intégrale surfacique est alors simplement remplacé par les expressions de la vitesse et de la normale dans  $\mathcal{R}_E$  :  $\rho(\mathbf{u}_A - \mathbf{s}_A)^\top \mathbf{n}_A = \rho[\mathbf{R}(\mathbf{u}_E - \mathbf{s}_E)]^\top \mathbf{R}\mathbf{n}_E = \rho(\mathbf{u}_E - \mathbf{s}_E)^\top (\mathbf{R}^\top \mathbf{R}) \mathbf{n}_E$ . Finalement, puisque la matrice de changement de base est orthogonale, la propriété (E.7a) permet de simplifier l'équation de bilan en

$$\boxed{\left. \frac{\partial}{\partial t} \right|_{\mathcal{X}} \int_{\Omega(t)} \rho \, d\Omega + \oint_{\partial\Omega(t)} \rho (\mathbf{u}_E - \mathbf{s}_E)^\top \mathbf{n}_E \, d\partial\Omega = 0}. \quad (\text{E.35})$$

### E.1.2.2 Équation de conservation de la quantité de mouvement en repère entraîné

L'écriture de l'équation de conservation de la quantité de mouvement demande plus de travail puisqu'il s'agit d'une équation vectorielle pour laquelle la dérivée temporelle doit être transformée au moyen de la formule (E.22). L'équation doit ensuite être projetée dans le repère entraîné en multipliant par  $\mathbf{R}^\top$ .

Dans un premier temps, la formule de dérivation en repère mobile appliquée au vecteur  $\mathbf{r}_A = \int \rho \mathbf{u}_A \, d\Omega$  dont les composantes sont évaluées dans la base  $\mathcal{B}_A$  donne

$$\left. \frac{\partial}{\partial t} \right|_{\mathcal{X}} \int_{\Omega(t)} \rho \mathbf{u}_A \, d\Omega = \left. \frac{\delta}{\delta t} \right|_{\mathcal{X}} \int_{\Omega(t)} \rho \mathbf{u}_A \, d\Omega + \boldsymbol{\omega}_A(\mathcal{R}_E/\mathcal{R}_A) \wedge \int_{\Omega(t)} \rho \mathbf{u}_A \, d\Omega. \quad (\text{E.36})$$

L'équation est ensuite projetée dans la base du repère entraîné  $\mathcal{B}_E$  en multipliant par  $\mathbf{R}^\top$ , puis le bilan de conservation de la quantité de mouvement donné à l'équation (4.27) est introduit dans la relation obtenue pour aboutir à l'expression

$$\begin{aligned} \mathbf{R}^\top \left. \frac{\delta}{\delta t} \right|_{\mathcal{X}} \int_{\Omega(t)} \rho \mathbf{u}_A \, d\Omega + \mathbf{R}^\top \boldsymbol{\omega}_A(\mathcal{R}_E/\mathcal{R}_A) \wedge \int_{\Omega(t)} \rho \mathbf{u}_A \, d\Omega = \\ -\mathbf{R}^\top \oint_{\partial\Omega(t)} [\rho \mathbf{u}_A \otimes (\mathbf{u}_A - \mathbf{s}_A)] \mathbf{n}_A \, d\partial\Omega + \mathbf{R}^\top \oint_{\partial\Omega(t)} \boldsymbol{\sigma}(\mathbf{u}_A) \mathbf{n}_A \, d\partial\Omega, \end{aligned} \quad (\text{E.37})$$

dans laquelle les différents termes doivent être modifiés pour faire apparaître les vitesses absolues  $\mathbf{u}_E$  exprimées dans le repère entraîné. Le premier terme est réécrit en explicitant la dérivée dans le repère entraîné  $\delta/\delta t$  définie à l'équation (E.18). Par ailleurs, l'orthogonalité de la matrice de changement de base  $\mathbf{R}$  donne

$$\mathbf{R}^\top \left. \frac{\delta}{\delta t} \right|_{\mathcal{X}} \int_{\Omega(t)} \rho \mathbf{u}_A \, d\Omega = \mathbf{R}^\top \left[ \left. \frac{\partial}{\partial t} \right|_{\mathcal{X}} \left( \mathbf{R}^\top \int_{\Omega(t)} \rho \mathbf{R} \mathbf{u}_E \, d\Omega \right) \right] = \left. \frac{\partial}{\partial t} \right|_{\mathcal{X}} \int_{\Omega(t)} \rho \mathbf{u}_E \, d\Omega. \quad (\text{E.38})$$

Pour le second terme de l'expression (E.37), il est nécessaire d'introduire une nouvelle matrice de vitesse de rotation  $\boldsymbol{\Omega}_E$  et un nouveau vecteur vitesse de rotation  $\boldsymbol{\omega}_E$  décrivant la rotation dans la base  $\mathcal{B}_E$ . En effet, le terme s'écrit avec la définition (E.12)

$$\mathbf{R}^\top \boldsymbol{\omega}_A(\mathcal{R}_E/\mathcal{R}_A) \wedge \int_{\Omega(t)} \rho \mathbf{u}_A \, d\Omega = \mathbf{R}^\top \boldsymbol{\Omega}_A \int_{\Omega(t)} \rho \mathbf{u}_A \, d\Omega = \mathbf{R}^\top \left[ \frac{d\mathbf{R}}{dt} \mathbf{R}^\top \right] \int_{\Omega(t)} \rho \mathbf{R} \mathbf{u}_E \, d\Omega. \quad (\text{E.39})$$

Le produit  $\mathbf{R}^\top \mathbf{R}$  se simplifie puisque la matrice est orthogonale, et on introduit alors la matrice de vitesse angulaire  $\boldsymbol{\Omega}_E$  ainsi que le vecteur vitesse angulaire  $\boldsymbol{\omega}_E$  définis pour tout  $\mathbf{r}_E \in \mathbb{R}^3$  par

$$\boldsymbol{\Omega}_E \mathbf{r}_E = \boldsymbol{\omega}_E \wedge \mathbf{r}_E \quad \text{avec} \quad \boldsymbol{\Omega}_E = \mathbf{R}^\top \frac{d\mathbf{R}}{dt}. \quad (\text{E.40})$$

La matrice de la vitesse angulaire exprimée dans la base  $\mathcal{B}_E$  aurait par ailleurs pu être simplement déduite de la relation de changement de base pour une matrice quelconque. En effet, la relation (A.11) donne  $\boldsymbol{\Omega}_E = \mathbf{R}^\top \boldsymbol{\Omega}_A \mathbf{R}$  puisque la matrice de changement de base  $\mathbf{R}$  est orthogonale et vérifie la propriété (E.7b). On retrouve alors l'expression (E.40) en développant :  $\boldsymbol{\Omega}_E = \mathbf{R}^\top \boldsymbol{\Omega}_A \mathbf{R} = \mathbf{R}^\top (\dot{\mathbf{R}} \mathbf{R}^\top) \mathbf{R} = \mathbf{R}^\top \dot{\mathbf{R}}$ . Avec ces nouvelles quantités, le second terme de (E.37) s'écrit donc

$$\mathbf{R}^\top \boldsymbol{\omega}_A(\mathcal{R}_E/\mathcal{R}_A) \wedge \int_{\Omega(t)} \rho \mathbf{u}_A d\Omega = \boldsymbol{\Omega}_E \int_{\Omega(t)} \rho \mathbf{u}_E d\Omega = \boldsymbol{\omega}_E(\mathcal{R}_E/\mathcal{R}_A) \wedge \int_{\Omega(t)} \rho \mathbf{u}_E d\Omega. \quad (\text{E.41})$$

Le troisième terme de l'expression (E.37) est facilement transformé à l'aide de la relation de changement de base (E.3) en

$$\begin{aligned} \mathbf{R}^\top \oint_{\partial\Omega(t)} [\rho \mathbf{u}_A \otimes (\mathbf{u}_A - \mathbf{s}_A)] \mathbf{n}_A d\partial\Omega &= \mathbf{R}^\top \oint_{\partial\Omega(t)} [\rho \mathbf{u}_A (\mathbf{u}_A - \mathbf{s}_A)^\top] \mathbf{n}_A d\partial\Omega \\ &= \mathbf{R}^\top \oint_{\partial\Omega(t)} [\rho \mathbf{R} \mathbf{u}_E (\mathbf{u}_E - \mathbf{s}_E)^\top \mathbf{R}^\top] \mathbf{R} \mathbf{n}_E d\partial\Omega \\ &= \oint_{\partial\Omega(t)} [\rho \mathbf{u}_E \otimes (\mathbf{u}_E - \mathbf{s}_E)] \mathbf{n}_E d\partial\Omega. \end{aligned} \quad (\text{E.42})$$

Reste finalement le terme diffusif à modifier. Il faut pour cela expliciter le tenseur des contraintes  $\boldsymbol{\sigma}(\mathbf{u}_A) = -p\mathbf{I} + \boldsymbol{\tau}(\mathbf{u}_A)$ . Le tenseur des contraintes visqueuses  $\boldsymbol{\tau}(\mathbf{u}_A)$  exprimé dans la base  $\mathcal{B}_E$  devient

$$\boldsymbol{\tau}(\mathbf{u}_A) = \lambda (\text{div}_A \mathbf{u}_A) \mathbf{I} + \mu (\nabla_A \mathbf{u}_A + \nabla_A^\top \mathbf{u}_A) = \lambda (\text{div}_E \mathbf{u}_E) \mathbf{R} \mathbf{R}^\top + \mu \mathbf{R} (\nabla_E \mathbf{u}_E + \nabla_E^\top \mathbf{u}_E) \mathbf{R}^\top \quad (\text{E.43})$$

en utilisant les relations (E.32) et (E.29) pour la transformation des opérateurs de divergence et du gradient. Si on définit par  $\boldsymbol{\tau}(\mathbf{u}_E) = \lambda (\text{div}_E \mathbf{u}_E) \mathbf{I} + \mu (\nabla_E \mathbf{u}_E + \nabla_E^\top \mathbf{u}_E)$  le tenseur des contraintes visqueuses exprimé dans la base  $\mathcal{B}_E$ , alors on peut écrire  $\boldsymbol{\tau}(\mathbf{u}_A) = \mathbf{R} \boldsymbol{\tau}(\mathbf{u}_E) \mathbf{R}^\top$  et le tenseur des contraintes  $\boldsymbol{\sigma}(\mathbf{u}_A)$  vérifie alors la relation

$$\boldsymbol{\sigma}(\mathbf{u}_A) = \mathbf{R} \boldsymbol{\sigma}(\mathbf{u}_E) \mathbf{R}^\top \quad (\text{E.44})$$

qui revient simplement à effectuer un changement de base pour la matrice  $\boldsymbol{\sigma}$  au moyen de la relation (A.11). En notant  $\boldsymbol{\sigma}(\mathbf{u}_E) = -p\mathbf{I} + \boldsymbol{\tau}(\mathbf{u}_E)$ , le quatrième terme de (E.37) se transforme donc en

$$\mathbf{R}^\top \oint_{\partial\Omega(t)} \boldsymbol{\sigma}(\mathbf{u}_A) \mathbf{n}_A d\partial\Omega = \mathbf{R}^\top \oint_{\partial\Omega(t)} [\mathbf{R} \boldsymbol{\sigma}(\mathbf{u}_E) \mathbf{R}^\top] \mathbf{R} \mathbf{n}_E d\partial\Omega = \oint_{\partial\Omega(t)} \boldsymbol{\sigma}(\mathbf{u}_E) \mathbf{n}_E d\partial\Omega. \quad (\text{E.45})$$

Finalement, le bilan de conservation de la quantité de mouvement exprimé dans le repère entraîné avec les composantes de la vitesse absolue prend la forme suivante en remplaçant les termes (E.38), (E.41), (E.42) et (E.45) dans (E.37) :

$$\boxed{\frac{\partial}{\partial t} \Big|_{\mathcal{X}} \int_{\Omega(t)} \rho \mathbf{u}_E d\Omega + \oint_{\partial\Omega(t)} [\rho \mathbf{u}_E \otimes (\mathbf{u}_E - \mathbf{s}_E)] \mathbf{n}_E d\partial\Omega = \oint_{\partial\Omega(t)} [-p\mathbf{I} + \boldsymbol{\tau}(\mathbf{u}_E)] \mathbf{n}_E d\partial\Omega - \boldsymbol{\omega}_E(\mathcal{R}_E/\mathcal{R}_A) \wedge \int_{\Omega(t)} \rho \mathbf{u}_E d\Omega} \quad (\text{E.46})$$

### E.1.2.3 Équation de conservation de l'énergie en repère entraîné

Pour terminer, il reste à établir le bilan de conservation de l'énergie totale dans le repère entraîné. Puisqu'il s'agit d'une équation scalaire, la dérivée temporelle dans le bilan de l'énergie (4.27) reste identique, et

il suffit de modifier les autres termes. Comme pour le bilan de conservation de la masse, le terme convectif sous l'intégrale surfacique s'écrit :  $[\rho e_T (\mathbf{u}_A - \mathbf{s}_A)^T] \mathbf{n}_A = [\rho e_T (\mathbf{u}_E - \mathbf{s}_E)^T \mathbf{R}^T] \mathbf{R} \mathbf{n}_E = [\rho e_T (\mathbf{u}_E - \mathbf{s}_E)^T] \mathbf{n}_E$ . Ensuite, le terme diffusif est transformé au moyen de la relation (E.44) pour le tenseur des contraintes et en notant que le flux de chaleur s'écrit  $\mathbf{q}_{\theta,A} = -k_\theta \nabla_A \theta$  on obtient

$$\begin{aligned} \oint_{\partial\Omega(t)} [\boldsymbol{\sigma}(\mathbf{u}_A) \mathbf{u}_A - \mathbf{q}_{\theta,A}]^T \mathbf{n}_A \, d\partial\Omega &= \oint_{\partial\Omega(t)} [(\mathbf{R} \boldsymbol{\sigma}(\mathbf{u}_E) \mathbf{R}^T) \mathbf{R} \mathbf{u}_E + k_\theta \mathbf{R} \nabla_E \theta]^T \mathbf{R} \mathbf{n}_E \, d\partial\Omega \\ &= \oint_{\partial\Omega(t)} [\boldsymbol{\sigma}(\mathbf{u}_E) \mathbf{u}_E + k_\theta \nabla_E \theta]^T \mathbf{n}_E \, d\partial\Omega : \end{aligned} \quad (\text{E.47})$$

Au final, le bilan de conservation de l'énergie s'écrit donc

$$\boxed{\left. \frac{\partial}{\partial t} \right|_{\mathbf{x}} \int_{\Omega(t)} \rho e_T \, d\Omega + \oint_{\partial\Omega(t)} [\rho e_T (\mathbf{u}_E - \mathbf{s}_E)^T] \mathbf{n}_E \, d\partial\Omega = \oint_{\partial\Omega(t)} [\boldsymbol{\sigma}(\mathbf{u}_E) \mathbf{u}_E - \mathbf{q}_{\theta,E}]^T \mathbf{n}_E \, d\partial\Omega} . \quad (\text{E.48})$$

Les expressions des bilans de conservation de la masse (E.35), de la quantité de mouvement (E.46) et de l'énergie (E.48) établis ici justifient donc la formulation FALE-REUE des équations de Navier-Stokes données à l'équation (4.31) au moyen d'une description ALE dans le repère entraîné  $\mathcal{R}_E$  avec les composantes de la vitesse absolue  $\mathbf{u}_E$ .

### E.1.3 Formulation intégrale des équations de Navier-Stokes en repère entraîné avec les composantes de la vitesse relative

La formulation FALE-REVE des équations (4.33) de Navier-Stokes dans le repère entraîné avec les composantes de la vitesse relative se déduit en partie des résultats précédents. L'obtention des équations requiert un travail fastidieux, en particulier pour l'équation de conservation de l'énergie. L'objectif est de parvenir à exprimer les équations en fonction de la vitesse relative exprimée dans le repère entraîné par  $\mathbf{v}_E = \mathbf{u}_E - \mathbf{s}_{e,E}$ , où  $\mathbf{s}_E = \mathbf{s}_{e,E} + \mathbf{s}_{d,E}$  est la vitesse des nœuds du maillage telle que  $\mathbf{s}_{e,E}$  est la vitesse d'entraînement correspondant au mouvement de corps rigide du domaine et  $\mathbf{s}_{d,E}$  est la vitesse de déformation des nœuds.

#### E.1.3.1 Équation de conservation de la masse en repère entraîné

Le bilan de conservation de la masse s'exprime trivialement en fonction de  $\mathbf{v}_E$  à partir de l'équation (E.35) en décomposant la vitesse du maillage en  $\mathbf{s}_E = \mathbf{s}_{e,E} + \mathbf{s}_{d,E}$ , ce qui conduit à

$$\boxed{\left. \frac{\partial}{\partial t} \right|_{\mathbf{x}} \int_{\Omega(t)} \rho \, d\Omega + \oint_{\partial\Omega(t)} \rho (\mathbf{v}_E - \mathbf{s}_{d,E})^T \mathbf{n}_E \, d\partial\Omega = 0} . \quad (\text{E.49})$$

Les termes convectifs apparaissant dans les équations de bilan de la quantité de mouvement et de l'énergie seront modifiés de façon similaire.

#### E.1.3.2 Équation de conservation de la quantité de mouvement en repère entraîné

L'écriture de l'équation de conservation de la quantité de mouvement avec la vitesse relative nécessite plus de travail. Tout d'abord, on exprime chacun des termes de l'équation (E.46) en fonction de la vitesse relative. Le premier terme donne

$$\left. \frac{\partial}{\partial t} \right|_{\mathbf{x}} \int_{\Omega(t)} \rho \mathbf{u}_E \, d\Omega = \left. \frac{\partial}{\partial t} \right|_{\mathbf{x}} \int_{\Omega(t)} \rho \mathbf{v}_E \, d\Omega + \left. \frac{\partial}{\partial t} \right|_{\mathbf{x}} \int_{\Omega(t)} \rho \mathbf{s}_{e,E} \, d\Omega. \quad (\text{E.50})$$

Le second terme de l'expression précédente donnera naissance à une partie de la force centrifuge définie à l'équation (4.34) comme on le verra par la suite.

Le terme de convection sous l'intégrale dans l'équation (E.46) se transforme en :  $\rho \mathbf{u}_E (\mathbf{u}_E - \mathbf{s}_E)^T =$

$\rho(\mathbf{v}_E + \mathbf{s}_{e,E})(\mathbf{v}_E + \mathbf{s}_{e,E} - \mathbf{s}_{d,E} - \mathbf{s}_{d,E})^\top = \rho \mathbf{v}_E (\mathbf{v}_E - \mathbf{s}_{d,E})^\top + \rho \mathbf{s}_{e,E} (\mathbf{v}_E - \mathbf{s}_{d,E})^\top$ . Par conséquent le second terme de l'équation (E.46) devient

$$\begin{aligned} \oint_{\partial\Omega(t)} \left[ \rho \mathbf{u}_E (\mathbf{u}_E - \mathbf{s}_E)^\top \right] \mathbf{n}_E d\partial\Omega &= \oint_{\partial\Omega(t)} \left[ \rho \mathbf{v}_E (\mathbf{v}_E - \mathbf{s}_{d,E})^\top \right] \mathbf{n}_E d\partial\Omega \\ &+ \oint_{\partial\Omega(t)} \left[ \rho \mathbf{s}_{e,E} (\mathbf{v}_E - \mathbf{s}_{d,E})^\top \right] \mathbf{n}_E d\partial\Omega. \end{aligned} \quad (\text{E.51})$$

Dans le terme diffusif de l'équation (E.46) les contraintes normales  $-p\mathbf{I}$  ne sont pas modifiées. En revanche, le tenseur des contraintes visqueuses  $\boldsymbol{\tau}(\mathbf{u}_E)$  doit être formulé en fonction de la vitesse relative. Pour y parvenir, les propriétés suivantes sont nécessaires

$$\nabla_E^\top \mathbf{s}_{e,E} = -\nabla_E \mathbf{s}_{e,E}, \quad (\text{E.52a})$$

$$\operatorname{div}_E \mathbf{s}_{e,E} = \mathbf{0}, \quad (\text{E.52b})$$

$$\nabla_E \mathbf{s}_{e,E} = \boldsymbol{\Omega}_E \Leftrightarrow \forall \mathbf{r} \in \mathbb{R}^3 \quad (\nabla_E \mathbf{s}_{e,E}) \mathbf{r} = \boldsymbol{\omega}_E \wedge \mathbf{r}. \quad (\text{E.52c})$$

La première propriété découle de la définition même d'un corps rigide donnée à l'équation (4.41). En effet, si  $\mathbf{s}_{e,E}$  est le champ de vitesse correspondant au mouvement de corps rigide, alors  $\mathbf{D}_E(\mathbf{s}_{e,E}) = 1/2(\nabla_E \mathbf{s}_{e,E} + \nabla_E^\top \mathbf{s}_{e,E}) = \mathbf{0}$ , d'où la première relation. Par ailleurs,  $\mathbf{D}(\mathbf{s}_{e,E}) : \mathbf{I} = \operatorname{div}_E \mathbf{s}_{e,E}$  et l'hypothèse de corps rigide impose l'annulation du tenseur des taux de déformation, d'où la seconde propriété. La dernière propriété provient du fait que  $\mathbf{s}_{e,E} = \mathbf{s}_{0,E} + \boldsymbol{\omega}_E \wedge \tilde{\mathbf{x}}_E$  et que  $\mathbf{s}_{0,E}$  est indépendant de la position d'un point  $M$  quelconque. Avec les propriétés précédentes, le tenseur des contraintes visqueuses s'écrit donc

$$\begin{aligned} \boldsymbol{\tau}(\mathbf{u}_E) &= \lambda (\operatorname{div}_E \mathbf{u}_E) \mathbf{I} + \mu (\nabla_E \mathbf{u}_E + \nabla_E^\top \mathbf{u}_E) \\ &= \lambda (\operatorname{div}_E \mathbf{v}_E + \underbrace{\operatorname{div}_E \mathbf{s}_{e,E}}_{=0}) \mathbf{I} + \mu (\nabla_E \mathbf{v}_E + \nabla_E \mathbf{s}_{e,E} + \nabla_E^\top \mathbf{v}_E + \underbrace{\nabla_E^\top \mathbf{s}_{e,E}}_{=-\nabla_E \mathbf{s}_{e,E}}) \\ &= \lambda (\operatorname{div}_E \mathbf{v}_E) \mathbf{I} + \mu (\nabla_E \mathbf{v}_E + \nabla_E^\top \mathbf{v}_E) \\ &= \boldsymbol{\tau}(\mathbf{v}_E). \end{aligned} \quad (\text{E.53})$$

Le tenseur des contraintes visqueuses reste ainsi invariant à cause des propriétés (E.52a) et (E.52b). Le terme diffusif se transforme donc en fonction de la vitesse relative  $\mathbf{v}_E$  en

$$\oint_{\partial\Omega(t)} \boldsymbol{\sigma}(\mathbf{u}_E) \mathbf{n}_E d\partial\Omega = \oint_{\partial\Omega(t)} \boldsymbol{\sigma}(\mathbf{v}_E) \mathbf{n}_E d\partial\Omega. \quad (\text{E.54})$$

Enfin, en décomposant le terme source de l'équation (E.46) avec la relation  $\mathbf{u}_E = \mathbf{v}_E + \mathbf{s}_{e,E}$  et en remplaçant les termes avec les équations (E.50), (E.51) et (E.54), le bilan s'écrit en fonction de la vitesse relative

$$\begin{aligned} \frac{\partial}{\partial t} \Big|_{\mathcal{X}} \int_{\Omega(t)} \rho \mathbf{v}_E d\Omega + \oint_{\partial\Omega(t)} \left[ \rho \mathbf{v}_E (\mathbf{v}_E - \mathbf{s}_{d,E})^\top \right] \mathbf{n}_E d\partial\Omega - \oint_{\partial\Omega(t)} \boldsymbol{\sigma}(\mathbf{v}_E) \mathbf{n}_E d\partial\Omega = \\ - \underbrace{\frac{\partial}{\partial t} \Big|_{\mathcal{X}} \int_{\Omega(t)} \rho \mathbf{s}_{e,E} d\Omega}_{\textcircled{1}} - \underbrace{\oint_{\partial\Omega(t)} \left[ \rho \mathbf{s}_{e,E} (\mathbf{v}_E - \mathbf{s}_{d,E})^\top \right] \mathbf{n}_E d\partial\Omega}_{\textcircled{2}} \\ - \underbrace{\boldsymbol{\omega}_E \wedge \int_{\Omega(t)} \rho \mathbf{v}_E d\Omega}_{\textcircled{3}} - \underbrace{\boldsymbol{\omega}_E \wedge \int_{\Omega(t)} \rho \mathbf{s}_{e,E} d\Omega}_{\textcircled{4}}. \end{aligned} \quad (\text{E.55})$$

Les différents termes du second membre sont alors réorganisés afin de faire apparaître les forces de Coriolis et centrifuge introduites à l'équation (4.34). Dans le terme  $\textcircled{1}$ , les symboles de dérivation et d'intégration doivent être intervertis. Toutefois, puisque le domaine d'intégration est variable au cours du temps, l'opération s'effectue au moyen du théorème de Reynolds écrit pour la description ALE à l'équation (4.10). Dans le cas présent, le changement du domaine est uniquement dû à la vitesse de déformation  $\mathbf{s}_{d,E}$  puisque les

équations sont formulées dans le repère entraîné. Par conséquent, la vitesse du maillage apparaissant dans l'équation (4.10) correspond seulement à  $\mathbf{s}_{d,E}$  et le terme ① s'écrit

$$\begin{aligned} \frac{\partial}{\partial t} \Big|_{\mathbf{x}} \int_{\Omega(t)} \rho \mathbf{s}_{e,E} \, d\Omega &= \int_{\Omega(t)} \frac{\partial \rho \mathbf{s}_{e,E}}{\partial t} \Big|_{\mathbf{x}} \, d\Omega + \oint_{\partial\Omega(t)} \rho \mathbf{s}_{e,E} \otimes \mathbf{s}_{d,E} \, d\Omega \\ &= \int_{\Omega(t)} \rho \frac{\partial \mathbf{s}_{e,E}}{\partial t} \Big|_{\mathbf{x}} \, d\Omega + \int_{\Omega(t)} \frac{\partial \rho}{\partial t} \Big|_{\mathbf{x}} \mathbf{s}_{e,E} \, d\Omega + \int_{\Omega(t)} \operatorname{div}_E (\rho \mathbf{s}_{e,E} \otimes \mathbf{s}_{d,E}) \, d\Omega. \end{aligned} \quad (\text{E.56})$$

Le premier terme dans l'équation précédente correspond à la première contribution dans l'accélération centrifuge (4.34) qui est donc conservé sous cette forme. Le second terme dans l'équation (E.56) est transformé en utilisant l'expression locale de la conservation de la masse en formulation eulérienne dans le repère relatif, de sorte que

$$\frac{\partial \rho}{\partial t} \Big|_{\mathbf{x}} \mathbf{s}_{e,E} = -(\mathbf{v}_E \cdot \nabla_E \rho + \rho \operatorname{div}_E \mathbf{v}_E) \mathbf{s}_{e,E}. \quad (\text{E.57})$$

Le dernier terme dans l'équation (E.56) est développé au moyen des relations (D.25e) et (D.25i) en

$$\begin{aligned} \operatorname{div}_E (\rho \mathbf{s}_{e,E} \otimes \mathbf{s}_{d,E}) &= \rho \operatorname{div}_E (\mathbf{s}_{e,E} \otimes \mathbf{s}_{d,E}) + (\mathbf{s}_{e,E} \otimes \mathbf{s}_{d,E}) \nabla_E \rho \\ &= \rho (\nabla_E \mathbf{s}_{e,E}) \mathbf{s}_{d,E} + \rho \mathbf{s}_{e,E} \operatorname{div}_E \mathbf{s}_{d,E} + (\mathbf{s}_{e,E} \otimes \mathbf{s}_{d,E}) \nabla_E \rho. \end{aligned} \quad (\text{E.58})$$

L'intégrale surfacique du terme ② est transformée avec le théorème de Green-Ostrogradski en une intégrale volumique afin de pouvoir ensuite combiner les termes avec ceux provenant de ①. L'intégrale sur le volume  $\Omega$  porte alors sur le terme

$$\begin{aligned} \operatorname{div}_E [\rho \mathbf{s}_{e,E} \otimes (\mathbf{v}_E - \mathbf{s}_{d,E})] &= \rho \mathbf{s}_{e,E} \operatorname{div}_E (\mathbf{v}_E - \mathbf{s}_{d,E}) + [\nabla_E (\rho \mathbf{s}_{e,E})] (\mathbf{v}_E - \mathbf{s}_{d,E}) \\ &= \rho \mathbf{s}_{e,E} \operatorname{div}_E \mathbf{v}_E - \rho \mathbf{s}_{e,E} \operatorname{div}_E \mathbf{s}_{d,E} \\ &\quad + \rho (\nabla_E \mathbf{s}_{e,E}) \mathbf{v}_E - \rho (\nabla_E \mathbf{s}_{e,E}) \mathbf{s}_{d,E} \\ &\quad + (\mathbf{s}_{e,E} \otimes \nabla_E \rho) \mathbf{v}_E - (\mathbf{s}_{e,E} \otimes \nabla_E \rho) \mathbf{s}_{d,E} \end{aligned} \quad (\text{E.59})$$

qui est transformé à l'aide des relations (D.25i) et (D.24c). Finalement, il reste à regrouper les termes de l'expression précédente avec ceux issus des équations (E.57) et (E.58). Le résultat se simplifie alors en

$$\begin{aligned} \textcircled{1} + \textcircled{2} &= -\cancel{[\mathbf{v}_E \cdot \nabla_E \rho] \mathbf{s}_{e,E}} - \cancel{\rho \mathbf{s}_{e,E} \operatorname{div}_E \mathbf{v}_E} \\ &\quad + \cancel{\rho (\nabla_E \mathbf{s}_{e,E}) \mathbf{s}_{d,E}} + \cancel{\rho \mathbf{s}_{e,E} \operatorname{div}_E \mathbf{s}_{d,E}} + \cancel{(\mathbf{s}_{e,E} \otimes \mathbf{s}_{d,E}) \nabla_E \rho} \\ &\quad + \cancel{\rho \mathbf{s}_{e,E} \operatorname{div}_E \mathbf{v}_E} - \cancel{\rho \mathbf{s}_{e,E} \operatorname{div}_E \mathbf{s}_{d,E}} \\ &\quad + \rho (\nabla_E \mathbf{s}_{e,E}) \mathbf{v}_E - \rho (\nabla_E \mathbf{s}_{e,E}) \mathbf{s}_{d,E} \\ &\quad + \cancel{(\mathbf{s}_{e,E} \otimes \nabla_E \rho) \mathbf{v}_E} - \cancel{(\mathbf{s}_{e,E} \otimes \nabla_E \rho) \mathbf{s}_{d,E}} \end{aligned} \quad (\text{E.60})$$

puisque la plupart des termes s'annulent en remarquant que  $(\mathbf{v}_E \cdot \nabla_E \rho) \mathbf{s}_{e,E}$  s'écrit en notation indicielle  $(v_E)_j \rho_{,j} (s_{e,E})_i$ , ce qui est équivalent à  $(\mathbf{s}_{e,E} \otimes \nabla_E \rho) \mathbf{v}_E$  dont le terme général est  $(s_{e,E})_i \rho_{,j} (v_E)_j$ . De la même façon,  $(\mathbf{s}_{e,E} \otimes \mathbf{s}_{d,E}) \nabla_E \rho$  et  $(\mathbf{s}_{e,E} \otimes \nabla_E \rho) \mathbf{s}_{d,E}$  sont tous deux des vecteurs de terme général  $(s_{e,E})_i (s_{d,E})_j \rho_{,j}$ . Enfin, le terme subsistant s'exprime en fonction du vecteur vitesse angulaire  $\boldsymbol{\omega}_E$  au moyen de la relation (E.52c) :  $\rho (\nabla_E \mathbf{s}_{e,E}) \mathbf{v}_E = \rho \boldsymbol{\omega}_E \wedge \mathbf{v}_E$ .

Ainsi la somme des termes ① à ④ devient

$$\textcircled{1} + \textcircled{2} + \textcircled{3} + \textcircled{4} = - \int_{\Omega(t)} \left[ \underbrace{\rho \frac{\partial \mathbf{s}_{e,E}}{\partial t} \Big|_{\mathbf{x}}}_{\mathbf{f}_{\text{cen}}} + \rho \boldsymbol{\omega}_E \wedge \mathbf{s}_{e,E} + \underbrace{2 \rho \boldsymbol{\omega}_E \wedge \mathbf{v}_E}_{\mathbf{f}_{\text{Cor}}} \right] d\Omega \quad (\text{E.61})$$

en utilisant les relations (E.56) à (E.60), et le bilan de conservation de la quantité de mouvement écrit avec la vitesse relative dans le repère entraîné au moyen d'une description ALE est donné par l'expression

$$\boxed{\begin{aligned} \frac{\partial}{\partial t} \Big|_{\mathbf{x}} \int_{\Omega(t)} \rho \mathbf{v}_E \, d\Omega + \oint_{\partial\Omega(t)} [\rho \mathbf{v}_E \otimes (\mathbf{v}_E - \mathbf{s}_{d,E})] \mathbf{n}_E \, d\partial\Omega = \\ \oint_{\partial\Omega(t)} \boldsymbol{\sigma}(\mathbf{v}_E) \mathbf{n}_E \, d\partial\Omega - \int_{\Omega(t)} \rho (\mathbf{f}_{\text{Cor}} + \mathbf{f}_{\text{cen}}) \, d\Omega \end{aligned}} \quad (\text{E.62})$$

Les termes sources liés à l'accélération de Coriolis et à l'accélération centrifuge sont définis aux équations (E.61) et (4.34). Le terme de *force centrifuge* est un abus de langage et il serait plus approprié de parler de force d'inertie d'entraînement (voir la remarque E.2 page 247).

### E.1.3.3 Équation de conservation de l'énergie en repère entraîné

L'énergie totale reste invariante lors d'un changement de repère puisque c'est une quantité scalaire. En effet,  $e_T = e_I + e_C = c_v \theta + 1/2 \|\mathbf{u}\|^2$  et le vecteur  $\mathbf{u}$  a la même norme dans  $\mathcal{B}_A$  et  $\mathcal{B}_E$  puisque le changement de base est une simple rotation<sup>1</sup>. Il n'y a donc pas de problème de définition de l'énergie cinétique lorsque la vitesse *absolue* est employée pour formuler le bilan de conservation de l'énergie comme on l'a fait pour l'équation (E.48).

Cependant, lorsque la vitesse *relative* est adoptée pour formuler les bilans de conservation, il est préférable de travailler avec une énergie totale « relative », définie à partir des composantes de la vitesse relative. Celle-ci découle naturellement de l'égalité  $\mathbf{u}^T \mathbf{u} = \mathbf{v}^T \mathbf{v} + 2 \mathbf{v}^T \mathbf{s}_e + \mathbf{s}_e^T \mathbf{s}_e$  qui permet de définir l'énergie cinétique « relative »  $e_{C,r}$ , l'énergie cinétique d'entraînement  $e_{C,e}$  et l'énergie cinétique fictive  $e_{C,f}$ , toutes trois données par

$$e_{C,r} = \frac{\|\mathbf{v}\|^2}{2}, \quad (\text{E.63a})$$

$$e_{C,e} = \frac{\|\mathbf{s}_{e,E}\|^2}{2}, \quad (\text{E.63b})$$

$$e_{C,f} = \mathbf{v}^T \mathbf{s}_e = \mathbf{s}_e^T \mathbf{v}, \quad (\text{E.63c})$$

si bien que l'énergie cinétique « absolue » est définie par

$$e_C = e_{C,r} + e_{C,f} + e_{C,e}. \quad (\text{E.64})$$

Pour obtenir une description cohérente du système dans le repère entraîné avec les composantes de la vitesse relative, il est nécessaire de travailler avec l'énergie totale « relative »  $e_{T,r}$  définie comme la somme de l'énergie interne et de l'énergie cinétique relative introduite à l'équation (E.63a). Les énergies totales « absolue »  $e_T$  et « relative »  $e_{T,r}$  sont donc reliées par les relations

$$e_T = e_I + e_C = e_{T,r} + e_{C,f} + e_{C,e} \quad \text{avec} \quad e_{T,r} = e_I + e_{C,r}, \quad (\text{E.65})$$

et l'équation de conservation de l'énergie totale « absolue » est donc transformée afin d'obtenir un bilan pour l'énergie totale relative. On procède terme par terme en décomposant le bilan (E.48) en

$$\underbrace{\frac{\partial}{\partial t} \Big|_{\mathbf{x}} \int_{\Omega(t)} \rho e_T \, d\Omega}_{\textcircled{1}} + \underbrace{\oint_{\partial\Omega(t)} [\rho e_T (\mathbf{u}_E - \mathbf{s}_E)^T] \mathbf{n}_E \, d\partial\Omega}_{\textcircled{2}} = \underbrace{\oint_{\partial\Omega(t)} [\boldsymbol{\sigma}(\mathbf{u}_E) \mathbf{u}_E - \mathbf{q}_{\theta,E}]^T \mathbf{n}_E \, d\partial\Omega}_{\textcircled{3}}. \quad (\text{E.66})$$

**Transformation du terme ①** À partir de la relation (E.65), le terme ① est scindé en trois contributions telles que

$$\frac{\partial}{\partial t} \Big|_{\mathbf{x}} \int_{\Omega(t)} \rho e_T \, d\Omega = \underbrace{\frac{\partial}{\partial t} \Big|_{\mathbf{x}} \int_{\Omega(t)} \rho e_{T,r} \, d\Omega}_{\textcircled{1.1}} + \underbrace{\frac{\partial}{\partial t} \Big|_{\mathbf{x}} \int_{\Omega(t)} \rho e_{C,f} \, d\Omega}_{\textcircled{1.2}} + \underbrace{\frac{\partial}{\partial t} \Big|_{\mathbf{x}} \int_{\Omega(t)} \rho e_{C,e} \, d\Omega}_{\textcircled{1.3}}. \quad (\text{E.67})$$

Le terme ①.① correspond à la variation de l'énergie totale relative qui est donc conservé sous cette forme.

Le terme ①.② est quant à lui transformé en

$$\frac{\partial}{\partial t} \Big|_{\mathbf{x}} \int_{\Omega(t)} \rho e_{C,f} \, d\Omega = \int_{\Omega(t)} \left( \underbrace{\frac{\partial \rho e_{C,f}}{\partial t} \Big|_{\mathbf{x}}}_{\textcircled{1.2.1}} + \underbrace{\text{div}_E [\rho e_{C,f} \mathbf{s}_{d,E}]}_{\textcircled{1.2.2}} \right) d\Omega \quad (\text{E.68})$$

<sup>1</sup>En effet :  $\|\mathbf{u}\|^2 = \mathbf{u}_A^T \mathbf{u}_A = (\mathbf{R} \mathbf{u}_E)^T (\mathbf{R} \mathbf{u}_E) = \mathbf{u}_E^T \mathbf{u}_E$ .

au moyen du théorème de Reynolds et en suivant le même procédé que pour établir l'équation (E.56). Le second terme ①.②.② est conservé pour l'instant sous cette forme. En revanche, le terme ①.②.① est développé à l'aide de l'expression de l'énergie cinétique fictive donnée à l'équation (E.63c) en

$$\left. \frac{\partial \rho e_{C,f}}{\partial t} \right|_{\mathbf{x}} = \rho \mathbf{s}_{e,E}^{\top} \left. \frac{\partial \mathbf{v}_E}{\partial t} \right|_{\mathbf{x}} + \left. \frac{\partial \rho}{\partial t} \right|_{\mathbf{x}} \mathbf{s}_{e,E}^{\top} \mathbf{v}_E + \rho \left. \frac{\partial \mathbf{s}_{e,E}^{\top}}{\partial t} \right|_{\mathbf{x}} \mathbf{v}_E. \quad (\text{E.69})$$

Les deux premiers termes du membre de droite peuvent encore être développés à l'aide des bilans locaux de conservation de la masse et de la quantité de mouvement. Le dernier terme est simplement modifié en intervertissant l'opération de transposition. Après calcul, l'expression (E.69) du terme ①.②.① se récrit

$$\begin{aligned} \left. \frac{\partial \rho e_{C,f}}{\partial t} \right|_{\mathbf{x}} &= \mathbf{s}_{e,E}^{\top} \left[ -\rho \mathbf{v}_E \cdot \nabla_E \mathbf{v}_E + \text{div}_E \boldsymbol{\sigma}(\mathbf{v}_E) + \rho (\mathbf{f}_{\text{cor}} + \mathbf{f}_{\text{cen}}) \right] \\ &- \left[ \mathbf{v}_E \cdot \nabla_E \rho + \rho \text{div}_E \mathbf{v}_E \right] \mathbf{s}_{e,E}^{\top} \mathbf{v}_E + \rho \mathbf{v}_E^{\top} \left. \frac{\partial \mathbf{s}_{e,E}}{\partial t} \right|_{\mathbf{x}}. \end{aligned} \quad (\text{E.70})$$

Comme pour le terme ①.②, le terme ①.③ est transformé avec le théorème de Reynolds et devient

$$\left. \frac{\partial}{\partial t} \right|_{\mathbf{x}} \int_{\Omega(t)} \rho e_{C,e} d\Omega = \int_{\Omega(t)} \left( \underbrace{\left. \frac{\partial \rho e_{C,e}}{\partial t} \right|_{\mathbf{x}}}_{\text{①.③.①}} + \underbrace{\text{div}_E [\rho e_{C,e} \mathbf{s}_{d,E}]}_{\text{①.③.②}} \right) d\Omega, \quad (\text{E.71})$$

puis la dérivée dans le terme ①.③.① est développée en plusieurs contributions à l'aide de la définition (E.63b) de l'énergie cinétique d'entraînement pour aboutir à l'expression

$$\left. \frac{\partial \rho e_{C,e}}{\partial t} \right|_{\mathbf{x}} = \frac{1}{2} \left[ 2\rho \mathbf{s}_{e,E}^{\top} \left. \frac{\partial \mathbf{s}_{e,E}}{\partial t} \right|_{\mathbf{x}} + \left. \frac{\partial \rho}{\partial t} \right|_{\mathbf{x}} \mathbf{s}_{e,E}^{\top} \mathbf{s}_{e,E} \right]. \quad (\text{E.72})$$

Le premier terme dans l'expression précédente est conservé sous cette forme tandis que le second est modifié à l'aide de l'équation de conservation de la masse. Le terme ①.③.② reste inchangé et par conséquent après calcul le terme ①.③ s'écrit

$$\begin{aligned} \left. \frac{\partial}{\partial t} \right|_{\mathbf{x}} \int_{\Omega(t)} \rho e_{C,e} d\Omega &= \int_{\Omega(t)} \rho \mathbf{s}_{e,E}^{\top} \left. \frac{\partial \mathbf{s}_{e,E}}{\partial t} \right|_{\mathbf{x}} d\Omega - \frac{1}{2} \int_{\Omega(t)} (\mathbf{v}_E \cdot \nabla_E \rho + \rho \text{div}_E \mathbf{v}_E) (\mathbf{s}_{e,E}^{\top} \mathbf{s}_{e,E}) d\Omega \\ &+ \int_{\Omega(t)} \text{div}_E (\rho e_{C,e} \mathbf{s}_{d,E}) d\Omega. \end{aligned} \quad (\text{E.73})$$

En regroupant les expressions (E.68), (E.70) et (E.73) le terme ① s'écrit donc

$$\begin{aligned} \left. \frac{\partial}{\partial t} \right|_{\mathbf{x}} \int_{\Omega(t)} \rho e_T d\Omega &= \left. \frac{\partial}{\partial t} \right|_{\mathbf{x}} \int_{\Omega(t)} \rho e_{T,r} d\Omega + \int_{\Omega(t)} \text{div}_E (\rho e_{C,f} \mathbf{s}_{d,E}) d\Omega \\ &+ \int_{\Omega(t)} \mathbf{s}_{e,E}^{\top} \left[ -\rho \mathbf{v}_E \cdot \nabla_E \mathbf{v}_E + \text{div}_E \boldsymbol{\sigma}(\mathbf{v}_E) + \rho (\mathbf{f}_{\text{cor}} + \mathbf{f}_{\text{cen}}) \right] d\Omega \\ &- \int_{\Omega(t)} (\mathbf{v}_E \cdot \nabla_E \rho + \rho \text{div}_E \mathbf{v}_E) \mathbf{s}_{e,E}^{\top} \mathbf{v}_E d\Omega + \int_{\Omega(t)} \rho \mathbf{v}_E^{\top} \left. \frac{\partial \mathbf{s}_{e,E}}{\partial t} \right|_{\mathbf{x}} d\Omega \\ &+ \int_{\Omega(t)} \rho \mathbf{s}_{e,E}^{\top} \left. \frac{\partial \mathbf{s}_{e,E}}{\partial t} \right|_{\mathbf{x}} d\Omega + \int_{\Omega(t)} \text{div}_E (\rho e_{C,e} \mathbf{s}_{d,E}) d\Omega \\ &- \int_{\Omega(t)} \frac{1}{2} (\mathbf{v}_E \cdot \nabla_E \rho + \rho \text{div}_E \mathbf{v}_E) (\mathbf{s}_{e,E}^{\top} \mathbf{s}_{e,E}) d\Omega. \end{aligned} \quad (\text{E.74})$$

**Transformation du terme ②** Dans le second terme ② défini à l'équation (E.66), on modifie l'énergie totale  $e_T$  avec la relation (E.65) et on remplace la vitesse absolue  $\mathbf{u}_E$  par  $\mathbf{v}_E + \mathbf{s}_{e,E}$ . Après avoir transformé l'intégrale surfacique en une intégrale volumique avec le théorème de Green-Ostrogradski, on aboutit à



$$\begin{aligned}
 \int_{\Omega(t)} \operatorname{div}_E [\rho e_T (\mathbf{u}_E - \mathbf{s}_E)] \, d\Omega &= \underbrace{\int_{\Omega(t)} \operatorname{div}_E [\rho e_{T,r} (\mathbf{v}_E - \mathbf{s}_{d,E})] \, d\Omega}_{\textcircled{2.1}} \\
 &+ \underbrace{\int_{\Omega(t)} \operatorname{div}_E [\rho e_{C,f} (\mathbf{v}_E - \mathbf{s}_{d,E})] \, d\Omega}_{\textcircled{2.2}} \\
 &+ \underbrace{\int_{\Omega(t)} \operatorname{div}_E [\rho e_{C,e} (\mathbf{v}_E - \mathbf{s}_{d,E})] \, d\Omega}_{\textcircled{2.3}}.
 \end{aligned} \tag{E.75}$$

Le terme  $\textcircled{2.1}$  reste sous la forme précédente puisqu'il ne fait intervenir que l'énergie totale relative et la vitesse relative. Le terme  $\textcircled{2.2}$  est partiellement développé en remplaçant  $e_{C,f}$  par son expression (E.63c). Le terme sous l'intégrale se réécrit alors au moyen des relations (D.25d) et (E.63c) sous la forme

$$\begin{aligned}
 \operatorname{div}_E [\rho e_{C,f} (\mathbf{v}_E - \mathbf{s}_{d,E})] &= \operatorname{div}_E (\rho e_{C,f} \mathbf{v}_E) - \operatorname{div}_E (\rho e_{C,f} \mathbf{s}_{d,E}) \\
 &= \rho (\mathbf{v}_E^\top \mathbf{s}_{e,E}) \operatorname{div}_E \mathbf{v}_E + (\mathbf{v}_E \cdot \nabla_E \rho) (\mathbf{v}_E^\top \mathbf{s}_{e,E}) \\
 &+ \rho \mathbf{s}_{e,E}^\top (\nabla_E \mathbf{v}_E) \mathbf{v}_E + \rho \mathbf{v}_E^\top (\nabla_E \mathbf{s}_{e,E}) \mathbf{v}_E \\
 &- \operatorname{div}_E (\rho e_{C,f} \mathbf{s}_{d,E}).
 \end{aligned} \tag{E.76}$$

De la même façon le terme sous l'intégrale de  $\textcircled{2.3}$  se réécrit

$$\begin{aligned}
 \operatorname{div}_E [\rho e_{C,e} (\mathbf{v}_E - \mathbf{s}_{d,E})] &= \operatorname{div}_E (\rho e_{C,e} \mathbf{v}_E) - \operatorname{div}_E (\rho e_{C,e} \mathbf{s}_{d,E}) \\
 &= \frac{1}{2} \rho (\mathbf{s}_{e,E}^\top \mathbf{s}_{e,E}) \operatorname{div}_E \mathbf{v}_E + \frac{1}{2} (\mathbf{v}_E \cdot \nabla_E \rho) (\mathbf{s}_{e,E}^\top \mathbf{s}_{e,E}) \\
 &+ \rho \mathbf{s}_{e,E}^\top (\nabla_E \mathbf{s}_{e,E}) \mathbf{v}_E - \operatorname{div}_E (\rho e_{C,e} \mathbf{s}_{d,E})
 \end{aligned} \tag{E.77}$$

en remplaçant  $e_{C,e}$  par l'expression (E.63b). Enfin, en rassemblant les expressions (E.75), (E.76) et (E.77), le terme  $\textcircled{2}$  s'écrit

$$\begin{aligned}
 \int_{\Omega(t)} \operatorname{div}_E [\rho e_T (\mathbf{u}_E - \mathbf{s}_E)] \, d\Omega &= \int_{\Omega(t)} \operatorname{div}_E [\rho e_{T,r} (\mathbf{v}_E - \mathbf{s}_{d,E})] \, d\Omega - \int_{\Omega(t)} \operatorname{div}_E (\rho e_{C,f} \mathbf{s}_{d,E}) \, d\Omega \\
 &+ \int_{\Omega(t)} [\rho (\mathbf{v}_E^\top \mathbf{s}_{e,E}) \operatorname{div}_E \mathbf{v}_E + (\mathbf{v}_E \cdot \nabla_E \rho) (\mathbf{v}_E^\top \mathbf{s}_{e,E})] \, d\Omega \\
 &+ \int_{\Omega(t)} [\rho \mathbf{s}_{e,E}^\top (\nabla_E \mathbf{v}_E) \mathbf{v}_E + \rho \mathbf{v}_E^\top (\nabla_E \mathbf{s}_{e,E}) \mathbf{v}_E] \, d\Omega \\
 &+ \int_{\Omega(t)} \left[ \frac{1}{2} \rho (\mathbf{s}_{e,E}^\top \mathbf{s}_{e,E}) \operatorname{div}_E \mathbf{v}_E + \frac{1}{2} (\mathbf{v}_E \cdot \nabla_E \rho) (\mathbf{s}_{e,E}^\top \mathbf{s}_{e,E}) \right] \, d\Omega \\
 &+ \int_{\Omega(t)} [\rho \mathbf{s}_{e,E}^\top (\nabla_E \mathbf{s}_{e,E}) \mathbf{v}_E - \operatorname{div}_E (\rho e_{C,e} \mathbf{s}_{d,E})] \, d\Omega.
 \end{aligned} \tag{E.78}$$

**Transformation du terme  $\textcircled{3}$**  Le terme  $\textcircled{3}$  relatif aux contributions diffusives est très peu modifié. Le tenseur des contraintes visqueuses n'est pas affecté par le passage de la vitesse absolue à la vitesse relative; en effet, on a montré à l'équation (E.53) que  $\boldsymbol{\tau}(\mathbf{u}_E) = \boldsymbol{\tau}(\mathbf{v}_E)$ . Par ailleurs le terme de flux de chaleur est aussi invariant puisqu'il est indépendant de la vitesse de l'écoulement. Ainsi, il suffit de remplacer  $\mathbf{u}_E$  par  $\mathbf{v}_E + \mathbf{s}_{e,E}$  dans le terme  $\textcircled{3}$ , ce qui conduit après transformation avec le théorème de Green-Ostrogradski à

$$\int_{\Omega(t)} \operatorname{div}_E [\boldsymbol{\sigma}(\mathbf{u}_E) \mathbf{u}_E - \mathbf{q}_{\theta,E}] \, d\Omega = \int_{\Omega(t)} \operatorname{div}_E [\boldsymbol{\sigma}(\mathbf{v}_E) \mathbf{v}_E - \mathbf{q}_{\theta,E}] \, d\Omega + \int_{\Omega(t)} \operatorname{div}_E [\boldsymbol{\sigma}(\mathbf{v}_E) \mathbf{s}_{e,E}] \, d\Omega. \tag{E.79}$$

**Simplification des termes** Le remplacement des termes ①, ② et ③ par les expressions (E.74), (E.78) et (E.79) dans l'équation (E.66) conduit à

$$\begin{aligned} \frac{\partial}{\partial t} \Big|_{\mathbf{x}} \int_{\Omega(t)} \rho e_{T,r} d\Omega + \oint_{\partial\Omega(t)} [\rho e_{T,r} (\mathbf{v}_E - \mathbf{s}_{d,E})^\top] \mathbf{n}_E d\partial\Omega = \\ \oint_{\partial\Omega(t)} [\boldsymbol{\sigma}(\mathbf{v}_E) \mathbf{v}_E - \mathbf{q}_{\theta,E}]^\top \mathbf{n}_E d\partial\Omega + \int_{\Omega(t)} \mathcal{I} d\Omega, \end{aligned} \quad (\text{E.80})$$

où le terme  $\mathcal{I}$  est donné par

$$\begin{aligned} \mathcal{I} = & -\text{div}_E [\rho e_{C,f} \mathbf{s}_{d,E}] + \rho \mathbf{s}_{e,E}^\top (\mathbf{v}_E \cdot \nabla_E \mathbf{v}_E) - \mathbf{s}_{e,E}^\top \text{div}_E \boldsymbol{\sigma}(\mathbf{v}_E) - \rho \mathbf{s}_{e,E}^\top (\mathbf{f}_{\text{Cor}} + \mathbf{f}_{\text{cen}}) \\ & + (\mathbf{v}_E \cdot \nabla_E \rho) (\mathbf{s}_{e,E}^\top \mathbf{v}_E) + \rho \text{div}_E \mathbf{v}_E (\mathbf{s}_{e,E}^\top \mathbf{v}_E) - \rho \mathbf{v}_E^\top \frac{\partial \mathbf{s}_{e,E}}{\partial t} \Big|_{\mathbf{x}} - \rho \mathbf{s}_{e,E}^\top \frac{\partial \mathbf{s}_{e,E}}{\partial t} \Big|_{\mathbf{x}} \\ & - \text{div}_E [\rho e_{C,e} \mathbf{s}_{d,E}] + \frac{1}{2} (\mathbf{v}_E \cdot \nabla_E \rho) (\mathbf{s}_{e,E}^\top \mathbf{s}_{e,E}) + \frac{1}{2} (\rho \text{div}_E \mathbf{v}_E) (\mathbf{s}_{e,E}^\top \mathbf{s}_{e,E}) \\ & + \text{div}_E [\rho e_{C,f} \mathbf{s}_{d,E}] - \rho (\mathbf{v}_E \mathbf{s}_{e,E}^\top) \text{div}_E \mathbf{v}_E - (\mathbf{v}_E \cdot \nabla_E \rho) (\mathbf{v}_E^\top \mathbf{s}_{e,E}) \\ & - \rho \mathbf{s}_{e,E}^\top (\nabla_E \mathbf{v}_E) \mathbf{v}_E - \rho \mathbf{v}_E^\top (\nabla_E \mathbf{s}_{e,E}) \mathbf{v}_E \\ & - \frac{1}{2} \rho (\mathbf{s}_{e,E}^\top \mathbf{s}_{e,E}) \text{div}_E \mathbf{v}_E - \frac{1}{2} (\mathbf{v}_E \cdot \nabla_E \rho) (\mathbf{s}_{e,E}^\top \mathbf{s}_{e,E}) \\ & - \rho \mathbf{s}_{e,E}^\top (\nabla_E \mathbf{s}_{e,E}) \mathbf{v}_E + \text{div}_E [\rho e_{C,e} \mathbf{s}_{d,E}] + \text{div}_E [\boldsymbol{\sigma}(\mathbf{v}_E) \mathbf{s}_{e,E}]. \end{aligned} \quad (\text{E.81})$$

Une grande partie des termes s'éliminent immédiatement, et les termes  $\rho \mathbf{s}_{e,E}^\top (\mathbf{v}_E \cdot \nabla_E \mathbf{v}_E)$  et  $\rho \mathbf{s}_{e,E}^\top (\nabla_E \mathbf{v}_E) \mathbf{v}_E$  se simplifient aussi puisque tous deux ont pour terme général  $\rho (s_e)_i (v_j v_{i,j})$  et  $\rho (s_e)_i (v_{i,j}) v_j$  respectivement. Par conséquent, le terme  $\mathcal{I}$  s'écrit

$$\begin{aligned} \mathcal{I} = & -\mathbf{s}_{e,E}^\top \text{div}_E \boldsymbol{\sigma}(\mathbf{v}_E) + \text{div}_E [\boldsymbol{\sigma}(\mathbf{v}_E) \mathbf{s}_{e,E}] \\ & - \rho \mathbf{s}_{e,E}^\top (\mathbf{f}_{\text{Cor}} + \mathbf{f}_{\text{cen}}) - \rho \mathbf{v}_E^\top \frac{\partial \mathbf{s}_{e,E}}{\partial t} \Big|_{\mathbf{x}} - \rho \mathbf{s}_{e,E}^\top \frac{\partial \mathbf{s}_{e,E}}{\partial t} \Big|_{\mathbf{x}} \\ & - \rho \mathbf{v}_E^\top (\nabla_E \mathbf{s}_{e,E}) \mathbf{v}_E - \rho \mathbf{s}_{e,E}^\top (\nabla_E \mathbf{s}_{e,E}) \mathbf{v}_E. \end{aligned} \quad (\text{E.82})$$

Dans la première ligne de l'équation précédente, le second terme est développé au moyen de la formule (D.25f) en utilisant le fait que le tenseur des contraintes est symétrique — donc  $\boldsymbol{\sigma}(\mathbf{v}_E) \mathbf{s}_{e,E} = \mathbf{s}_{e,E}^\top \boldsymbol{\sigma}(\mathbf{v}_E)$ . Ainsi la première ligne devient

$$-\mathbf{s}_{e,E}^\top \text{div}_E \boldsymbol{\sigma}(\mathbf{v}_E) + \text{div}_E [\boldsymbol{\sigma}(\mathbf{v}_E) \mathbf{s}_{e,E}] = -\cancel{\mathbf{s}_{e,E}^\top \text{div}_E \boldsymbol{\sigma}(\mathbf{v}_E)} + \cancel{\mathbf{s}_{e,E}^\top \text{div}_E \boldsymbol{\sigma}(\mathbf{v}_E)} + \nabla_E \mathbf{s}_{e,E} : \boldsymbol{\sigma}(\mathbf{v}_E), \quad (\text{E.83})$$

et le terme subsistant  $\nabla_E \mathbf{s}_{e,E} : \boldsymbol{\sigma}(\mathbf{v}_E)$  est en réalité nul. En effet, d'après la propriété (E.52a), le tenseur  $\nabla_E \mathbf{s}_{e,E}$  est antisymétrique. Par ailleurs, le tenseur des contraintes visqueuses est symétrique. Or le double produit de contraction de deux tenseurs  $\mathbf{A}$  et  $\mathbf{B}$  de terme général  $a_{ij}$  et  $b_{ij}$  est égal à celui des tenseurs transposés :  $\mathbf{A} : \mathbf{B} = \mathbf{A}^\top : \mathbf{B}^\top$ , donc  $(\mathbf{A} : \mathbf{B}) + (\mathbf{A}^\top : \mathbf{B}^\top) = 2(\mathbf{A} : \mathbf{B})$ ; les propriétés de symétrie de  $\mathbf{B}$  et d'antisymétrie de  $\mathbf{A}$  conduisent à  $(\mathbf{A} : \mathbf{B}) + (\mathbf{A}^\top : \mathbf{B}^\top) = a_{ij}b_{ij} + a_{ji}b_{ji} = a_{ij}b_{ij} - a_{ij}b_{ij} = 0$  et puisque  $(\mathbf{A} : \mathbf{B}) + (\mathbf{A}^\top : \mathbf{B}^\top) = 2(\mathbf{A} : \mathbf{B})$  alors on a bien  $\mathbf{A} : \mathbf{B} = 0$ . Il n'y a donc pas d'énergie supplémentaire créée lorsque les équations sont formulées avec les composantes de la vitesse relative.

Cependant, une partie des termes liés aux forces fictives subsiste. Le premier terme de la deuxième ligne de l'équation (E.82) est développé avec les expressions (4.34) de la force de Coriolis et de la force centrifuge, et ceux de la troisième ligne sont modifiés en utilisant la propriété (E.52c) ce qui donne

$$\begin{aligned} \mathcal{I} = & -\rho \mathbf{s}_{e,E}^\top \left( -2\boldsymbol{\omega}_E \wedge \mathbf{v}_E - \frac{\partial \mathbf{s}_{e,E}}{\partial t} - \boldsymbol{\omega}_E \wedge \mathbf{s}_{e,E} \right) - \rho \mathbf{v}_E^\top \frac{\partial \mathbf{s}_{e,E}}{\partial t} \Big|_{\mathbf{x}} - \rho \mathbf{s}_{e,E}^\top \frac{\partial \mathbf{s}_{e,E}}{\partial t} \Big|_{\mathbf{x}} \\ & - \rho \mathbf{v}_E^\top (\boldsymbol{\omega}_E \wedge \mathbf{v}_E) - \rho \mathbf{s}_{e,E}^\top (\boldsymbol{\omega}_E \wedge \mathbf{v}_E) \\ = & \rho \mathbf{s}_{e,E}^\top (\boldsymbol{\omega}_E \wedge \mathbf{v}_E) - \rho \mathbf{v}_E^\top (\boldsymbol{\omega}_E \wedge \mathbf{v}_E) + \rho \mathbf{s}_{e,E}^\top (\boldsymbol{\omega}_E \wedge \mathbf{s}_{e,E}) - \rho \mathbf{v}_E^\top \frac{\partial \mathbf{s}_{e,E}}{\partial t} \Big|_{\mathbf{x}} \\ = & \rho [\boldsymbol{\omega}_E, \mathbf{v}_E, \mathbf{s}_{e,E}] - \rho [\boldsymbol{\omega}_E, \mathbf{v}_E, \mathbf{v}_E] + \rho [\boldsymbol{\omega}_E, \mathbf{s}_{e,E}, \mathbf{s}_{e,E}] - \rho \mathbf{v}_E^\top \frac{\partial \mathbf{s}_{e,E}}{\partial t} \Big|_{\mathbf{x}}. \end{aligned} \quad (\text{E.84})$$

Dans la dernière égalité, on a simplement introduit les produits mixtes tels que  $\mathbf{w}^\top(\mathbf{u} \wedge \mathbf{v}) = (\mathbf{u} \wedge \mathbf{v}) \cdot \mathbf{w} = [\mathbf{u}, \mathbf{v}, \mathbf{w}]$ . Le produit mixte est nul si il contient au moins deux vecteurs colinéaires, donc les deux derniers produits mixtes dans l'expression précédente s'annulent. Enfin, le produit mixte est invariant par toute permutation circulaire des vecteurs qui le composent, et est égal mais de signe opposé si une permutation non circulaire de ses éléments est effectuée. On a alors  $[\boldsymbol{\omega}_E, \mathbf{v}_E, \mathbf{s}_{e,E}] = -[\boldsymbol{\omega}_E, \mathbf{s}_{e,E}, \mathbf{v}_E] = -(\boldsymbol{\omega}_E \wedge \mathbf{s}_{e,E}) \cdot \mathbf{v}_E$ . Par conséquent le terme  $\mathcal{I}$  se réduit finalement à

$$\mathcal{I} = -\rho \mathbf{v}_E^\top (\boldsymbol{\omega}_E \wedge \mathbf{s}_{e,E}) - \rho \mathbf{v}_E^\top \left. \frac{\partial \mathbf{s}_{e,E}}{\partial t} \right|_{\mathbf{x}} = \rho \mathbf{v}_E^\top \mathbf{f}_{\text{cen}}, \quad (\text{E.85})$$

et on retrouve bien le bilan de conservation de l'énergie donné à l'équation (4.33) en remplaçant l'expression précédente de  $\mathcal{I}$  dans l'équation (E.80) :

$$\boxed{\frac{\partial}{\partial t} \Big|_{\mathbf{x}} \int_{\Omega(t)} \rho e_{T,r} \, d\Omega + \oint_{\partial\Omega(t)} \rho e_{T,r} (\mathbf{v}_E - \mathbf{s}_{d,E}) \cdot \mathbf{n}_E \, d\partial\Omega = \oint_{\partial\Omega(t)} [\boldsymbol{\sigma}(\mathbf{v}_E) \mathbf{v}_E - \mathbf{q}_{\theta,E}] \cdot \mathbf{n}_E \, d\partial\Omega + \int_{\Omega(t)} \rho \mathbf{v}_E \cdot \mathbf{f}_{\text{cen}} \, d\Omega} \quad (\text{E.86})$$

Les expressions des bilans de conservation de la masse (E.49), de la quantité de mouvement (E.62) et de l'énergie (E.86) établies ici justifient donc la formulation FALÉ-REVE des équations de Navier-Stokes écrites à l'équation (4.33) au moyen de la description ALE dans le repère entraîné  $\mathcal{R}_E$  avec les composantes de la vitesse relative  $\mathbf{v}_E$ .

## E.2 Impact du jeu de variables sur la forme des équations de Navier-Stokes

### E.2.1 Écriture générique des équations du fluide

Dans le cas général, le fluide est décrit par un jeu de  $n_v = d + 2$  variables indépendantes où  $d$  est la dimension du domaine spatial  $\Omega \subset \mathbb{R}^d$ . On notera  $\mathbf{q} = [q_1 \dots q_{n_v}]^\top$  le vecteur des inconnues qui est supposé être un élément de l'espace  $H^{n_v} = (L^2(\Omega))^{n_v}$ . Les équations du fluide forment un système de  $n_v$  équations non-linéaires formulées pour les variables  $q_i$ . Les non-linéarités sont souvent polynomiales quadratiques (termes en  $q_i q_j$ ), cubiques (termes en  $q_i q_j q_k$ ), voire même tétraédriques (termes en  $q_i q_j q_k q_l$ ) suivant le jeu de variables employé. Pour simplifier l'écriture des équations, les non-linéarités polynomiales sont interprétées comme provenant de l'action d'opérateurs multilinéaires sur le vecteur des inconnues  $\mathbf{q}$  en suivant le même raisonnement qu'à l'équation (2.58).

Pour clarifier les choses, on considère l'exemple simple d'une application  $\Xi : H^{n_v} \rightarrow H$  avec  $n_v \geq 2$  qui à un vecteur  $\mathbf{q} \in H^{n_v}$  associe une fonction continue quadratique telle que  $\Xi(\mathbf{q}) = q_1 q_2$ . L'application  $\Xi$  n'est évidemment pas linéaire par rapport au vecteur  $\mathbf{q}$  puisque pour tout  $\lambda \in \mathbb{R}$  et  $\mathbf{r} \in H^{n_v}$ ,  $\Xi(\mathbf{q} + \lambda \mathbf{r}) = (q_1 + \lambda r_1)(q_2 + \lambda r_2) = q_1 q_2 + \lambda^2 r_1 r_2 + \lambda(q_1 r_2 + q_2 r_1) \neq \Xi(\mathbf{q}) + \lambda \Xi(\mathbf{r})$ . Cependant, si on considère à présent l'application  $\Xi' : H \times H \rightarrow H$  comme une fonction des composantes  $q_1$  et  $q_2$  telle que  $\Xi'(q_1, q_2) = q_1 q_2$ , alors on obtient une application linéaire par rapport à chaque variable, *i. e.* une application multilinéaire, puisque pour tout  $\lambda \in \mathbb{R}$  et  $r_1 \in H$ ,  $\Xi'(q_1 + \lambda r_1, q_2) = (q_1 + \lambda r_1) q_2 = q_1 q_2 + \lambda r_1 q_2 = \Xi'(q_1, q_2) + \lambda \Xi'(r_1, q_2)$ . Plutôt que de définir l'opérateur multilinéaire par rapport à chacune des composantes intervenant dans les termes non-linéaires, il est possible de le définir directement par rapport au vecteur  $\mathbf{q}$  contenant toutes les composantes. On introduit donc finalement l'application  $\Xi'' : H^{n_v} \times H^{n_v} \rightarrow H$  telle que  $\Xi''(\mathbf{q}, \mathbf{r}) = q_1 r_2$ . Cette dernière application est bien linéaire par rapport à chaque argument puisque pour tout  $\lambda \in \mathbb{R}$  et  $\mathbf{s} \in H^{n_v}$ ,  $\Xi''(\mathbf{q} + \lambda \mathbf{s}, \mathbf{r}) = (q_1 + \lambda s_1) r_2 = q_1 r_2 + \lambda s_1 r_2 = \Xi''(\mathbf{q}, \mathbf{r}) + \lambda \Xi''(\mathbf{s}, \mathbf{r})$ . Pour retrouver l'image  $q_1 q_2$  de la fonction initiale  $\Xi$  au moyen de la fonction  $\Xi''$ , il suffit d'interpréter le terme quadratique comme l'action de l'opérateur  $\Xi''(\mathbf{q}, \mathbf{q})$  dont les deux arguments correspondent au même vecteur d'inconnues  $\mathbf{q}$ . Le principe exposé ici pour un terme quadratique se généralise bien entendu pour n'importe quel degré de non-linéarité polynomiale : dans ce cas, le nombre d'arguments de l'opérateur multilinéaire est égal au degré de non-linéarité du terme à interpréter.

En appliquant ce formalisme au système d'équations de la mécanique des fluides, on verra qu'il est possible de parvenir à l'écriture synthétique

$$\mathbf{M}(\mathbf{q}) \dot{\mathbf{q}} = \mathbf{K} + \mathbf{L}(\mathbf{q}) + \mathbf{Q}(\mathbf{q}, \mathbf{q}) + \mathbf{C}(\mathbf{q}, \mathbf{q}, \mathbf{q}) + \dots \quad (\text{E.87})$$

faisant intervenir une somme d'opérateurs multilinéaires dont le nombre dépend du degré de non-linéarité des termes impliqués dans les équations. L'expression précédente suppose que les non-linéarités des équations sont au maximum cubiques, mais il n'est pas impossible que des non-linéarités d'ordre supérieur apparaissent. Au premier membre,  $\mathbf{M}$  est une matrice de  $\mathcal{M}(n_v, n_v)$ , alors que les quantités du second membre sont toutes des vecteurs de  $\mathbb{R}^{n_v}$ ;  $\mathbf{K}$  correspond au terme constant,  $\mathbf{L}$  contient les opérateurs linéaires en  $\mathbf{q}$ ,  $\mathbf{Q}$  les opérateurs quadratiques et enfin  $\mathbf{C}$  les opérateurs cubiques.

Toutefois, les équations (4.28) écrites au moyen des variables conservatives ne se mettent pas sous cette forme comme on l'a fait remarquer au paragraphe 4.4.1. Dans cette annexe, plusieurs écritures équivalentes des équations de Navier-Stokes sont donc développées en adoptant d'autres jeux de variables. L'objectif est de parvenir au système d'équations polynomial le plus simple possible afin de pouvoir construire ensuite un modèle d'ordre réduit POD-Galerkin au moyen d'une projection de ces équations. Pour les différents jeux de variables envisagés, on écrit donc les équations en faisant apparaître explicitement dans leurs expressions uniquement les inconnues choisies. Cependant, pour éviter de trop surcharger les expressions, le tenseur des contraintes visqueuses  $\boldsymbol{\tau}$  sera interprété comme un opérateur linéaire en  $\mathbf{q}$  puisque celui-ci dépend linéairement de la vitesse selon la définition  $\boldsymbol{\tau}(\mathbf{u}) = \mu(\nabla \mathbf{u} + \nabla^T \mathbf{u}) + \lambda(\text{div } \mathbf{u}) \mathbf{I}$ .

### E.2.2 Variables quasi-conservatives $\mathbf{q} = [\rho, \mathbf{u}, e_I]$ et $\mathbf{q} = [\rho, \mathbf{u}, e_T]$

Les variables conservatives posent problème en raison de l'apparition de termes isolés  $\mathbf{u}$  et  $e_T$  dans les équations alors que les variables sont  $\rho \mathbf{u}$  et  $\rho e_T$ . Par conséquent l'utilisation de variables « quasi-conservatives » est un moyen d'écrire les équations sous forme polynomiale. L'expression des bilans de conservation (4.14), (4.16) et (4.19) est donc développée pour faire apparaître exclusivement les variables  $\mathbf{q} = [\rho, \mathbf{u}, e_I]$ . Pour cela, le tenseur des contraintes  $\boldsymbol{\sigma}$  est dans un premier temps décomposé en  $\boldsymbol{\sigma} = -p\mathbf{I} + \boldsymbol{\tau}$ . La pression  $p$  participant aux contraintes normales s'exprime au moyen de l'équation d'état (4.26). Enfin, la température apparaissant dans le bilan d'énergie se déduit de l'équation des gaz parfaits (4.25) et de la relation précédente qui donnent  $\theta = e_I/c_v$ . Le système d'équations utilisant le jeu de variables quasi-conservatives s'écrit alors

$$\begin{cases} \frac{\partial \rho}{\partial t} + \text{div}(\rho \mathbf{u}) & = 0 \\ \rho \frac{\partial \mathbf{u}}{\partial t} + \rho \mathbf{u} \cdot \nabla \mathbf{u} & = -\nabla[\rho(\gamma - 1)e_I] + \text{div } \boldsymbol{\tau}(\mathbf{u}) + \rho \mathbf{f}_{\text{vol}} \\ \rho \frac{\partial e_I}{\partial t} + \rho \mathbf{u} \cdot \nabla e_I & = -\rho(\gamma - 1)e_I \text{div } \mathbf{u} + \boldsymbol{\tau}(\mathbf{u}) : \nabla \mathbf{u} + \text{div} \left( \frac{k_\theta}{c_v} \nabla e_I \right) \end{cases}, \quad (\text{E.88})$$

et des non-linéarités cubiques font leur apparition. Le système prend donc la forme synthétique

$$\mathbf{M}(\mathbf{q}) \dot{\mathbf{q}} = \mathbf{L}(\mathbf{q}) + \mathbf{Q}(\mathbf{q}, \mathbf{q}) + \mathbf{C}(\mathbf{q}, \mathbf{q}, \mathbf{q}), \quad (\text{E.89})$$

où les opérateurs de la forme générique (E.89) associée aux variables quasi-conservatives  $\mathbf{q} = [\rho, \mathbf{u}, e_I]^T$  sont définis par

$$\begin{aligned} \mathbf{M}(\mathbf{q}) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \rho & 0 \\ 0 & 0 & \rho \end{bmatrix}, \quad \mathbf{L}(\mathbf{q}) = \begin{bmatrix} 0 \\ \text{div } \boldsymbol{\tau}(\mathbf{u}) + \rho \mathbf{f}_{\text{vol}} \\ \text{div} (k_\theta/c_v \nabla e_I) \end{bmatrix}, \quad \mathbf{Q}(\mathbf{q}, \mathbf{q}) = \begin{bmatrix} -\text{div}(\rho \mathbf{u}) \\ -\nabla[\rho(\gamma - 1)e_I] \\ \boldsymbol{\tau}(\mathbf{u}) : \nabla \mathbf{u} \end{bmatrix} \\ \mathbf{C}(\mathbf{q}, \mathbf{q}, \mathbf{q}) &= \begin{bmatrix} 0 \\ -\rho \mathbf{u} \cdot \nabla \mathbf{u} \\ -\rho \mathbf{u} \cdot \nabla e_I - \rho(\gamma - 1)e_I \text{div } \mathbf{u} \end{bmatrix}. \end{aligned} \quad (\text{E.90})$$

Avec les variables quasi-conservatives  $\mathbf{q} = [\rho, \mathbf{u}, e_T]^T$  faisant intervenir l'énergie totale  $e_T$  au lieu de l'énergie interne, le bilan de conservation de l'énergie totale (4.21) remplace alors celui portant sur l'énergie

interne. Les relations donnant la pression  $p = \rho(\gamma - 1)e_I$  et la température  $\theta = e_I/c_v$  se transforment en remplaçant l'énergie interne par la différence  $e_T - e_C$ . Or l'énergie cinétique est une forme quadratique de la vitesse ( $e_C = 1/2\mathbf{u}^T\mathbf{u}$ ), ce qui complique les équations et augmente les non-linéarités. Par conséquent, la forme générique des équations écrites avec les variables  $\mathbf{q} = [\rho, \mathbf{u}, e_T]^T$  présente en plus une non-linéarité tétraédrique telle que

$$\mathbf{M}(\mathbf{q}) \dot{\mathbf{q}} = \mathbf{L}(\mathbf{q}) + \mathbf{Q}(\mathbf{q}, \mathbf{q}) + \mathbf{C}(\mathbf{q}, \mathbf{q}, \mathbf{q}) + \mathbf{T}(\mathbf{q}, \mathbf{q}, \mathbf{q}, \mathbf{q}). \quad (\text{E.91})$$

Les opérateurs de la forme générique (E.91) associée aux variables quasi-conservatives  $\mathbf{q} = [\rho, \mathbf{u}, e_T]^T$  contiennent plus de termes que lorsque les variables  $\mathbf{q} = [\rho, \mathbf{u}, e_I]^T$  sont employées, et en outre, ces termes sont fortement non-linéaires. La matrice  $\mathbf{M}(\mathbf{q})$  reste identique et les autres opérateurs sont donnés par les expressions

$$\begin{aligned} \mathbf{L}(\mathbf{q}) &= \begin{bmatrix} 0 \\ \text{div } \boldsymbol{\tau}(\mathbf{u}) + \rho \mathbf{f}_{\text{vol}} \\ \text{div } (k_\theta/c_v \nabla e_T) \end{bmatrix}, \quad \mathbf{Q}(\mathbf{q}, \mathbf{q}) = \begin{bmatrix} -\text{div } (\rho \mathbf{u}) \\ -\nabla[\rho(\gamma - 1)e_T] \\ \rho \mathbf{u} \cdot \mathbf{f}_{\text{vol}} + \text{div } [\boldsymbol{\tau}(\mathbf{u}) \mathbf{u}] - \text{div } [k_\theta/c_v \nabla(\mathbf{u}^T \mathbf{u})/2] \end{bmatrix}, \\ \mathbf{C}(\mathbf{q}, \mathbf{q}, \mathbf{q}) &= \begin{bmatrix} 0 \\ -\rho \mathbf{u} \cdot \nabla \mathbf{u} + \nabla[\rho(\gamma - 1)(\mathbf{u}^T \mathbf{u})]/2 \\ -\rho \mathbf{u} \cdot \nabla e_T - \text{div } [\rho \mathbf{u}(\gamma - 1)e_T] \end{bmatrix}, \quad \mathbf{T}(\mathbf{q}, \mathbf{q}, \mathbf{q}, \mathbf{q}) = \begin{bmatrix} 0 \\ 0 \\ \text{div } [\rho \mathbf{u}(\gamma - 1)(\mathbf{u}^T \mathbf{u})/2] \end{bmatrix}. \end{aligned} \quad (\text{E.92})$$

L'utilisation des variables quasi-conservatives est peu appropriée — notamment lorsque l'énergie totale est employée — si l'objectif est ensuite de construire un modèle d'ordre réduit POD-Galerkin sur la base de ces équations puisque des non-linéarités cubiques voire tétraédriques apparaissent. En outre, le système dynamique prend une forme implicite  $f(\dot{\mathbf{q}}, \mathbf{q}, t) = 0$ .

### E.2.3 Variables thermiques $\mathbf{q} = [\rho, \mathbf{u}, \theta]$

L'énergie interne est remplacée par la température  $\theta$  pour formuler le dernier bilan. La pression provenant du tenseur des contraintes  $\boldsymbol{\sigma}$  s'exprime simplement en fonction de la température d'après la loi des gaz parfaits (4.25) :  $p = \rho R_s \theta$ . Par ailleurs, l'énergie interne s'écrit simplement  $e_I = c_v \theta$ , si bien que le système d'équations pour les variables thermiques devient

$$\begin{cases} \frac{\partial \rho}{\partial t} + \text{div } (\rho \mathbf{u}) &= 0 \\ \rho \frac{\partial \mathbf{u}}{\partial t} + \rho \mathbf{u} \cdot \nabla \mathbf{u} &= -\nabla(\rho R_s \theta) + \text{div } \boldsymbol{\tau}(\mathbf{u}) + \rho \mathbf{f}_{\text{vol}} \\ \rho \frac{\partial \theta}{\partial t} + \rho \mathbf{u} \cdot \nabla \theta &= -\rho(\gamma - 1)\theta \text{div } \mathbf{u} + \frac{1}{c_v} \boldsymbol{\tau}(\mathbf{u}) : \nabla \mathbf{u} + \text{div } \left( \frac{k_\theta}{c_v} \nabla \theta \right) \end{cases}. \quad (\text{E.93})$$

La structure du système reste identique à celle obtenue avec l'énergie interne (E.89). La matrice de masse est identique et les autres opérateurs multilinéaires conservent une forme proche donnée par

$$\begin{aligned} \mathbf{L}(\mathbf{q}) &= \begin{bmatrix} 0 \\ \text{div } \boldsymbol{\tau}(\mathbf{u}) + \rho \mathbf{f}_{\text{vol}} \\ \text{div } (k_\theta/c_v \nabla \theta) \end{bmatrix}, \quad \mathbf{Q}(\mathbf{q}, \mathbf{q}) = \begin{bmatrix} -\text{div } (\rho \mathbf{u}) \\ -\nabla(\rho R_s \theta) \\ \boldsymbol{\tau}(\mathbf{u}) : \nabla \mathbf{u}/c_v \end{bmatrix}, \\ \mathbf{C}(\mathbf{q}, \mathbf{q}, \mathbf{q}) &= \begin{bmatrix} 0 \\ -\rho \mathbf{u} \cdot \nabla \mathbf{u} \\ -\rho \mathbf{u} \cdot \nabla \theta - \rho(\gamma - 1)\theta \text{div } \mathbf{u} \end{bmatrix}. \end{aligned} \quad (\text{E.94})$$

L'utilisation des variables thermiques ne présente donc pas plus d'intérêt que le jeu de variables basé sur l'énergie interne du système puisque les non-linéarités cubiques persistent, de même que le caractère implicite des équations aux dérivées partielles du système dynamique.

### E.2.4 Variables primitives $\mathbf{q} = [\rho, \mathbf{u}, p]$

Le choix de la pression comme seconde variable thermodynamique se justifie par le fait qu'elle apparaît naturellement dans l'expression du tenseur des contraintes normales. Les bilans de conservation de la masse (4.14) et de la quantité de mouvement (4.16) ne nécessitent aucune modification. Le bilan de pression peut s'obtenir en substituant  $\rho = p/(\theta R_s)$  dans le bilan de conservation de la masse  $D\rho/Dt = \rho \operatorname{div} \mathbf{u}$  où  $D \cdot /Dt$  correspond à la dérivée matérielle. En explicitant la dérivée et en réorganisant les termes on obtient une relation entre la dérivée de la température et de la pression :

$$\frac{D}{Dt} \left( \frac{p}{\theta R_s} \right) + \frac{p}{\theta R_s} \operatorname{div} \mathbf{u} = 0 \iff \frac{D\theta}{Dt} = \frac{1}{\rho R_s} \left( p \operatorname{div} \mathbf{u} + \frac{Dp}{Dt} \right). \quad (\text{E.95})$$

Le premier membre du bilan de température peut donc être remplacé par l'expression précédente ; enfin, les autres termes impliquant la température  $\theta$  sont modifiés à l'aide de la relation  $\theta = p/(\rho R_s)$ . Après calcul, on parvient au système d'équations

$$\begin{cases} \frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{u}) & = 0 \\ \rho \frac{\partial \mathbf{u}}{\partial t} + \rho \mathbf{u} \cdot \nabla \mathbf{u} & = -\nabla p + \operatorname{div} \boldsymbol{\tau}(\mathbf{u}) + \rho \mathbf{f}_{\text{vol}} \\ \frac{\partial p}{\partial t} + \mathbf{u} \cdot \nabla p & = -\gamma p \operatorname{div} \mathbf{u} + (\gamma - 1) \boldsymbol{\tau}(\mathbf{u}) : \nabla \mathbf{u} + \operatorname{div} \left( \frac{k_\theta}{c_v} \nabla \left( \frac{p}{\rho} \right) \right) \end{cases}, \quad (\text{E.96})$$

qui est intéressant puisque la quasi-totalité des termes est quadratique. Cependant, le terme de diffusion de la chaleur s'exprime à présent en fonction du rapport  $p/\rho$  et introduit par conséquent une fraction rationnelle. Il n'est donc pas possible d'écrire le bilan de pression sous une forme polynomiale simple et propice à l'élaboration du modèle d'ordre réduit.

### E.2.5 Variables primitives modifiées $\mathbf{q} = [\vartheta, \mathbf{u}, p]$

Pour résumer, le principal problème de la mise en forme des équations découle de la masse volumique qui apparaît soit dans le membre de gauche et rend ainsi le système d'équations aux dérivées partielles implicite, soit au dénominateur de certains termes si bien que les équations ne sont pas polynomiales. Le jeu de variables primitives modifiées employé à l'origine par Vigo [240] et Iollo *et al.* [117] permet de remédier à ces deux problèmes par l'introduction du *covolume* — aussi appelé *volume spécifique* —  $\vartheta = 1/\rho$ . Les bilans de pression et de quantité de mouvement dans (E.96) s'expriment facilement en fonction du covolume. En revanche, le changement de variable  $\rho = 1/\vartheta$  dans l'équation (4.14) nécessite de transformer le bilan en

$$\frac{\partial}{\partial t} \left( \frac{1}{\vartheta} \right) + \frac{1}{\vartheta} \operatorname{div} \mathbf{u} + \mathbf{u} \cdot \nabla \left( \frac{1}{\vartheta} \right) = 0 \iff - \left( \frac{1}{\vartheta^2} \right) \frac{\partial \vartheta}{\partial t} + \frac{1}{\vartheta} \operatorname{div} \mathbf{u} - \left( \frac{1}{\vartheta^2} \right) \mathbf{u} \cdot \nabla \vartheta = 0. \quad (\text{E.97})$$

Après simplification dans le bilan du covolume, les équations de Navier-Stokes écrites avec les variables primitives modifiées prennent la forme

$$\begin{cases} \frac{\partial \vartheta}{\partial t} + \mathbf{u} \cdot \nabla \vartheta & = \vartheta \operatorname{div} \mathbf{u} \\ \frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} & = -\vartheta \nabla p + \vartheta \operatorname{div} \boldsymbol{\tau}(\mathbf{u}) + \mathbf{f}_{\text{vol}} \\ \frac{\partial p}{\partial t} + \mathbf{u} \cdot \nabla p & = -\gamma p \operatorname{div} \mathbf{u} + (\gamma - 1) \boldsymbol{\tau}(\mathbf{u}) : \nabla \mathbf{u} + \operatorname{div} \left( \frac{k_\theta}{c_v} \nabla (\vartheta p) \right) \end{cases}. \quad (\text{E.98})$$

Le système présente une forme particulièrement intéressante puisque les équations s'écrivent sous la forme polynomiale explicite

$$\dot{\mathbf{q}} = \mathbf{K} + \mathbf{Q}(\mathbf{q}, \mathbf{q}), \quad (\text{E.99})$$

dans laquelle les non-linéarités sont au maximum quadratiques. Ce nouveau jeu de variables ne fait donc intervenir que des termes quadratiques — mis à part le terme d'efforts volumiques externes  $\mathbf{f}_{\text{vol}}$  qui est la seule contribution au terme constant  $\mathbf{K}$  — puisque les termes linéaires et cubiques apparaissant dans les autres formulations ont donc tous été rendus quadratiques. Les opérateurs multilinéaires associés à ce jeu de variables sont donc définis par

$$\mathbf{K} = \begin{bmatrix} 0 \\ \mathbf{f}_{\text{vol}} \\ 0 \end{bmatrix} \quad \text{et} \quad \mathbf{Q}(\mathbf{q}, \mathbf{q}) = \begin{bmatrix} -\mathbf{u} \cdot \nabla \vartheta + \vartheta \operatorname{div} \mathbf{u} \\ -\mathbf{u} \cdot \nabla \mathbf{u} - \vartheta \nabla p + \vartheta \operatorname{div} \boldsymbol{\tau}(\mathbf{u}) \\ -\mathbf{u} \cdot \nabla p - \gamma p \operatorname{div} \mathbf{u} + (\gamma - 1) \boldsymbol{\tau}(\mathbf{u}) : \nabla \mathbf{u} + \operatorname{div} [k_\theta / c_v \nabla(\vartheta p)] \end{bmatrix}. \quad (\text{E.100})$$

### E.3 Discrétisation par la méthode des Volumes Finis

On décrit brièvement dans cette section les techniques employées dans le logiciel *elsA* pour mener à bien la résolution des équations de Navier-Stokes. Pour simplifier, on considère ici la formulation eulérienne FEUL-RAUA écrite avec les variables conservatives. On détaille ensuite le calcul du gradient d'une grandeur physique quelconque par une approche Volumes Finis.

#### E.3.1 Discrétisation des équations de Navier-Stokes

On considère un domaine de calcul  $\Omega \subset \mathbb{R}^d$  de dimension  $d$ , de frontière  $\partial\Omega$  et de normale extérieure  $\mathbf{n}$ . Les lois de bilan (4.27) s'écrivent avec  $\mathbf{s}_A = \mathbf{0}$  dans le repère  $\mathcal{R}_A$  sous la forme

$$\frac{d}{dt} \int_{\Omega} \mathbf{q} \, d\Omega + \oint_{\partial\Omega} \mathcal{F}^C(\mathbf{q}) \cdot \mathbf{n} \, d\partial\Omega + \oint_{\partial\Omega} \mathcal{F}^D(\mathbf{q}, \nabla \mathbf{q}) \cdot \mathbf{n} \, d\partial\Omega = \int_{\Omega} \mathbf{T}(\mathbf{q}, \nabla \mathbf{q}) \, d\Omega \quad (\text{E.101})$$

qui sera le point de départ de la méthode des Volumes Finis. Le terme source  $\mathbf{T}$  est la plupart du temps nul — en l'absence de forces volumiques — sauf par exemple lorsque les équations de Navier-Stokes sont formulées dans un repère relatif entraîné, auquel cas le terme source contiendra des termes associés aux forces centrifuge et de Coriolis. La formulation (E.101) des équations de Navier-Stokes fait alors apparaître le flux convectif  $\oint_{\partial\Omega} \mathcal{F}^C(\mathbf{q}) \cdot \mathbf{n} \, d\partial\Omega$  et le flux diffusif  $\oint_{\partial\Omega} \mathcal{F}^D(\mathbf{q}, \nabla \mathbf{q}) \cdot \mathbf{n} \, d\partial\Omega$  définis respectivement par

$$\mathcal{F}^C(\mathbf{q}) = \begin{bmatrix} \rho \mathbf{u}_A \\ \rho \mathbf{u}_A \otimes \mathbf{u}_A + p \mathbf{I} \\ \rho e_T \mathbf{u}_A + p \mathbf{u}_A \end{bmatrix} \quad \text{et} \quad \mathcal{F}^D(\mathbf{q}, \nabla \mathbf{q}) = \begin{bmatrix} 0 \\ -\boldsymbol{\tau}(\mathbf{u}_A) \\ -\boldsymbol{\tau}(\mathbf{u}_A) \mathbf{u}_A - \mathbf{q}_{\theta, A} \nabla_A \theta \end{bmatrix}. \quad (\text{E.102})$$

La discrétisation des équations (E.101) s'effectue pour le domaine de calcul  $\Omega$ , partitionné en un ensemble de cellules de contrôle  $\Omega_i$  de centre  $M_i$  telles que  $\Omega = \cup_{i=1}^N \Omega_i$  et de volume

$$\mathcal{V}(\Omega_i) = \int_{\Omega_i} d\Omega. \quad (\text{E.103})$$

Les cellules de contrôle  $\Omega_i$  du maillage structuré sont des quadrilatères lorsque  $d = 2$  ou des hexaèdres lorsque  $d = 3$ . Par conséquent, la surface  $\partial\Omega_i$  délimitant la cellule  $\Omega_i$  peut être définie comme la réunion des  $n_f$  faces de la cellule<sup>2</sup>. En notant  $\Sigma_{i,j}$  la face commune entre la cellule  $\Omega_i$  et la cellule  $\Omega_j$  le contour  $\partial\Omega_i$  de la cellule  $\Omega_i$  s'écrit

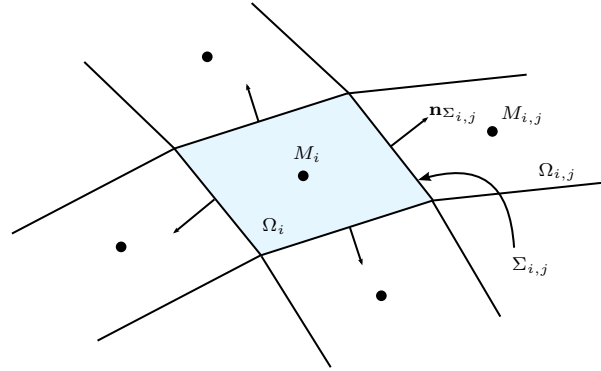
$$\partial\Omega_i = \bigcup_{j=1}^{n_f} \Sigma_{i,j}, \quad (\text{E.104})$$

et l'intégration sur la surface fermée  $\oint_{\partial\Omega_i}$  pourra être remplacée par une somme d'intégrales sur les faces  $\Sigma_{i,j}$  de la cellule au moyen de la relation

$$\oint_{\partial\Omega_i} \bullet \, d\partial\Omega = \sum_{j=1}^{n_f} \int_{\Sigma_{i,j}} \bullet \, d\Sigma. \quad (\text{E.105})$$

Le schéma de la figure E.2 représente les différentes quantités géométriques intervenant dans le formalisme Volumes Finis pour la discrétisation du problème.

<sup>2</sup>Le nombre de face  $n_f$  vaut 4 lorsque  $d = 2$  puisque les cellules sont des quadrilatères, et  $n_f = 6$  pour  $d = 3$  lorsque les cellules sont des hexaèdres.



**Figure E.2 :** Schématisation d'une cellule de contrôle en Volumes Finis et représentation des notations associées à la discrétisation spatiale du domaine  $\Omega$ .

En négligeant le terme source  $\mathbf{T}$  et en utilisant les relations précédentes, les lois de bilan (E.101) s'écrivent

$$\frac{d}{dt} \int_{\Omega_i} \mathbf{q} \, d\Omega = - \sum_{j=1}^{n_f} \int_{\Sigma_{i,j}} \left[ \mathcal{F}^C(\mathbf{q}) + \mathcal{F}^D(\mathbf{q}, \nabla \mathbf{q}) \right] \cdot \mathbf{n} \, d\Sigma \quad (\text{E.106})$$

pour chaque cellule de contrôle  $\Omega_i$ . On définit alors la valeur moyenne des variables conservatives  $\mathbf{q}$  dans la cellule  $\Omega$  par

$$\bar{\mathbf{q}}_i = \frac{1}{\mathcal{V}(\Omega_i)} \int_{\Omega_i} \mathbf{q} \, d\Omega, \quad (\text{E.107})$$

et de la même manière, on définit la valeur moyenne du flux à travers la surface  $\Sigma_{i,j}$  par

$$\bar{\mathcal{F}}_{\Sigma_{i,j}} = \int_{\Sigma_{i,j}} \left[ \mathcal{F}^C(\mathbf{q}) + \mathcal{F}^D(\mathbf{q}, \nabla \mathbf{q}) \right] \cdot \mathbf{n} \, d\Sigma. \quad (\text{E.108})$$

Pour chaque cellule de contrôle  $\Omega_i$ , le système d'équations (E.101) prend ainsi la forme

$$\frac{d}{dt} (\mathcal{V}(\Omega_i) \bar{\mathbf{q}}_i) = - \sum_{j=1}^{n_f} \bar{\mathcal{F}}_{\Sigma_{i,j}}. \quad (\text{E.109})$$

Puisque le maillage est supposé fixe, le volume  $\mathcal{V}(\Omega_i)$  peut être extrait de l'opérateur de dérivation dans le premier membre. La résolution numérique du système d'équations précédent consiste à chercher une approximation numérique  $\mathbf{w}_i$  de la valeur moyenne  $\bar{\mathbf{q}}_i$  des variables conservatives dans les cellules  $\Omega_i$ . Pour cela, le flux exact  $\bar{\mathcal{F}}_{\Sigma_{i,j}}$  à travers une des faces  $\Sigma_{i,j}$  de la cellule  $\Omega_i$  est approché par une fonction flux numérique  $\mathbf{F}(\mathbf{w}_i, \mathbf{w}_{i,j}) = \mathbf{F}^C(\mathbf{w}_i, \mathbf{w}_{i,j}) + \mathbf{F}^D(\mathbf{w}_i, \mathbf{w}_{i,j})$  qui dépend des valeurs moyennes approchées  $\mathbf{w}_i$  des variables conservatives dans la cellule  $\Omega_i$  et des valeurs  $\mathbf{w}_{i,j}$  dans un certain nombre de cellules voisines. En utilisant ces approximations, le système d'équations (E.109) se récrit

$$\frac{d}{dt} \mathbf{w}_i = - \frac{1}{\mathcal{V}(\Omega_i)} \sum_{j=1}^{n_f} \mathbf{F}(\mathbf{w}_i, \mathbf{w}_{i,j}) \cdot \mathbf{n}_{\Sigma_{i,j}} = - \frac{1}{\mathcal{V}(\Omega_i)} \mathbf{R}_{\Omega_i}, \quad (\text{E.110})$$

où  $\mathbf{n}_{\Sigma_{i,j}}$  est la normale extérieure (non unitaire) à la face  $\Sigma_{i,j}$ , telle que

$$\mathbf{n}_{\Sigma_{i,j}} = \int_{\Sigma_{i,j}} \mathbf{n} \, d\Sigma \quad (\text{E.111})$$

et  $\mathbf{R}_{\Omega_i}$  est le résidu numérique de la modélisation donné par  $\mathbf{R}_{\Omega_i} = \sum_{j=1}^{n_f} \mathbf{F}(\mathbf{w}_i, \mathbf{w}_{i,j}) \cdot \mathbf{n}_{\Sigma_{i,j}}$ . Afin de résoudre le système d'équations (E.110), il reste encore à préciser d'une part comment l'intégration temporelle est effectuée, et d'autre part comment la fonction flux numérique est définie.



### E.3.1.1 Discrétisation en temps

Le système d'équations (E.110) peut être intégré en temps au moyen de la méthode de Runge-Kutta à 1, 3, 4 ou 5 pas par exemple. L'intégration en temps est explicite et s'adresse plutôt à des problèmes instationnaires rapides, faisant intervenir des phénomènes à haute fréquence ne justifiant pas l'utilisation de grands pas de temps.

Le schéma employé pour l'intégration en temps des équations semi-discrétisées (E.110) est le schéma d'Euler rétrograde, qui est un cas particulier de la méthode de Runge-Kutta avec un seul pas. Les variables conservatives  $\mathbf{w}_i^{n+1}$  à l'instant  $t^{n+1} = t^n + \Delta t$  se déduisent alors des valeurs connues des variables  $\mathbf{w}_i^n$  à l'instant  $t^n$  par la relation

$$\mathbf{w}_i^{n+1} - \mathbf{w}_i^n = -\frac{\Delta t}{\mathcal{V}(\Omega_i)} \mathbf{R}_{\Omega_i}^n. \quad (\text{E.112})$$

### E.3.1.2 Discrétisation des flux convectifs et diffusifs

Dans ce paragraphe, on donne les approximations employées pour la fonction flux numérique permettant de décrire les flux convectifs et diffusifs. Pour cela, on considère dans un premier temps l'approximation du flux convectif, pour lequel un schéma centré avec viscosité artificielle est adopté. La discrétisation dite « centrée simple » sur chaque face  $\Sigma_{i,j}$  fait intervenir une moyenne des flux calculés aux centres des deux cellules de contrôles adjacentes à  $\Sigma_{i,j}$ . Ce type de discrétisation conduit à un schéma inconditionnellement instable et doit par conséquent être stabilisé par l'adjonction d'une dissipation artificielle pour le calcul des ondes de chocs notamment. Plusieurs choix sont possibles dans *elsA* pour modéliser le terme de dissipation artificielle, et on optera ici pour la formulation proposée par Jameson [120].

L'approximation « centrée simple » de la partie convective du flux numérique stabilisé par un flux de dissipation artificielle s'écrit sur la face  $\Sigma_{i,j}$  de la cellule  $\Omega_i$

$$\mathbf{F}_{\text{Jameson}}^C(\mathbf{w}_i, \mathbf{w}_{i,j}) \cdot \mathbf{n}_{\Sigma_{i,j}} = \frac{1}{2} [\mathbf{F}^C(\mathbf{w}_i) + \mathbf{F}^C(\mathbf{w}_{i,j})] \cdot \mathbf{n}_{\Sigma_{i,j}} - \mathbf{D}_{\Sigma_{i,j}}, \quad (\text{E.113})$$

où  $\mathbf{w}_i$  et  $\mathbf{w}_{i,j}$  désignent respectivement les inconnues conservatives dans la cellule considérée  $\Omega_i$  et dans la cellule voisine  $\Omega_{i,j}$ , adjacente à la face  $\Sigma_{i,j}$ . Le dernier terme  $\mathbf{D}_{\Sigma_{i,j}}$  correspond à la dissipation artificielle pour la face  $\Sigma_{i,j}$ . L'approximation du flux convectif avec le schéma de Jameson [120] requiert donc tout d'abord l'évaluation aux centres des cellules  $\Omega_i$  et  $\Omega_{i,j}$ , de la densité de flux  $\mathbf{F}^C(\mathbf{w}_i)$  définie à l'équation (E.102), et ensuite le calcul du terme de dissipation artificiel. Ce dernier peut se mettre sous la forme  $\sum_{j=1}^{n_f} \mathbf{D}_{\Sigma_{i,j}} = \mathbf{D}_i + \mathbf{D}_j + \mathbf{D}_k$  en remplaçant (E.113) dans (E.110), ce qui permet de définir les contributions de l'opérateur de dissipation dans chacune des directions  $i, j$  et  $k$  du maillage structuré. Sans entrer dans le détail du principe de calcul du terme de dissipation artificiel, on peut signaler tout de même que les formulations adoptées comprennent un terme de dissipation non-linéaire du second ordre pour capter correctement les discontinuités de l'écoulement, et un terme de dissipation linéaire du quatrième ordre car le schéma n'est pas dissipatif au sens de Kreiss [24, 157].

Dans un second temps, les flux diffusifs doivent aussi être approchés en fonction des inconnues conservatives connues aux centres des cellules de calcul. La discrétisation nécessite deux opérations successives. La première étape consiste à évaluer les densités de flux diffusif (E.102), dont l'expression fait intervenir le gradient des vitesses et de la température. Cette première étape exige donc l'évaluation numérique de ces gradients à l'aide d'un schéma de discrétisation. Une fois ces densités de flux connues, la seconde opération a pour but de discrétiser les flux de diffusion et d'en effectuer le bilan. La partie diffusif du flux numérique s'obtient par un schéma similaire à celui de la formulation de Jameson, en effectuant une moyenne arithmétique des densités de flux de diffusion dans la cellule  $\Omega_i$  et la cellule  $\Omega_{i,j}$  adjacente. Le flux diffusif est alors approché par l'expression

$$\mathbf{F}_{\text{Jameson}}^D(\mathbf{w}_i, \mathbf{w}_{i,j}) \cdot \mathbf{n}_{\Sigma_{i,j}} = \frac{1}{2} [\mathbf{F}^D(\mathbf{w}_i) + \mathbf{F}^D(\mathbf{w}_{i,j})] \cdot \mathbf{n}_{\Sigma_{i,j}} \quad (\text{E.114})$$

qui requiert l'évaluation des densités de flux (E.102) au centre de la cellule  $\Omega_i$  et des cellules  $\Omega_{i,j}$  adjacentes. Pour cela, il est nécessaire d'évaluer les gradients des grandeurs conservatives aux centres des cellules de contrôle.

Finalement, les équations de Navier-Stokes (E.110) écrites avec les variables conservatives et discrétisées en temps par un schéma d'Euler rétrograde et en espace avec le schéma de Jameson prennent la forme

$$\begin{aligned} \mathbf{w}_i^{n+1} - \mathbf{w}_i^n &= -\frac{\Delta t}{\mathcal{V}(\Omega_i)} \sum_{j=1}^{n_f} \frac{1}{2} [\mathbf{F}^C(\mathbf{w}_i^n) + \mathbf{F}^C(\mathbf{w}_{i,j}^n)] \cdot \mathbf{n}_{\Sigma_{i,j}} \\ &\quad - \frac{\Delta t}{\mathcal{V}(\Omega_i)} \sum_{j=1}^{n_f} \frac{1}{2} [\mathbf{F}^D(\mathbf{w}_i^n) + \mathbf{F}^D(\mathbf{w}_{i,j}^n)] \cdot \mathbf{n}_{\Sigma_{i,j}} \\ &\quad - \frac{\Delta t}{\mathcal{V}(\Omega_i)} (\mathbf{D}i + \mathbf{D}j + \mathbf{D}k) \end{aligned} \quad (\text{E.115})$$

lorsque les bilans sont écrits pour une cellule de contrôle  $\Omega_i$ . Le processus de discrétisation a donc permis de passer des variables conservatives  $\mathbf{q} = [\rho, \rho \mathbf{u}_A, \rho e_T]$  correspondant à un vecteur de  $n_v$  variables appartenant à l'espace  $(L^2(\Omega_i))^{n_v}$  en supposant que les variables sont des fonctions de carré intégrable, à une approximation numérique  $\mathbf{w}_i \in \mathbb{R}^{n_v}$  représentant les valeurs discrètes moyennes des variables  $\mathbf{q}$  au centre de la cellule  $\Omega_i$ . Afin d'obtenir une solution du problème sur l'ensemble du domaine  $\Omega = \cup_{i=1}^N \Omega_i$ , il reste encore à assembler le système matriciel qui tiendra compte de l'ensemble des contributions  $\mathbf{w}_i$  pour chaque cellule. Le champ aérodynamique  $\mathbf{q} \in (L^2(\Omega))^{n_v}$  sur l'ensemble du domaine  $\Omega$  sera donc approché par le vecteur  $\mathbf{W} \in \mathbb{R}^{N n_v}$  avec  $N n_v = N n_v$  qui rassemble les contributions discrètes  $\mathbf{w}_i$  sur chaque cellule.

### E.3.2 Approximation du gradient en Volumes Finis

Lors de la construction du modèle d'ordre réduit par une approche POD-Galerkin, il est nécessaire de disposer des gradients des modes POD et de la partie moyenne des clichés. Puisque les clichés ont été calculés au moyen d'un code de calcul reposant sur une discrétisation en Volumes Finis, il est conseillé de conserver le même formalisme pour évaluer les gradients.

On considère donc le champ  $\psi = [\psi_1 \cdots \psi_N] \in \mathbb{R}^N$  regroupant les approximations des valeurs de la fonction scalaire continue  $\psi$  aux centres des cellules du maillage Volumes Finis. L'évaluation du gradient du champ scalaire  $\psi$  s'effectue au centre  $M_i$  d'une cellule de discrétisation  $\Omega_i$  délimitée par les faces  $\Sigma_{i,j}$ . On suppose que la valeur du gradient au centre de la cellule de discrétisation est la valeur moyenne dans toute la cellule. Ainsi le théorème de Green-Ostrogradski (D.29) s'applique à  $\psi$  et conduit à la relation

$$(\nabla \psi)_{M_i} = \frac{1}{\mathcal{V}(\Omega_i)} \int_{\Omega_i} \nabla \psi \, d\Omega_i = \frac{1}{\mathcal{V}(\Omega_i)} \sum_{j=1}^{n_f} \int_{\Sigma_{i,j}} \psi \, \mathbf{n} \, d\Sigma_{i,j} \quad (\text{E.116})$$

qui permet de ramener le calcul du gradient dans la cellule à un calcul sur la frontière  $\partial\Omega_i = \cup_{j=1}^{n_f} \Sigma_{i,j}$ . Dans l'expression précédente,  $\mathcal{V}(\Omega_i)$  représente le volume de la cellule de calcul, et  $\mathbf{n}$  correspond à la normale sortante à  $\Omega_i$  sur la face  $\Sigma_{i,j}$ . Pour être en mesure d'évaluer le gradient, on suppose que le champ  $\psi$  est uniforme sur chaque face  $\Sigma_{i,j}$  et la valeur représentative de  $\psi$  utilisée diffère ensuite suivant que la face est interne au domaine ou externe — *i. e.* sur le bord  $\partial\Omega$  du domaine. Pour une face interne au domaine, la valeur de  $\psi$  sur la face  $\Sigma_{i,j}$  est simplement approchée par la demi-somme de  $\psi$  évalué au centre  $M_i$  de la cellule  $\Omega_i$  et au centre  $M_{i,j}$  de sa voisine  $\Omega_{i,j}$ , si bien que

$$\int_{\Sigma_{i,j}} \psi \, \mathbf{n} \, d\Sigma_{i,j} \approx \frac{1}{2} (\psi_{M_i} + \psi_{M_{i,j}}) \mathbf{n}_{\Sigma_{i,j}}. \quad (\text{E.117})$$

De la même manière que pour le champ  $\psi$ , on suppose que la normale  $\mathbf{n}$  sur la face  $\Sigma_{i,j}$  est uniforme; on utilise alors dans l'approximation précédente la normale  $\mathbf{n}_{\Sigma_{i,j}}$  définie à l'équation (E.111) comme la normale moyenne sur la face  $\Sigma_{i,j}$ . En trois dimensions, la normale moyenne à la face  $\Sigma_{i,j}$  délimitée par les quatre sommets  $A, B, C$  et  $D$  est donnée par l'expression

$$\mathbf{n}_{\Sigma_{i,j}} = \epsilon \frac{1}{2} (\mathbf{OC} - \mathbf{OA}) \wedge (\mathbf{OD} - \mathbf{OB}), \quad (\text{E.118})$$

où  $\mathbf{OA}, \mathbf{OB}, \mathbf{OC}$  et  $\mathbf{OD}$  sont les vecteurs distance par rapport au point d'origine  $O$  du repère  $\mathcal{R}_A$  et  $\epsilon = \pm 1$  afin de toujours avoir une normale sortante.

Dans le cas d'une face  $\Sigma_{i,j}$  externe, le champ  $\psi$  sur la face est calculé à partir d'un état satisfaisant les conditions aux limites. C'est la méthode employée dans *elsA* puisque les conditions limites sont connues et s'appliquent aux champs dont on cherche les gradients. Cependant, lorsqu'il s'agit d'évaluer le gradient d'un champ tel que celui d'un mode POD, les conditions aux limites ne sont pas explicitement connues.

Plusieurs solutions peuvent alors être envisagées pour calculer les valeurs du gradient sur le bord du domaine. La méthode la plus intuitive consiste à extrapoler le champ  $\psi$  à partir des valeurs connues au centre des cellules à l'intérieur du domaine. Une variante de cette méthode repose sur un premier calcul du gradient à l'intérieur du domaine, suivi d'une extrapolation du gradient au centre des cellules de bord. Toutefois ces deux méthodes reposent sur une technique d'extrapolation qui n'est généralement pas recommandée. La dernière méthode proposée consiste à calculer le gradient des modes POD à partir du gradient des clichés en tirant profit de la décomposition linéaire (2.20) des modes POD sur les clichés. Les gradients des modes POD sont alors calculés comme la combinaison linéaire des gradients des clichés

$$\nabla\varphi^{(j)} = \frac{1}{\lambda_j} \sum_{m=1}^M c_m^{(j)} \left( \nabla u^{(m)} - \nabla \bar{u} \right). \quad (\text{E.119})$$

Le gradient de la partie moyenne  $\bar{u}$  s'obtient lui aussi à partir du gradient des clichés puisqu'il s'agit simplement de la moyenne des clichés pondérés par les coefficients  $\alpha_m$  relatifs à l'opérateur de moyenne temporelle  $E[\cdot]$ . Le gradient de la partie moyenne s'écrit alors

$$\nabla \bar{u} = E \left[ \nabla u^{(m)} \right] = \sum_{m=1}^M \alpha_m \nabla u^{(m)}. \quad (\text{E.120})$$

L'avantage de cette approche est qu'il est possible d'extraire d'*elsA* les valeurs des clichés sur le bord du domaine si bien que les valeurs des champs sur le bord sont exactement connues et n'ont pas besoin d'être extrapolées. La méthode peut néanmoins devenir assez coûteuse lorsque le nombre  $M$  de clichés employés dans la base de données initiale est élevé. En effet, il est indispensable d'évaluer le gradient des  $M$  clichés pour effectuer ensuite la combinaison linéaire, alors que les méthodes précédentes reposant sur une extrapolation ne nécessitent que l'évaluation de  $q$  gradients, *i. e.* autant que le nombre de modes POD qui est toujours inférieur au nombre  $M$  de clichés.

# Analyse des contributions des coefficients impliqués dans les modèles d'ordre réduit des équations de Navier-Stokes

## Sommaire

<b>F.1</b>	<b>Analyse des coefficients du modèle d'ordre réduit autonome</b>	<b>267</b>
F.1.1	Comparaison des contributions convectives et diffusives	267
F.1.2	Comparaison des contributions surfaciques et volumiques	268
<b>F.2</b>	<b>Analyse des coefficients du modèle d'ordre réduit non-autonome</b>	<b>271</b>
F.2.1	Comparaison des coefficients analytiques et identifiés	271
F.2.1.1	Comparaison des coefficients autonomes	271
F.2.1.2	Comparaison des coefficients non-autonomes	273
F.2.2	Comparaison des contributions autonomes et non-autonomes	274

Dans cette annexe, les contributions relatives des différents coefficients intervenant dans les modèles d'ordre réduit des équations du fluide développés aux chapitres 5 et 6 sont analysés afin d'évaluer leur participation à la réponse globale.

## F.1 Analyse des coefficients du modèle d'ordre réduit autonome

### F.1.1 Comparaison des contributions convectives et diffusives

Les expressions des coefficients du modèle d'ordre réduit font apparaître plusieurs contributions correspondant d'une part aux termes convectifs ou diffusifs (5.9), et d'autre part aux termes volumiques et surfaciques issus des intégrations par parties pratiquées sur les termes diffusifs (5.21). On se propose donc ici d'analyser les contributions relatives de ces deux groupes, au moyen du modèle d'ordre réduit employé au paragraphe 5.5 avec une base de projection constituée de  $q = 10$  modes POD.

Tout d'abord les contributions convectives ( $\mathcal{K}_i^C, \mathcal{L}_{ij}^C$  et  $\mathcal{Q}_{ijk}^C$ ) et diffusives ( $\mathcal{K}_i^D, \mathcal{L}_{ij}^D$  et  $\mathcal{Q}_{ijk}^D$ ) contribuant aux coefficients du modèle d'ordre réduit sont comparées en fonction de leur participation aux coefficients totaux ( $\mathcal{K}_i, \mathcal{L}_{ij}$  et  $\mathcal{Q}_{ijk}$ ). Pour cela, on définit les rapports

$$\Pi^C(\mathcal{C}) = \log \left| 100 \frac{\mathcal{C}^C}{\mathcal{C}} \right| \quad \text{et} \quad \Pi^D(\mathcal{C}) = \log \left| 100 \frac{\mathcal{C}^D}{\text{Re} \mathcal{C}} \right|, \quad (\text{F.1})$$

où le terme  $\mathcal{C}$  représente selon le cas la contribution constante, linéaire ou quadratique. Le graphique F.1 présente ainsi les contributions relatives  $\Pi^C$  et  $\Pi^D$  des coefficients constants convectifs  $\mathcal{K}_i^C$  et diffusifs  $\mathcal{K}_i^D$ .

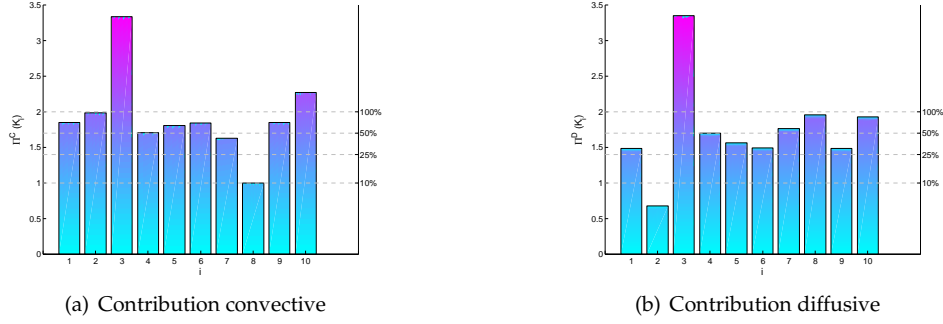


Figure F.1 : Comparaison des contributions relatives des termes convectifs et diffusifs constants  $\mathcal{K}_i^C$  et  $\mathcal{K}_i^D$ .

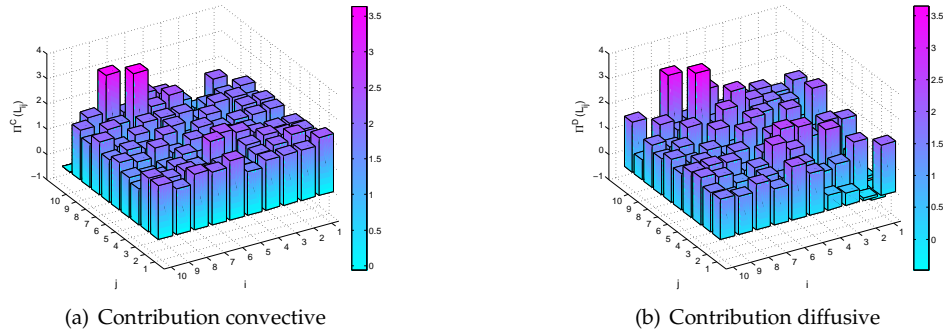


Figure F.2 : Comparaison des contributions relatives des termes convectifs et diffusifs linéaires  $\mathcal{L}_{ij}^C$  et  $\mathcal{L}_{ij}^D$ .

Parfois, la contribution dépasse largement les 100% puisque les contributions considérées séparément ont des signes opposés et leur somme produit une contribution totale inférieure en valeur absolue.

Le graphique F.2 représente les contributions relatives des termes linéaires convectifs  $\mathcal{L}_{ij}^C$  et diffusifs  $\mathcal{L}_{ij}^D$ . Pour quelques couples d'indices  $(i, j)$ , les contributions relatives des termes linéaires dépassent les 100%, pour la même raison que celle évoquée pour les coefficients constants. Comme pour ces derniers représentés sur le graphique F.1, les contributions convectives et diffusives des termes linéaires jouent chacune un rôle significatif.

Enfin les graphiques de la figure F.3 représentent, pour trois tranches ( $k = 1, 5$  et  $10$ ) du tenseur  $\mathcal{Q}$ , les contributions relatives des termes convectifs  $\mathcal{Q}_{ijk}^C$  et diffusifs  $\mathcal{Q}_{ijk}^D$ . Cette fois-ci, la contribution des termes convectifs est largement prépondérante par rapport à celle des termes diffusifs. Pour la majorité des coefficients, la contribution convective avoisine les 100% alors que la contribution diffusive est non négligeable seulement pour quelques triplets  $(i, j, k)$ . Ceci signifie que la contribution diffusive participe surtout à la partie moyenne de l'écoulement, et joue un rôle moindre dans les fluctuations. En effet, les termes constants et linéaires proviennent de la partie moyenne  $\bar{\mathbf{q}}$  de l'écoulement alors que la partie quadratique est purement liée aux fluctuations de la solution.

## F.1.2 Comparaison des contributions surfaciques et volumiques

Par analogie, les contributions provenant des termes surfaciques et volumiques sont comparées en se basant sur leur participation relative au coefficient diffusif. En effet, les contributions surfaciques et volumiques apparaissent comme résultat de l'intégration par parties des termes diffusifs (voir le paragraphe 5.2.2.2) de sorte que  $\mathcal{Q}_{ijk}^D = \mathcal{Q}_{\text{Vol}}^D + \mathcal{Q}_{\text{Surf}}^D$ <sup>1</sup>. La comparaison s'effectue au moyen des rapports  $\Pi^S$  et  $\Pi^V$ , dé-

<sup>1</sup>Des coefficients surfaciques et volumiques pour les termes constants ( $\mathcal{K}_{\text{Surf}}^D$  et  $\mathcal{K}_{\text{Vol}}^D$ ) et linéaires ( $\mathcal{L}_{\text{Surf}}^D$  et  $\mathcal{L}_{\text{Vol}}^D$ ) apparaissent aussi lorsque l'opérateur quadratique diffusif  $\mathbf{Q}^D(\mathbf{q}, \mathbf{q})$  de l'équation (5.5) s'applique respectivement aux couples  $(\bar{\mathbf{q}}, \bar{\mathbf{q}})$  et  $(\varphi^{(j)}, \bar{\mathbf{q}})$ .

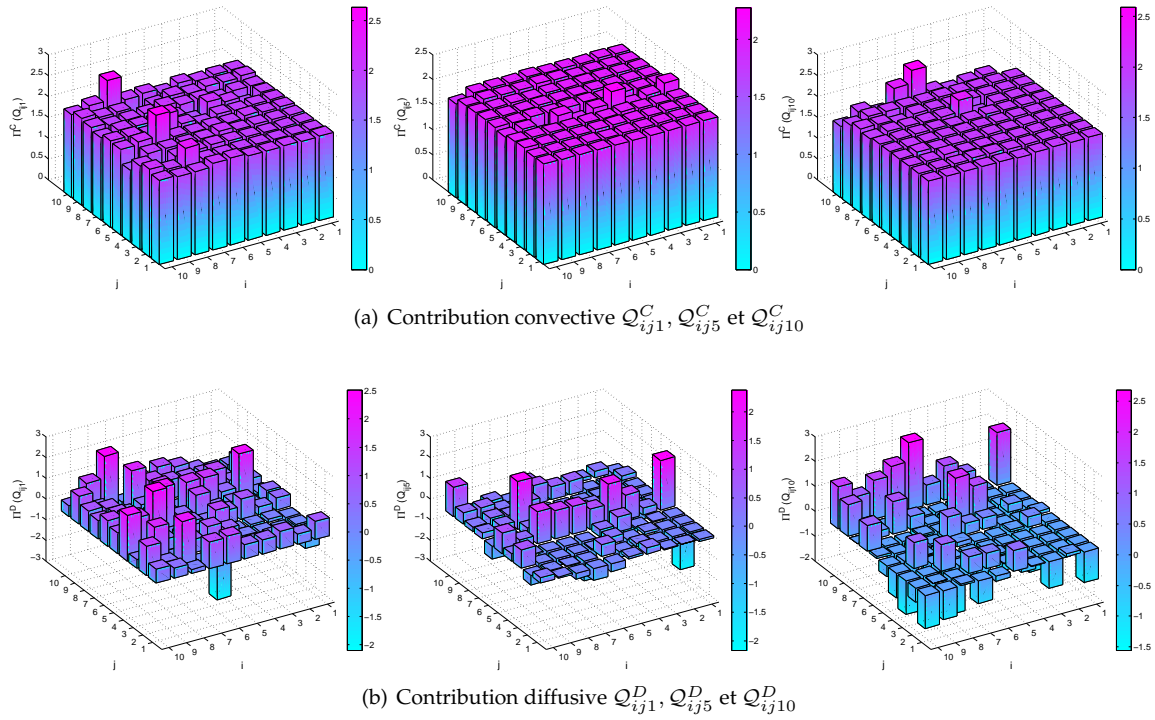


Figure F.3 : Comparaison des contributions relatives des termes convectifs et diffusifs quadratiques  $Q_{ijk}^C$  et  $Q_{ijk}^D$ .

finis de la même manière que le rapport  $\Pi^C$  donné à l'équation (F.1), en choisissant à la place de  $C^C$  la contribution surfacique ou volumique et au lieu du terme  $C$  la contribution diffusive totale.

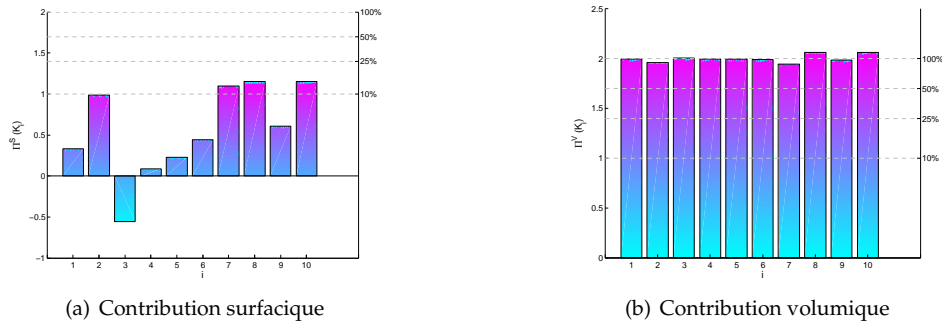


Figure F.4 : Comparaison des contributions relatives des termes surfaciques et volumiques constants  $K_{Surr}^D$  et  $K_{Vol}^D$ .

Le graphique F.4 présente ainsi les contributions relatives  $\Pi^S$  et  $\Pi^V$  des coefficients constants surfaciques  $K_{Surr}^D$  et volumiques  $K_{Vol}^D$ . On constate ici que les contributions surfaciques sont peu prononcées et ne dépassent jamais 25% du coefficient total. Le même phénomène est observé sur les coefficients linéaires et quadratiques représentés sur les figures F.5 et F.6. Les contributions surfaciques ne représentent à chaque fois qu'une faible proportion des coefficients totaux et semblent même être dans la plupart des cas négligeables.

La comparaison effectuée ci-dessus laisse supposer que les contributions surfaciques pourraient être négligées lors du calcul des coefficients du modèle d'ordre réduit, à l'instar des simplifications pratiquées par

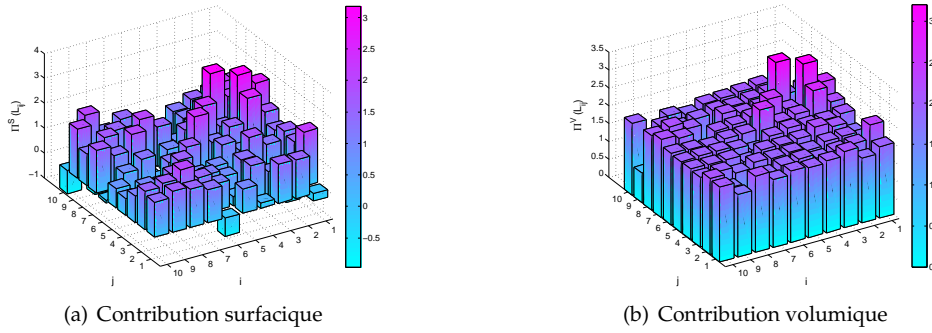


Figure F.5 : Comparaison des contributions relatives des termes surfaciques et volumiques linéaires  $\mathcal{L}_{\text{Surf}}^D$  et  $\mathcal{L}_{\text{Vol}}^D$ .

Iollo *et al.* [117]. La faible participation des termes surfaciques provient des conditions aux limites imposées sur les frontières du domaine où les quantités aérodynamiques varient peu. Par conséquent, les modes POD calculés à partir des fluctuations des variables aérodynamiques sont quasiment nuls sur les bords du domaine, d'où les faibles contributions des termes surfaciques dans cette zone. Cependant, on choisit ici de ne pas négliger le calcul des termes surfaciques afin d'éviter au maximum les erreurs d'évaluation des coefficients pouvant déstabiliser la réponse du modèle d'ordre réduit.

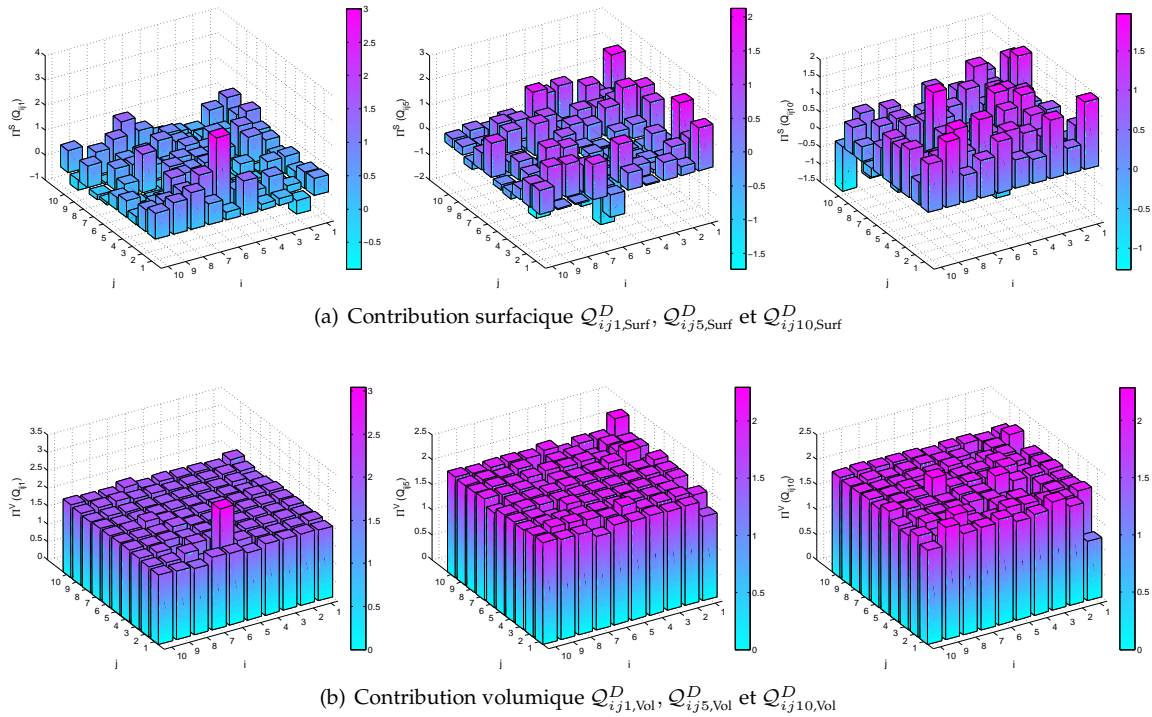


Figure F.6 : Comparaison des contributions relatives des termes surfaciques et volumiques quadratiques  $Q_{ijk,\text{Surf}}^D$  et  $Q_{ijk,\text{Vol}}^D$ .

## F.2 Analyse des coefficients du modèle d'ordre réduit non-autonome

Dans cette section, les coefficients du modèle d'ordre réduit développé au chapitre 6 sont dans un premier temps comparés lorsqu'ils sont évalués au moyen des expressions analytiques ou identifiés à partir des clichés. Ensuite, les contributions des termes autonomes et non-autonomes sont analysées.

### F.2.1 Comparaison des coefficients analytiques et identifiés

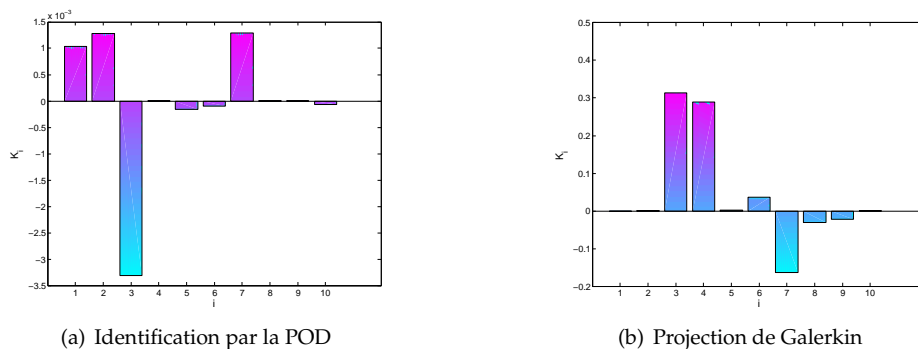
Le modèle d'ordre réduit considéré ici est celui du paragraphe 6.5 construit à partir de  $M = 120$  clichés extraits sur  $n_p \approx 2,5$  périodes d'oscillation et une base de projection comportant  $q = 10$  modes POD. Les coefficients analytiques sont d'abord corrigés avec la correction de [Cazemier et al.](#) pour stabiliser le modèle d'ordre réduit sur la durée d'échantillonnage puis dans un second temps, un calibrage des coefficients `KdL_E` est pratiqué puisque celui-ci avait permis d'améliorer considérablement la réponse sur une durée légèrement plus longue (voir le tableau 6.5 p.196). En ce qui concerne l'identification des coefficients, le problème inverse est résolu en approchant le pseudo-inverse avec  $p = 9$  valeurs singulières (voir le paragraphe 6.5.2 p.198).

#### F.2.1.1 Comparaison des coefficients autonomes

Les histogrammes des figures F.7, F.8 et F.9 représentent les coefficients des termes autonomes constants  $\mathcal{K}_i$ , linéaires  $\mathcal{L}_{ij}$  et quadratiques  $\mathcal{Q}_{ijk}$  respectivement. D'une manière générale, les valeurs des coefficients calculés par la résolution du problème d'identification sont beaucoup plus faibles que celles produites par les expressions analytiques. Par ailleurs les coefficients prépondérants diffèrent suivant la manière dont les coefficients sont évalués, et les signes sont souvent opposés pour les coefficients constants comme on peut l'observer sur la figure F.7.

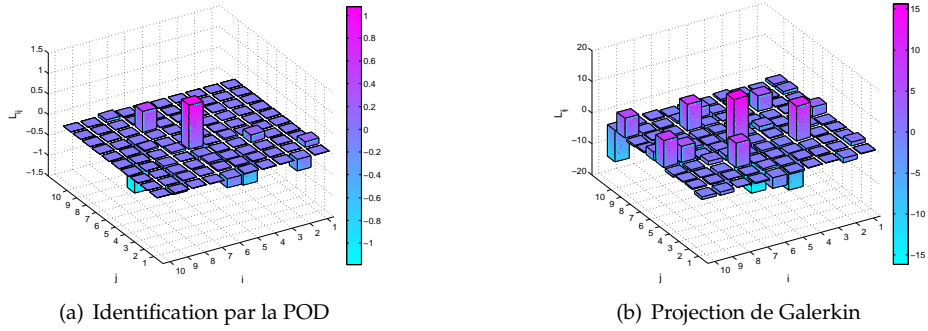
La comparaison des histogrammes des coefficients autonomes linéaires représentés sur la figure F.8 révèle que l'identification a tendance à aplatir l'histogramme. Les coefficients prépondérants sont moins nombreux que lorsque les expressions analytiques sont employées pour évaluer les coefficients, ce qui laisse penser que l'identification permet de lisser, ou de filtrer en quelque sorte l'influence de certains termes.

Ce comportement s'observe aussi pour les coefficients quadratiques représentés sur la figure F.9. Seuls quelques termes correspondant aux premiers indices  $i$  de l'opérateur  $\mathcal{Q}$  identifié contribuent de manière non négligeable à la réponse quadratique. L'effet est d'autant plus flagrant sur les tranches  $\mathcal{Q}_{ij5}$  et  $\mathcal{Q}_{ij10}$  où les contributions quadratiques sont quasiment nulles au delà de  $i = 3$ . Au contraire, les contributions des termes quadratiques évalués avec les expressions analytiques sont beaucoup plus réparties.

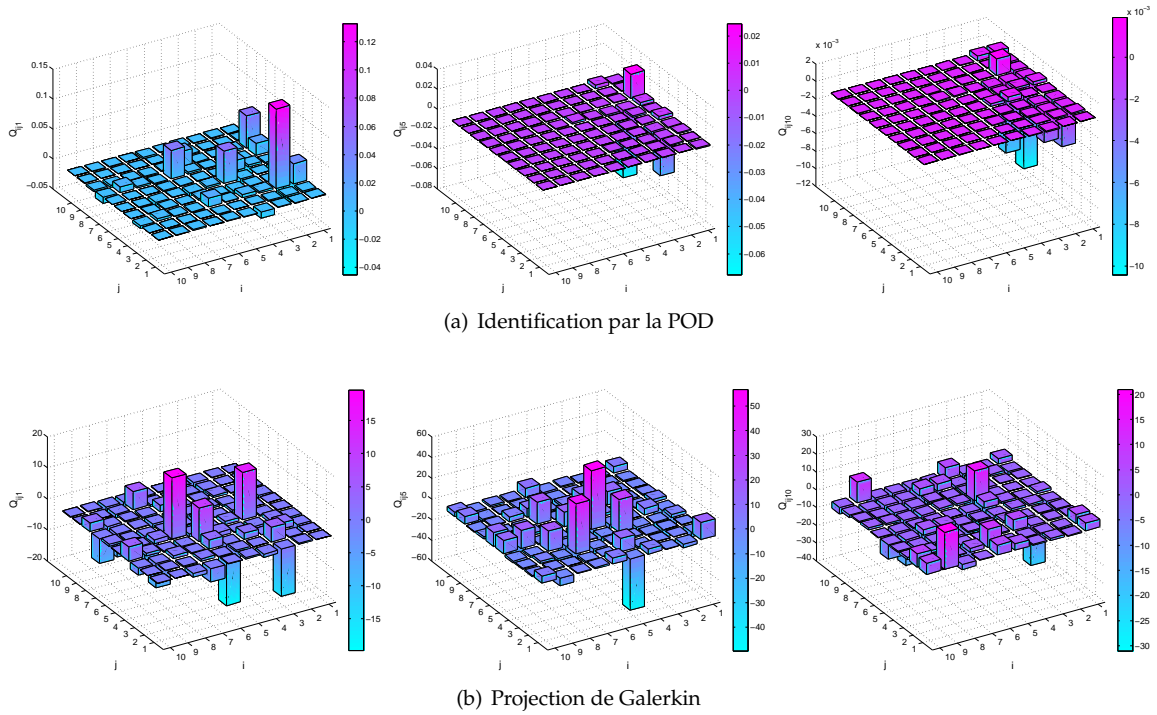


**Figure F.7 :** Comparaison des coefficients autonomes constants  $\mathcal{K}_i$  du modèle d'ordre réduit identifiés par la POD et calculés par les expressions analytiques issues de la projection de Galerkin.





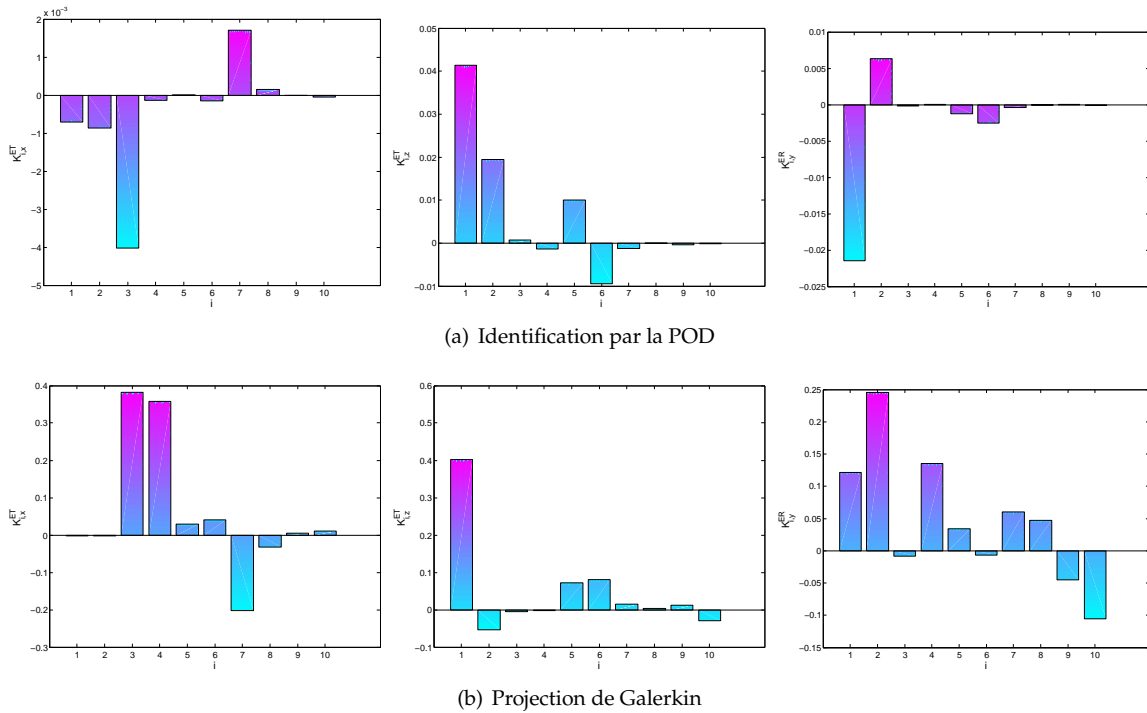
**Figure F.8 :** Comparaison des coefficients autonomes linéaires  $\mathcal{L}_{ij}$  du modèle d'ordre réduit identifiés par la POD et calculés par les expressions analytiques issues de la projection de Galerkin.



**Figure F.9 :** Comparaison des coefficients autonomes quadratiques du modèle d'ordre réduit identifiés par la POD et calculés par les expressions analytiques issues de la projection de Galerkin. Seules les trois tranches  $Q_{ij1}$ ,  $Q_{ij5}$  et  $Q_{ij10}$  de l'opérateur  $\mathcal{Q}$  sont représentées.

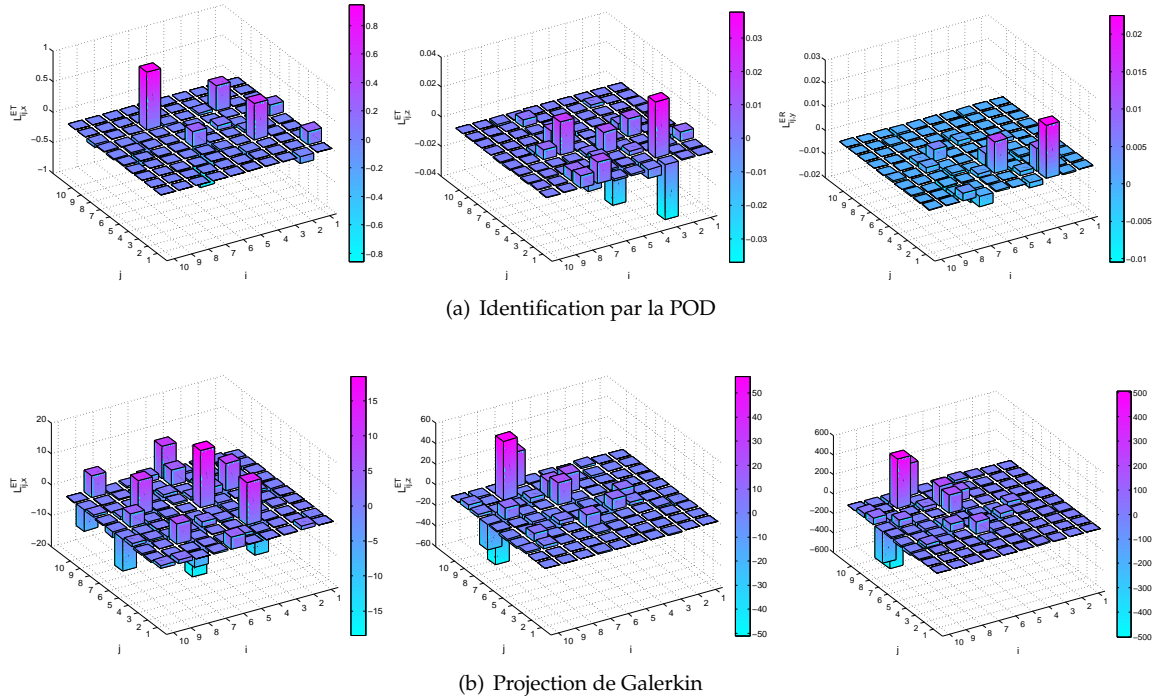
## F.2.1.2 Comparaison des coefficients non-autonomes

Il en va de même pour les coefficients non-autonomes définis à l'équation (6.70) pour lesquels seules six composantes sont non nulles. Les trois composantes constantes  $\mathcal{K}_{i,x}^{ET}$ ,  $\mathcal{K}_{i,z}^{ET}$  et  $\mathcal{K}_{i,y}^{ER}$  sont représentées sur la figure F.10 lorsqu'elles sont identifiées à partir de l'équation (2.65), ou lorsqu'elles sont calculées explicitement à partir des expressions analytiques. L'ordre de grandeur des coefficients identifiés est toujours beaucoup plus faible que dans le cas de coefficients calculés à partir des expressions de la projection de Galerkin et les signes diffèrent souvent. Le même type d'observation peut être formulé pour les coefficients linéaires non-autonomes représentés sur la figure F.11. Les coefficients prépondérants ne sont pas du tout localisés sur les mêmes indices. En outre, l'ordre de grandeur des coefficients calculés par les expressions analytiques prend une importance considérable par rapport à celui des coefficients identifiés, en particulier pour les coefficients linéaires associés à la rotation.



**Figure F.10 :** Comparaison des coefficients non-autonomes constants du modèle d'ordre réduit identifiés par la POD (a) et calculés par les expressions analytiques issues de la projection de Galerkin (b). Seules les composantes non nulles  $\mathcal{K}_{i,x}^{ET}$ ,  $\mathcal{K}_{i,z}^{ET}$  et  $\mathcal{K}_{i,y}^{ER}$  des opérateurs constants  $\mathcal{K}_i^{ET}$  et  $\mathcal{K}_i^{ER}$  sont représentées.

La conclusion de cette comparaison des coefficients calculés par les deux approches est que l'identification des coefficients permet en général de filtrer les erreurs commises dans l'évaluation des coefficients à partir des expressions analytiques de la projection de Galerkin. Le résultat est particulièrement visible pour les coefficients linéaires et quadratiques autonomes (voir les figures F.8 et F.9). L'identification des coefficients quadratiques conduit aussi à des termes prépondérants pour les premières amplitudes modales, ce qui n'est pas le cas lorsque les coefficients sont calculés au moyen des expressions analytiques. Les termes dominants dans les opérateurs du modèle d'ordre réduit diffèrent considérablement en fonction de l'approche choisie pour calculer les coefficients. Toutefois, la réponse globale du système est relativement proche lorsque la fenêtre d'échantillonnage  $I_e$  est considérée. Les modifications apportées par l'identification des coefficients concernent surtout le comportement à long terme du modèle d'ordre réduit. Il est cependant difficile de déterminer quels sont précisément les coefficients qui permettent de stabiliser la réponse à long terme.



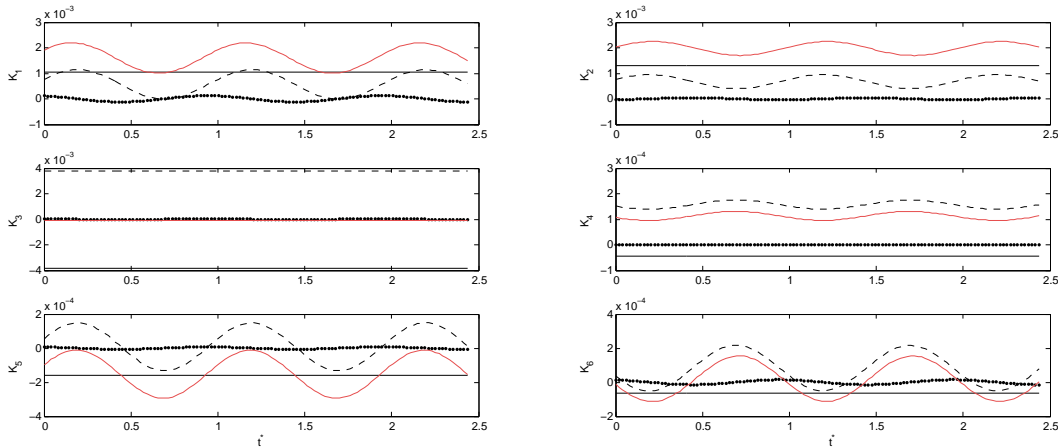
**Figure F.11 :** Comparaison des coefficients non-autonomes linéaires du modèle d'ordre réduit identifiés par la POD (a) et calculés par les expressions analytiques issues de la projection de Galerkin (b). Seules les composantes non nulles  $\mathcal{L}_{ij,x}^{ET}$ ,  $\mathcal{L}_{ij,z}^{ET}$  et  $\mathcal{L}_{ij,y}^{ER}$  des opérateurs linéaires  $\mathcal{L}_{ij}^{ET}$  et  $\mathcal{L}_{ij}^{ER}$  sont représentées.

## F.2.2 Comparaison des contributions autonomes et non-autonomes

Le modèle d'ordre réduit est ici construit à partir de  $M = 120$  clichés avec une base de projection contenant  $q = 6$  modes POD. Le problème inverse pour l'identification des coefficients est approché avec  $p = 7$  valeurs singulières puisque dans ce cas la réponse reste stable à très long terme (voir le tableau 6.7 p. 202). Dans l'exemple du profil NACA0012, les contributions convectives et diffusives, puis surfaciques et volumiques avaient été comparées afin d'analyser leur participation relative aux coefficients du modèle d'ordre réduit. Dans l'exemple traité ici, cette comparaison n'a pas lieu d'être puisque l'écoulement est supposé parfait et par conséquent les coefficients sont seulement donnés par les contributions convectives. Il est toutefois intéressant de distinguer cette fois les contributions autonomes et non-autonomes.

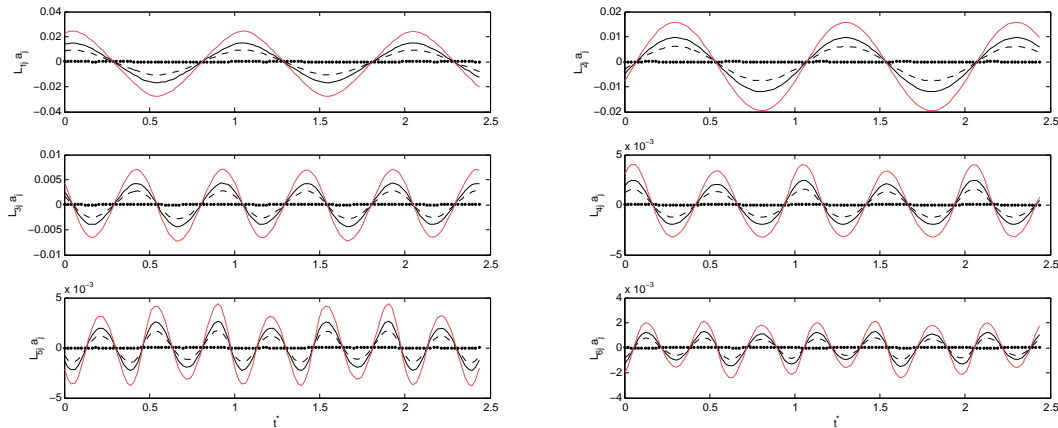
Les contributions des termes autonomes  $\mathcal{K}_i$  indépendantes du temps sont comparées à chaque instant  $t_m$  aux contributions  $\mathcal{K}_i^{ET} \cdot \mathbf{s}_{0,E}(t_m)$  et  $\mathcal{K}_i^{ER} \cdot \boldsymbol{\omega}_E(t_m)$  sur la figure F.12. Globalement, on constate que les contributions liées à la rotation sont relativement faibles en comparaison avec celles du terme autonome et de la contribution du mouvement de translation. Les contributions autonomes et non-autonomes issues de la translation sont en revanche du même ordre de grandeur et participent activement à la réponse totale due aux termes constants. Le comportement des différentes contributions varie en fonction des modes POD auxquels les équations sont associées : ainsi, pour les deux premières équations, la contribution autonome  $\mathcal{K}_i$  s'ajoute à celle produite par la translation  $\mathcal{K}_i^{ET} \cdot \mathbf{s}_{0,E}(t)$  puisque les deux contributions ont le même signe. En revanche pour les équations du modèle d'ordre réduit relatives aux indices  $i = 4, 5$  et  $6$ , c'est l'inverse qui se produit et les contributions autonomes atténuent les effets du mouvement d'entraînement en translation. La transition s'effectue pour l'équation d'indice  $i = 3$  où les contributions autonome et d'entraînement en translation sont pratiquement égales et opposées.

De la même manière, les coefficients autonomes  $\mathcal{L}_{ij}$  sont comparés aux coefficients non-autonomes  $\mathcal{L}_{ij}^{ET} \cdot \mathbf{s}_{0,E}(t_m)$  et  $\mathcal{L}_{ij}^{ER} \cdot \boldsymbol{\omega}_E(t_m)$ . Les graphiques de la figure F.13 représentent ainsi la contribution autonome  $\sum_{j=1}^q \mathcal{L}_{ij} a_j(t_m)$  à chaque instant d'échantillonnage, et les contributions non-autonomes en trans-



**Figure F.12** : Comparaison des contributions des coefficients constants du modèle d'ordre réduit calculés par identification sur l'intervalle de temps d'échantillonnage adimensionnel  $t^* \in [t_0; t_0 + T_e]/T_\alpha$ . Les termes autonomes  $\mathcal{K}_i$  tracés en trait plein (—) sont constants au cours du temps. Les contributions non-autonomes issues des termes de translation  $\mathcal{K}_i^{ET} \cdot \mathbf{s}_{0,E}(t_m)$  sont tracées en traits discontinus (– –) et celles issues des termes de rotation  $\mathcal{K}_i^{ER} \cdot \boldsymbol{\omega}_E(t_m)$  sont tracées en pointillés (· · ·). La résultante des termes constants est tracée en traits rouges (—).

lation  $\sum_{j=1}^q (\mathcal{L}_{ij}^{ET} \cdot \mathbf{s}_{0,E}(t_m)) a_j(t_m)$  et en rotation  $\sum_{j=1}^q (\mathcal{L}_{ij}^{ER} \cdot \boldsymbol{\omega}_E(t_m)) a_j(t_m)$ . Les contributions dues à la rotation sont encore une fois négligeables par rapport aux contributions autonomes et à celles dues à la translation. Pour toutes les amplitudes modales, on peut constater que les différentes contributions linéaires sont en phase. Par conséquent, la contribution due à la translation vient s'ajouter à celle du terme autonome dans la réponse du système.



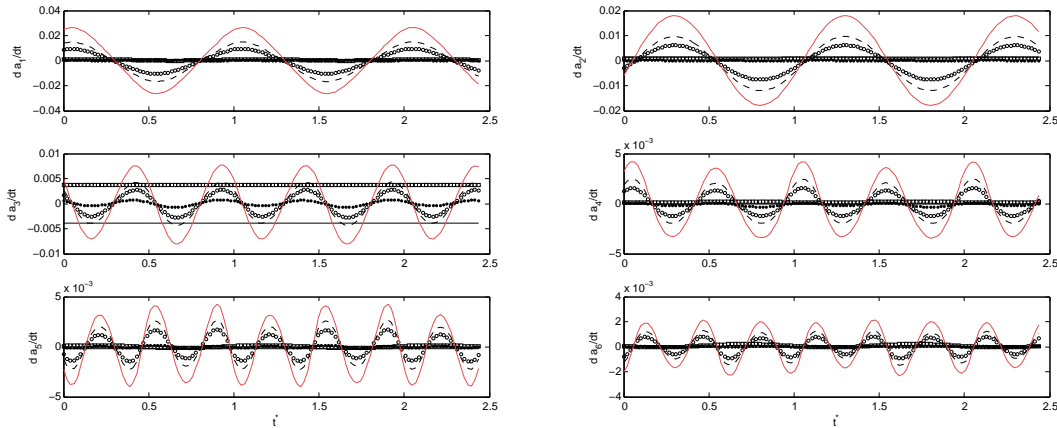
**Figure F.13** : Comparaison des contributions des coefficients linéaires du modèle d'ordre réduit calculés par identification sur l'intervalle de temps d'échantillonnage adimensionnel  $t^* \in [t_0; t_0 + T_e]/T_\alpha$ . Les contribution autonomes  $\sum_{j=1}^q \mathcal{L}_{ij} a_j(t_m)$  sont tracés en trait plein (—). Les contributions non-autonomes issues des termes de translation  $\sum_{j=1}^q (\mathcal{L}_{ij}^{ET} \cdot \mathbf{s}_{0,E}(t_m)) a_j(t_m)$  sont tracées en traits discontinus (– –) et celles issues des termes de rotation  $\sum_{j=1}^q (\mathcal{L}_{ij}^{ER} \cdot \boldsymbol{\omega}_E(t_m)) a_j(t_m)$  sont tracées en pointillés (· · ·). La résultante des termes linéaires est tracée en traits rouges (—).

Finalement, la figure F.14 nous renseigne sur la façon dont les contributions des termes autonomes et non-autonomes participent globalement à la réponse du modèle d'ordre réduit. Chaque graphique représente en trait plein rouge la dérivée  $\dot{a}_i$  des amplitudes modales qui correspond au membre de gauche de l'équation (6.14) gouvernant le modèle d'ordre réduit. D'après l'équation (6.14), cette quantité est la somme des contributions du terme constant  $\mathcal{K}$ , du terme linéaire  $\sum_{j=1}^q \mathcal{L}_{ij} a_j(t)$ , du terme quadratique  $\sum_{j,k=1}^q \mathcal{Q}_{ij} a_j(t) a_k(t)$ , du terme constant d'entraînement  $\mathcal{K}_i^E(t)$  et enfin du terme linéaire d'entraînement

$\sum_{j=1}^q \mathcal{L}_{ij}^E(t) a_j(t)$ . Les courbes des graphiques de la figure F.14 mettent en évidence le fait que les contributions linéaires autonomes et non-autonomes sont prépondérantes pour l'exemple traité ici, ce qui en soit n'est pas surprenant puisque les oscillations imposées au profil sont de faible amplitude et que la réponse aérodynamique est principalement linéaire. Toutefois, les contributions quadratiques ne sont pas négligeables : dès la troisième amplitude modale, les oscillations du terme quadratique sont en phase avec celles des contributions linéaires et participent donc à l'entretien des oscillations du système. Enfin, il s'avère que les contributions constantes sont négligeables : les graphiques de la figure F.12 avaient permis de montrer que l'ordre de grandeur des termes constants était très faible par rapport à ceux des termes linéaires représentés sur les graphiques de la figure F.13. Lorsqu'elle n'est pas négligeable comme c'est le cas pour  $\dot{a}_3$ , la contribution constante autonome s'annule avec celle produite par le terme constant d'entraînement, si bien que la contribution constante totale est bien négligeable. La contribution constante participe peu à la réponse du modèle d'ordre réduit puisque la décomposition des clichés pratiquée est une décomposition affine et que le modèle d'ordre réduit est formulé pour la partie fluctuante autour de la partie moyenne.

En conclusion, l'analyse de la contribution des différents coefficients a permis de mettre en évidence le fait que la réponse du modèle d'ordre réduit est principalement gouvernée par les termes linéaires, bien que les termes quadratiques ne soient pas négligeables pour les amplitudes  $a_{i,i \geq 3}$ . Par ailleurs, les termes non-autonomes participent principalement à la réponse du modèle d'ordre réduit via les termes d'entraînement en translation. Ceci s'explique par le fait que la vitesse de translation est beaucoup plus élevée que celle du mouvement de rotation puisque l'écoulement est transsonique. Toutefois, il faut mentionner que l'effet de la rotation intervient aussi dans le terme dit de « translation » puisque le vecteur vitesse  $s_{0,E}$  est exprimé dans le repère entraîné  $\mathcal{R}_E$  au moyen de la relation (E.8) dans laquelle apparaît la matrice de rotation  $\mathbf{R}$ .

La même analyse aurait pu être menée avec le modèle d'ordre réduit dont les coefficients sont évalués à partir des expressions analytiques. Les résultats, qui ne sont pas présentés ici, révèlent que la correction introduite dans le modèle d'ordre réduit POD-Galerkin a pour but de compenser entre elles les contributions autonomes et non-autonomes afin de retrouver la réponse macroscopique correcte. Ce résultat rejoint d'une certaine manière l'observation effectuée par Sirisup et Karniadakis [215] qui avaient dû modifier la valeur du facteur de pénalité pour imposer une condition limite non-homogène afin que la réponse globale du système soit correctement reproduite. Ainsi, il semble ici que la condition limite ne doit pas être imposée directement avec les valeurs de  $s_{0,E}$  et  $\omega_E$  mais avec des valeurs corrigées pour reproduire correctement la condition limite.



**Figure F.14 :** Participation des différentes contributions du modèle d'ordre réduit à la réponse. La dérivée  $\dot{a}_i$  des amplitudes modales est tracée pour  $i = 1, \dots, q$  en trait plein rouge (—). On trace ensuite les contributions du terme constant  $\mathcal{K}_i$  en trait plein (—), du terme linéaire  $\sum_{j=1}^q \mathcal{L}_{ij} a_j(t_m)$  en traits discontinus (---), du terme quadratique  $\sum_{j,k=1}^q \mathcal{Q}_{ij} a_j(t_m) a_k(t_m)$  en pointillés (⋯), du terme constant d'entraînement  $\mathcal{K}^E(t_m)$  avec les carrés (□) et enfin du terme linéaire d'entraînement  $\sum_{j=1}^q \mathcal{L}_{ij}^E(t_m) a_j(t_m)$  avec les cercles (○).

---

# Bibliographie

---

- [1] R.J. ADRIAN : Stochastic estimation of conditional structure : a review. *Applied Scientific Research*, Vol. **53**, No. 3–4, p. 291–303, 1994.
- [2] A. ADROVER et M. GIONA : Modal reduction of PDE models by means of Snapshot Archetypes. *Physica D*, Vol. **182**, No. 1–2, p. 23–45, 2003.
- [3] AGARD Manual vol.1 (1987). *Aeroelasticity in Axial-Flow Turbomachines : Unsteady Turbomachinery Aerodynamics*, 1987. AGARD-AG-297.
- [4] C.B. ALLEN, N.V. TAYLOR, C.L. FENWICK, A.L. GAITONDE et D.P. JONES : A comparison of full non-linear and reduced order aerodynamic models in control law design using a two-dimensional aerofoil model. *International Journal for Numerical Methods in Engineering*, Vol. **64**, No. 12, p. 1628–48, 2005.
- [5] C. ALLERY : *Contribution à l'identification des bifurcations et à l'étude des écoulements fluides par des systèmes dynamiques d'ordre faible (P.O.D.)*. Thèse de doctorat, Université de Poitiers, 2002.
- [6] C. ALLERY, C. BÉGHEIN et A. HAMDOUNI : Applying proper orthogonal decomposition to the computation of particle dispersion in a two-dimensional ventilated cavity. *Communications in Nonlinear Science and Numerical Simulation*, Vol. **10**, No. 8, p. 907–20, 2005.
- [7] C. ALLERY, S. GUERIN, A. HAMDOUNI et A. SAKOUT : Experimental and numerical POD study of the Coanda effect used to reduce self-sustained tones. *Mechanics Research Communications*, Vol. **31**, No. 1, p. 105–20, 2004.
- [8] M. AMABILI, A. SARKAR et M.P. PAÏDOUSSIS : Chaotic vibrations of circular cylindrical shells : Galerkin versus reduced-order models via the proper orthogonal decomposition method. *Journal of Sound and Vibration*, Vol. **290**, No. 3–5, p. 736–62, 2006.
- [9] A. AMMAR, B. MOKDAD, F. CHINESTA et R. KEUNINGS : A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids. *Journal of Non-Newtonian Fluid Mechanics*, Vol. **139**, No. 3, p. 153–76, 2006.
- [10] D. AMSALLEM, J. CORTIAL, K. CARLBERG et C. FARHAT : A method for interpolating on manifolds structural dynamics reduced-order models. *International Journal for Numerical Methods in Engineering*, Vol. **80**, No. 9, p. 1241–58, 2009.
- [11] A.C. ANTOULAS : *Approximation of Large-Scale Dynamical Systems*. Cambridge University Press, 2005.
- [12] J.S.R. ANTTONEN, P.I. KING et P.S. BERAN : POD-based reduced-order models with deforming grids. *Mathematical and Computer Modelling*, Vol. **38**, No. 1–2, p. 41–62, 2003.
- [13] J.S.R. ANTTONEN, P.I. KING et P.S. BERAN : Application of multi-POD to a pitching and plunging airfoil. *Mathematical and Computer Modelling*, Vol. **42**, No. 3–4, p. 245–59, 2005.
- [14] W. AQUINO : An object-oriented framework for reduced-order models using proper orthogonal decomposition (POD). *Computer Methods in Applied Mechanics and Engineering*, Vol. **196**, No. 41–44, p. 4375–90, 2007.

- [15] W.E. ARNOLDI : The principle of minimized iterations in the solution of the matrix eigenvalue problem. *Quarterly of Applied Mathematics*, Vol. **9**, p. 17–29, 1951.
- [16] N. AUBRY : On the hidden beauty of proper orthogonal decomposition. *Theoretical and Computational Fluid Dynamics*, Vol. **2**, p. 339–52, 1991.
- [17] N. AUBRY, R. GUYONNET et R. LIMA : Spatio-temporal symmetries and bifurcations via bi-orthogonal decompositions. *Journal of Nonlinear Science*, Vol. **2**, p. 183–215, 1992.
- [18] N. AUBRY, P. HOLMES, J.L. LUMLEY et E. STONE : The dynamics of coherent structures in the wall region of a turbulent boundary layer. *Journal of Fluid Mechanics*, Vol. **192**, p. 115–73, 1988.
- [19] M.F.A. AZEEZ et A.F. VAKAKIS : Proper orthogonal decomposition (POD) of a class of vibroimpact oscillations. *Journal of Sound and Vibration*, Vol. **240**, No. 5, p. 859–89, 2001.
- [20] Z. BAI : Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems. *Applied Numerical Mathematics*, Vol. **43**, No. 1–2, p. 9–44, 2002.
- [21] Z. BAI et R.W. FREUND : A partial Padé-Via-Lanczos method for reduced-order modeling. *Linear Algebra and its Applications*, Vol. **332–334**, p. 139–64, 2001.
- [22] O. BALIMA, Y. FAVENNEC, M. GIRAULT et D. PETIT : Comparison between the modal identification method and the POD-Galerkin method for model reduction in nonlinear diffusive systems. *International Journal for Numerical Methods in Engineering*, Vol. **67**, No. 7, p. 895–915, 2006.
- [23] M.F. BARONE, I. KALASHNIKOVA, D.J. SEGALMAN et H.K. THORNQUIST : Stable Galerkin reduced order models for linearized compressible flows. *Journal of Computational Physics*, Vol. **228**, No. 6, p. 1932–46, 2009.
- [24] G. BASTIN, S. BEAUGRAND, Ph. BEAUMIER, Ph. BENOIT, B. BERDE, G. BILLONET, J.-Ch. BONIFACE, M. BORREL, J.-F. BOUSSUGE, V. BRUNET, J. BRÉZILLON, H. BÉZARD, L. CAMBIER, B. CANTALOUBE, L. CASTILLON, S. CHAMPAGNEUX, Y. COLIN, R. COLLERCANDY, V. COUAILLIER, T. DARIS, M. GAZAIX, A. GAZAIX-JOLLÈS, V. GLEIZE, S. HEIB, R. HOUEVILLE, H.-P. JENNY, J.-C. JOUHAUD, X. JUVIGNY, M. LAZAREFF, M.-C. LE PAPE, P. LEMÉE, C. MARMIGNON, J. MAYEUR, B. MICHEL, M. MONTAGNAC, G. NGO BOUM, J. PETER, S. PLOT, A. ROURE, O. ROUZAUD, F. SICOT, M. THIERY, J.-P. VEUILLOT, A.-M. VUILLOT et X. de SAINT VICTOR : *elsA Theoretical Manual*. ONERA, 2.0 édition, 2008.
- [25] G.K. BATCHELOR : *An Introduction to Fluid Dynamics*. Cambridge University Press, 2000.
- [26] K.-J. BATHE : *Finite Element Procedures*. Prentice-Hall, 1996.
- [27] D.L. BELL et L. HE : Three-dimensionnal unsteady flow for an oscillating turbine blade and the influence of tip leakage. *ASME Journal of Turbomachinery*, Vol. **122**, No. 1, p. 93–101, 2000.
- [28] S. BELLIZZI et R. SAMPAIO : POMs analysis of randomly vibrating systems obtained from Karhunen-Loève expansion. *Journal of Sound and Vibration*, Vol. **297**, No. 3–5, p. 774–93, 2006.
- [29] S. BELLIZZI et R. SAMPAIO : Smooth Karhunen–Loève decomposition to analyze randomly vibrating systems. *Journal of Sound and Vibration*, Vol. **325**, No. 3, p. 491–8, 2009.
- [30] P.S. BERAN, D.J. LUCIA et C.L. PETTIT : Reduced-order modelling of limit-cycle oscillation for aeroelastic systems. *Journal of Fluids and Structures*, Vol. **19**, No. 5, p. 575–90, 2004.
- [31] M. BERGMANN et L. CORDIER : Optimal control of the cylinder wake in the laminar regime by trust-region methods and POD reduced-order models. *Journal of Computational Physics*, Vol. **227**, p. 7813–40, 2008.
- [32] G. BERKOOZ, P. HOLMES et J.L. LUMLEY : The proper orthogonal decomposition in the analysis of turbulent flows. *Annual Reviews of Fluid Mechanics*, Vol. **25**, p. 539–75, 1993.

- [33] F. BONNANS : A variant of a projected variable metric method for bound constrained optimization problems. Rapport de Recherche No. 242, INRIA, 1983.
- [34] R. BOURGUET, M. BRAZA, A. DERVIEUX et A. SÉVRAIN : Transition features in transonic flow around a NACA0012 airfoil by Navier-Stokes and low-order modelling. *In West-East High Speed Flow Field Conference*, Moscou (Russie), 19–22 Novembre 2007.
- [35] M. BUFFONI, S. CAMARRI, A. IOLLO et M.V. SALVETTI : Low-dimensional modelling of a confined three-dimensional wake flow. *Journal of Fluid Mechanics*, Vol. **569**, p. 141–50, 2006.
- [36] T. BUI-THANH, K. WILLCOX et O. GHATTAS : Model reduction for large-scale systems with high-dimensional parametric input space. *SIAM Journal on Scientific Computing*, Vol. **30**, No. 6, p. 3270–88, 2008.
- [37] J. BURKARDT, M. GUNZBURGER et H.-C. LEE : POD and CVT-based reduced-order modeling of Navier-Stokes flows. *Computer Methods in Applied Mechanics and Engineering*, Vol. **196**, No. 1–3, p. 337–55, 2006.
- [38] V. CARSTENS, R. KEMME et S. SCHMITT : Coupled simulation of flow-structure interaction in turbomachinery. *Aerospace Science and Technology*, Vol. **7**, No. 4, p. 298–306, 2003.
- [39] F. CASADEI, J. P. HALLEUX, A. SALA et F. CHILLÉ : Transient fluid–structure interaction algorithms for large industrial applications. *Computer Methods in Applied Mechanics and Engineering*, Vol. **190**, No. 24–25, p. 3081–110, 2001.
- [40] L. CAVAGNA, G. QUARANTA et P. MANTEGAZZA : Application of Navier-Stokes simulations for aeroelastic stability assessment in transonic regimes. *Computers & Structures*, Vol. **85**, No. 11–14, p. 818–32, 2007.
- [41] W. CAZEMIER : *Proper Orthogonal Decomposition and low dimensional models for turbulent flows*. Thèse de doctorat, Université de Groningen, Hollande, 1997.
- [42] W. CAZEMIER, R.W. VERSTAPPEN et A.E.P. VELDMAN : Proper orthogonal decomposition and low dimensional models for driven cavity flows. *Physics of Fluids*, Vol. **10**, No. 7, p. 1685–99, 1998.
- [43] Y. CHAHLAOUI, K. GALLIVAN et P. VAN DOOREN : Recursive calculation of dominant singular subspaces. *SIAM Journal on Matrix Analysis and Applications*, Vol. **25**, No. 2, p. 445–63, 2003.
- [44] D. CHELIDZE et W. ZHOU : Smooth orthogonal decomposition-based vibration mode identification. *Journal of Sound and Vibration*, Vol. **292**, No. 3–5, p. 461–73, 2006.
- [45] F. CHINESTA, A. AMMAR, F. LEMARCHAND, P. BEAUCHENE et F. BOUST : Alleviating mesh constraints : Model reduction, parallel time integration and high resolution homogenization. *Computer Methods in Applied Mechanics and Engineering*, Vol. **197**, No. 5, p. 400–13, 2008.
- [46] E.A. CHRISTENSEN, M. BRØNS et J.N. SØRENSEN : Evaluation of proper orthogonal decomposition-based decomposition techniques applied to parameter-dependent nonturbulent flows. *SIAM Journal on Scientific Computing*, Vol. **21**, No. 4, p. 1419–34, 2000.
- [47] P.G.A. CIZMAS et A. PALACIOS : Proper orthogonal decomposition of turbine rotor-stator interaction. *AIAA Journal of Propulsion and Power*, Vol. **19**, No. 2, p. 268–81, 2003.
- [48] L. CORDIER, B. ABOU EL MADJ et J. FAVIER : Calibration of POD reduced-order models using Tikhonov regularization. *International Journal for Numerical Methods in Fluids*, 2009. In Press, doi 10.1002/flid.2074.
- [49] L. CORDIER et M. BERGMANN : Réduction de dynamique par Décomposition Orthogonale aux Valeurs Propres (POD). *In École de Printemps Optimisation et Contrôle des Écoulements et des Transferts*, Aussois (France), 12–17 mars 2006.



## BIBLIOGRAPHIE

---

- [50] M. COUPLET : *Modélisation POD-Galerkine réduite pour le contrôle des écoulements instationnaires*. Thèse de doctorat, Université de Paris 13, 2005.
- [51] M. COUPLET, C. BASDEVANT et P. SAGAUT : Calibrated reduced-order POD-Galerkin system for fluid flow modelling. *Journal of Computational Physics*, Vol. **207**, No. 1, p. 192–220, 2005.
- [52] M. COUPLET, P. SAGAUT et C. BASDEVANT : Intermodal energy transfers in a proper orthogonal decomposition-Galerkin representation of a turbulent separated flows. *Journal of Fluid Mechanics*, Vol. **491**, p. 275–84, 2003.
- [53] R.R. Jr CRAIG et M.C.C. BAMPTON : Coupling of substructures for dynamic analysis. *AIAA Journal*, Vol. **6**, No. 7, p. 1313–9, 1968.
- [54] A. CUTLER et L. BREIMAN : Archetypal analysis. *Technometrics*, Vol. **36**, No. 4, p. 338–47, 1994.
- [55] F. DE VUYST et C. AUDOUZE : *Optimisation multidisciplinaire en mécanique 2 : réduction de modèles, robustesse, fiabilité, réalisations logicielles*, chapitre 2 Réduction de modèles Éléments Finis par POD pour les problèmes paramétrés aux EDP. Hermès Sciences, 2009.
- [56] P.G. del SASTRE et R. BERMEJO : Error estimates of proper orthogonal decomposition eigenvectors and Galerkin projection for a general dynamical system arising in fluid models. *Numerische Mathematik*, Vol. **110**, No. 1, p. 49–81, 2008.
- [57] J. DONEA, A. HUERTA, J.-Ph. PONTHOT et A. RODRÌGUEZ-FERRAN : *Encyclopedia of Computational Mechanics*, volume 1 : Fundamentals, chapitre 14 : Arbitrary Lagrangian-Eulerian Methods. John Wiley & Sons, 2004.
- [58] E.H. DOWELL, E.F. CRAWLEY, H.C. CURTISS, D.A. PETERS, R.H. SCANLAN et F. SISTO : *A Modern Course in Aeroelasticity*. Kluwer Academic, 3e édition, 1995.
- [59] E.H. DOWELL, K.C. HALL et M.C. ROMANOWSKI : Eigenmode analysis in unsteady aerodynamics : Reduced order models. *Applied Mechanics Review*, Vol. **50**, No. 6, p. 371–86, 1997.
- [60] E.H. DOWELL, K.C. HALL, J.P. THOMAS, R. FLOREA, B.I. EPUREANU et J. HEGG : Reduced order models in unsteady aerodynamics. *In Structures, Structural Dynamics and Materials Conference*, 1999.
- [61] G. DREYFUS, J.-M. MARTINEZ, M. SAMUELIDES, M. GORDON, F. BADRAN, S. THIRIA et L. HÉRAULT : *Réseaux de neurones, méthodologie et applications*. Eyrolles, 2002.
- [62] Q. DU, V. FABER et M. GUNZBURGER : Centroidal Voronoi tessellations : Applications and algorithms. *SIAM Review*, Vol. **41**, No. 4, p. 637–76, 1999.
- [63] Q. DU, M. GUNZBURGER et L. JU : Voronoi-based finite volume methods, optimal Voronoi meshes, and PDEs on the sphere. *Computer Methods in Applied Mechanics and Engineering*, Vol. **192**, No. 35–36, p. 3933–57, 2003.
- [64] A. DUGEAI : Aeroelastic developments in the elsA code and unsteady RANS applications. *In International Forum on Aeroelasticity and Structural Dynamics*, Munich (Allemagne), 28 Juin – 1er Juillet 2005.
- [65] A. DUGEAI : Nonlinear numerical aeroelasticity with the ONERA elsA solver. *In AVT-152 Limit-Cycle Oscillations and Other Amplitude-Limited, Self-Excited Vibrations*, Loen (Norvège), 5–8 Mai 2008.
- [66] C. ECKART et G. YOUNG : The approximation of one matrix by another of lower rank. *Psychometrika*, Vol. **1**, No. 3, p. 211–8, 1936.
- [67] M. EL-KAMALI, J.-S. SCHOTTÉ et R. OHAYON : Computation of the free surface of a liquid with surface tension for any tank geometry. *In International Conference on Noise and Vibration Engineering*, Louvain (Belgique), 15–17 Septembre 2008.

- [68] B.I. EPUREANU, E.H. DOWELL et K.C. HALL : Reduced-order models of unsteady transonic viscous flows in turbomachinery. *Journal of Fluids and Structures*, Vol. **14**, No. 8, p. 1215–34, 2000.
- [69] L.-E. ERIKSSON : Practical three-dimensional mesh generation using transfinite interpolation. *SIAM Journal on Scientific and Statistical Computing*, Vol. **6**, No. 3, p. 712–41, 1985.
- [70] M. FAHL : Computation of POD basis functions for fluid flows with Lanczos method. *Mathematical and Computer Modelling*, Vol. **34**, No. 1–2, p. 91–107, 2001.
- [71] C. FARHAT, M. LESOINNE et P. LE TALLEC : Load and motion transfer algorithms for fluid/structure interaction problems with non-matching discrete interfaces : Momentum and energy conservation, optimal discretization and application to aeroelasticity. *Computer Methods in Applied Mechanics and Engineering*, Vol. **157**, No. 1–2, p. 95–114, 1998.
- [72] C. FARHAT, M. LESOINNE et P. STERN : High performance solution of three-dimensional nonlinear aeroelastic problems via parallel partitioned algorithms : Methodology and preliminary results. *Advances in Engineering Software*, Vol. **28**, No. 1, p. 43–61, 1997.
- [73] J. FAVIER, A. KOURTA et L. CORDIER : Accurate POD reduced-order models of separated flows. In *IUTAM Symposium on Unsteady Separated Flows and their Control*, Corfu (Grèce), 2007.
- [74] B.F. FEENY et R. KAPPAGANTU : On the physical interpretation of proper orthogonal modes in vibrations. *Journal of Sound and Vibration*, Vol. **211**, No. 4, p. 607–16, 1998.
- [75] Z. FENG et A. SOULAIMANI : Nonlinear aeroelasticity modeling using a reduced order model based on proper orthogonal decomposition. In *ASME Pressure Vessels and Piping Division Conference*, No. 26006, San Antonio (USA), 22–26 juillet 2007.
- [76] J.H. FERZIGER et M. PERIĆ : *Computational Methods for Fluid Dynamics*. Springer, 1996.
- [77] R. FLETCHER : *Practical Methods of Optimization*. Wiley, 2003.
- [78] R. FLOREA et K.C. HALL : Eigenmode analysis of unsteady flows about airfoils. *Journal of Computational Physics*, Vol. **147**, No. 2, p. 568–93, 1998.
- [79] A.I.J. FORRESTER et A.J. KEANE : Recent advances in surrogate-based optimization. *Progress in Aerospace Science*, Vol. **45**, No. 1–3, p. 50–79, 2009.
- [80] R.W. FREUND : Model reduction methods based on Krylov subspaces. Numerical Analysis Manuscript 03-4-01, Bell Laboratories, 2003. <http://cm.bell-labs.com/cs/doc/03>.
- [81] B. GALLETTI, A. BOTTARO, C.H. BRUNEAU et A. IOLLO : Accurate model reduction of transient and forced wakes. *European Journal of Mechanics - B/Fluids*, Vol. **26**, No. 3, p. 354–66, 2007.
- [82] B. GALLETTI, C.H. BRUNEAU, L. ZANNETTI et A. IOLLO : Low-order modelling of laminar flow regimes past a confined square cylinder. *Journal of Fluid Mechanics*, Vol. **503**, p. 161–70, 2004.
- [83] M. GÉRADIN et D. RIXEN : *Théorie des Vibrations*. Masson, 1992.
- [84] Z. GETE et EVANS : An experimental investigation of unsteady turbulent-wake / boundary-layer interaction. *Journal of Fluids and Structures*, Vol. **17**, No. 1, p. 43–55, 2003.
- [85] P. GIRODROUX-LAVIGNE et A. DUGEAI : Transonic aeroelastic computations using Navier-Stokes equations. In *International Forum on Aeroelasticity and Structural Dynamics*, Amsterdam (Pays-Bas), 4–6 Juin 2003.
- [86] P. GIRODROUX-LAVIGNE et A. DUGEAI : Fluid-structure coupling with elsA : status of developments and examples of application. In *42<sup>e</sup> Colloque d'Aérodynamique Appliquée*, Nice (France), 19–21 Mars 2007.

- [87] X. GLOERFELT : Compressible POD-Galerkin reduced-order model of self-sustained oscillations in a cavity. *AIAA Journal*, Vol. **1**, p. 417–44, 2006.
- [88] R. GLOWINSKI, T.-W. PAN, T.I. HESLA, D. D. JOSEPH et J. PÉRIAUX : A distributed Lagrange multiplier/fictitious domain method for the simulation of flow around moving rigid bodies : Application to particulate flow. *Computer Methods in Applied Mechanics and Engineering*, Vol. **184**, No. 2–4, p. 241–67, 2000.
- [89] R. GLOWINSKI, T.-W. PAN et J. PÉRIAUX : Distributed Lagrange multiplier methods for incompressible viscous flow around moving rigid bodies. *Computer Methods in Applied Mechanics and Engineering*, Vol. **151**, No. 1–2, p. 181–94, 1998.
- [90] M. GLÜCK, M. BREUER, F. DURST, A. HALFMANN et E. RANK : Computation of wind-induced vibrations of flexible shells and membranous structures. *Journal of Fluids and Structures*, Vol. **17**, No. 5, p. 739–65, 2003.
- [91] G. GOLUB et W. KAHAN : Calculating the singular values and pseudo-inverse of a matrix. *Journal of the Society for Industrial and Applied Mathematics, Series B : Numerical Analysis*, Vol. **2**, No. 2, p. 205–24, 1965.
- [92] G. GOLUB et C. REINSCH : Singular value decomposition and least squares solutions. *Numerische Mathematik*, Vol. **14**, No. 5, p. 403–20, 1970.
- [93] G.H. GOLUB et C.F. VAN LOAN : *Matrix Computations*. Johns Hopkins University Press, 3e édition, 1996.
- [94] A. K. GOPINATH et A. JAMESON : Application of the Time Spectral Method to periodic unsteady vortex shedding. In *40<sup>th</sup> AIAA Aerospace Sciences Meeting and Exhibit*, Reno (USA), 9–12 Janvier 2006.
- [95] A.K. GOPINATH, E. Van der WEIDE, J.J. ALONSO, A. JAMESON, K. EKICI et K.C. HALL : Three-dimensional unsteady multi-stage turbomachinery simulations using the harmonic balance technique. In *45<sup>th</sup> AIAA Aerospace Sciences Meeting and Exhibit*, No. 892, Reno (USA), 8–11 janvier 2007.
- [96] N. GOURDAIN, S. BURGUBURU, F. LEOEUF et H. MITON : Numerical simulation of rotating stall in a subsonic compressor. *Aerospace Science and Technology*, Vol. **10**, No. 1, p. 9–18, 2006.
- [97] B. GRÜBER et V. CARSTENS : The impact of viscous effects on the aerodynamic damping of vibrating transonic compressor blades - A numerical study. *ASME Journal of Turbomachinery*, Vol. **123**, No. 2, p. 409–17, 2001.
- [98] J. GUCKENHEIMER et P. HOLMES : *Nonlinear oscillations, dynamical systems, and bifurcation of vector fields*. Springer Verlag, 1993.
- [99] T. GUDMUNDSSON et A.J. LAUB : Approximate solution of large sparse Lyapunov equations. *IEEE Transactions on Automatic Control*, Vol. **39**, No. 5, p. 1110–4, 1994.
- [100] S. GUGERCIN : An iterative SVD-Krylov based method for model reduction of large-scale dynamical systems. *Linear Algebra and its Applications*, Vol. **428**, No. 8–9, p. 1964–86, 2008.
- [101] M.D. GUNZBURGER, J.S. PETERSON et J.N. SHADID : Reduced-order modeling of time-dependent PDEs with multiple parameters in the boundary data. *Computer Methods in Applied Mechanics and Engineering*, Vol. **196**, No. 4–6, p. 1030–47, 2007.
- [102] G.P. GURUSWAMY et C. BUYN : Direct coupling of Euler flow equations with plate finite element structures. *AIAA Journal*, Vol. **33**, No. 2, p. 375–7, 1994.
- [103] K.C. HALL, J.P. THOMAS et W.S. CLARK : Computation of unsteady nonlinear flows in cascades using a harmonic balance technique. *AIAA Journal*, Vol. **40**, No. 5, p. 879–86, 2002.

- 
- [104] K.C. HALL, J.P. THOMAS et E.H. DOWELL : Reduced-order modelling of unsteady small-disturbance flows using a frequency-domain proper orthogonal decomposition technique. *In 37<sup>th</sup> Aerospace Sciences Meeting and Exhibit*, Reno (Nevada), 1999.
- [105] P.C. HANSEN : Truncated singular value decomposition solutions to discrete ill-posed problems with ill-determined numerical rank. *SIAM Journal on Scientific and Statistical Computing*, Vol. **11**, No. 3, p. 503–18, 1990.
- [106] P.C. HANSEN : Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Review*, Vol. **34**, No. 4, p. 561–80, 1992.
- [107] K. HASSELMANN : PIPs and POPs : The reduction of complex dynamical systems using principal interaction and oscillation patterns. *Journal of Geophysical Research*, Vol. **93**, No. D9, p. 11015–21, 1988.
- [108] H.J. HASSIG : An approximate true damping solution of the flutter equation by determinant iteration. *Journal of Aircraft*, Vol. **8**, No. 11, p. 885–9, 1971.
- [109] P. HÉMON et F. SANTI : Applications of biorthogonal decompositions in fluid-structure interactions. *Journal of Fluids and Structures*, Vol. **17**, No. 8, p. 1123–43, 2003.
- [110] T. HENRI et J.-P. YVON : *System Modeling and Optimization*, volume 166 de *IFIP International Federation for Information Processing*, chapitre Convergence Estimates of POD-Galerkin Methods for Parabolic Problems, pages 295–306. Springer Boston, 2005.
- [111] A.C. HINDMARSH : ODEPACK, a systematized collection of ODE solvers. *IMACS Transactions on Scientific Computation*, Vol. **1**, p. 55–64, 1983.
- [112] P. HOLMES, J.L. LUMLEY et G. BERKOOZ : *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge University Press, 1996.
- [113] H. HOTELLING : Analysis of a complex statistical variables into principal components. *Journal of Educational Psychology*, Vol. **24**, p. 417–41, 1933.
- [114] W.C. HURTY : Dynamic analysis of structural systems using component mode. *AIAA Journal*, Vol. **3**, No. 4, p. 678–85, 1965.
- [115] M. ILAK et C.W. ROWLEY : Reduced-order modeling of channel flow using traveling POD and balanced POD. *In 3<sup>rd</sup> AIAA Flow Control Conference*, No. 3194, San Francisco (USA), 5–8 juin 2006.
- [116] A. IOLLO : Remarks on the approximation of the Euler equations by a low order model. Rapport de Recherche No. 3329, INRIA, 1997.
- [117] A. IOLLO, S. LANTERI et J.-A. DÉSIDÉRI : Stability properties of POD-Galerkin approximations for the compressible Navier-Stokes equations. *Theoretical and Computational Fluid Dynamics*, Vol. **13**, No. 6, p. 377–96, 2000.
- [118] K. ITO et S.S. RAVINDRAN : A reduced-order method for simulation and control of fluid flows. *Journal of Computational Physics*, Vol. **143**, No. 2, p. 403–25, 1998.
- [119] I.M. JAÏMOUKHA et E.M. KASENALLY : Krylov subspace methods for solving large Lyapunov equations. *SIAM Journal on Numerical Analysis*, Vol. **31**, No. 1, p. 227–51, 1994.
- [120] A. JAMESON : Time dependent calculations using multigrid, with application to unsteady flows past airfoils and wings. *In 10<sup>th</sup> Computational Fluid Dynamics Conference*, No. 1596, Honolulu (Hawaï), 24–26 Juin 1991.
- [121] D.A. JONES, L.G. MARGOLIN et E.S. TITI : On the effectiveness of the approximate inertial manifold — a computational study. *Theoretical and Computational Fluid Dynamics*, Vol. **7**, p. 243–60, 1995.

## BIBLIOGRAPHIE

---

- [122] B.H. JØRGENSEN, J.N. SØRENSEN et M. BRØNS : Low-dimensional modeling of a driven cavity flow with two free parameters. *Theoretical and Computational Fluid Dynamics*, Vol. **16**, No. 4, p. 299–317, 2003.
- [123] J.-N. JUANG et R.S. PAPPAS : An eigensystem realization algorithm for modal parameter identification and model reduction. *Journal of Guidance, Control, and Dynamics*, Vol. **8**, No. 5, p. 620–7, 1985.
- [124] V.L. KALB et A.E. DEANE : An intrinsic stabilization scheme for proper orthogonal decomposition based low-dimensional models. *Physics of Fluids*, Vol. **19**, No. 5, p. 054106, 2007.
- [125] R. KAMAKOTI et W. SHYY : Fluid-structure interaction for aeroelastic applications. *Progress in Aerospace Science*, Vol. **40**, No. 8, p. 535–58, 2004.
- [126] K. KARHUNEN : Zur Spektraltheorie stochastischer Prozesse. *Annales Academiae Scientiarum*, Vol. **37**, 1946. Série A1.
- [127] M. KARPEL : Design for active flutter suppression and gust alleviation using state-space aeroelastic modeling. *AIAA Journal of Aircraft*, Vol. **19**, p. 221–7, 1982.
- [128] M. KARPEL : Time-domain aeroservoelastic modeling using weighted unsteady aerodynamic forces. *Journal of Guidance, Control, and Dynamics*, Vol. **13**, No. 1, p. 30–7, 1990.
- [129] G. KERSCHEN et J.-C. GOLINVAL : Physical interpretation of the proper orthogonal modes using the singular value decomposition. *Journal of Sound and Vibration*, Vol. **249**, No. 5, p. 849–65, 2002.
- [130] G. KERSCHEN, J.-C. GOLINVAL, A.F. VAKAKIS et L.A. BERGMAN : The method of proper orthogonal decomposition for dynamical characterization and order reduction of mechanical systems : An overview. *Nonlinear Dynamics*, Vol. **41**, p. 147–69, 2005.
- [131] T. KIM : Frequency-domain Karhunen-Loève method and its application to linear dynamic system. *AIAA Journal*, Vol. **36**, No. 11, p. 2117–23, 1998.
- [132] M. KIRBY : Minimal dynamical systems from PDEs using Sobolev eigenfunctions. *Physica D*, Vol. **57**, No. 3–4, p. 466–75, 1992.
- [133] K.A. KOUSEN et O.O. BENDIKSEN : Nonlinear aspects of the transonic aeroelastic stability problem. In *29<sup>th</sup> Structures, Structural Dynamics and Material Conference*, No. 88-2306, pages 760–9, Denver (USA), 1988.
- [134] D. KRIGE : A statistical approach to some basic mine valuation problems on the Witwater-strand. *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, Vol. **52**, No. 6, p. 119–39, 1951.
- [135] K. KUNISCH et S. VOLKWEIN : Galerkin proper orthogonal decomposition methods for parabolic problems. *Numerische Mathematik*, Vol. **90**, No. 1, p. 117–48, 2001.
- [136] K. KUNISCH et S. VOLKWEIN : Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. *SIAM Journal on Numerical Analysis*, Vol. **40**, No. 2, p. 492–515, 2002.
- [137] F. KWASNIOK : The reduction of complex dynamical systems using principal interaction patterns. *Physica D*, Vol. **92**, No. 1–2, p. 28–60, 1996.
- [138] M. LALANNE, P. BERTHIER et J. der HAGOPIAN : *Mécanique des vibrations linéaires*. Masson, 1980.
- [139] S. LALL, J.E. MARSDEN et S. GLAVAŠKI : A subspace approach to balanced truncation for model reduction of nonlinear control systems. *International Journal of Robust and Nonlinear Control*, Vol. **12**, No. 6, p. 519–35, 2002.
- [140] C. LANCZOS : An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *Journal of the National Bureau of Standards*, Vol. **45**, No. 4, p. 255–82, 1950.

- [141] T. LE GARREC et X. GLOERFELT : Simulation directe du bruit d'un profil en incidence. In *18<sup>e</sup> Congrès Français de Mécanique*, Grenoble, France, 2007.
- [142] P. LE TALLEC et J. MOURO : Fluid structure interaction with large structural displacements. *Computer Methods in Applied Mechanics and Engineering*, Vol. **190**, No. 24–25, p. 3039–67, 2001.
- [143] K. LEVENBERG : A method for the solution of certain problems in least squares. *Quarterly of Applied Mathematics*, Vol. **2**, p. 164–68, 1944.
- [144] Y.C. LIANG, H.P. LEE, S.P. LIM, W.Z. LIN, K.H. LEE et C.G. WU : Proper orthogonal decomposition and its applications - Part I : Theory. *Journal of Sound and Vibration*, Vol. **252**, No. 3, p. 527–44, 2002.
- [145] E. LIBERGE : *Réduction de modèle par POD-Galerkin pour les problèmes d'interaction fluide-structure*. Thèse de doctorat, Université de La Rochelle, 2008.
- [146] E. LIBERGE et A. HAMDOUNI : Reduced order modelling method via proper orthogonal decomposition (POD) for flow around an oscillating cylinder. *Journal of Fluids and Structures*, Vol. **26**, No. 2, p. 292–311, 2010.
- [147] T. LIEU et C. FARHAT : Adaptation of aeroelastic reduced-order models and application to an F-16 configuration. *AIAA Journal*, Vol. **45**, No. 6, p. 1244–57, 2007.
- [148] T. LIEU, C. FARHAT et M. LESOINNE : Reduced-order fluid/structure modeling of a complete aircraft configuration. *Computer Methods in Applied Mechanics and Engineering*, Vol. **195**, No. 41–43, p. 5730–42, 2006.
- [149] L. LIU, E.H. DOWELL et J.P. THOMAS : A high dimensional harmonic balance approach for an aeroelastic airfoil with cubic restoring forces. *Journal of Fluids and Structures*, Vol. **23**, No. 3, p. 351–63, 2006.
- [150] M. LOÈVE : Fonctions aléatoires du second ordre. *Comptes-Rendus de l'Académie des Sciences*, Vol. **220**, 1945.
- [151] E.N. LORENZ : Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, Vol. **20**, No. 2, p. 130–41, 1963.
- [152] D.J. LUCIA et P.S. BERAN : Projection methods for reduced order models of compressible flows. *Journal of Computational Physics*, Vol. **188**, p. 252–80, 2003.
- [153] D.J. LUCIA, P.S. BERAN et W.A. SILVA : Reduced-order modeling : New approaches for computational physics. *Progress in Aerospace Science*, Vol. **40**, p. 51–117, 2004.
- [154] D.J. LUCIA, P.S. BERAN et W.A. SILVA : Aeroelastic system development using proper orthogonal decomposition and Volterra theory. *AIAA Journal of Aircraft*, Vol. **42**, No. 2, p. 509–18, 2005.
- [155] D.J. LUCIA, P.I. KING et P.S. BERAN : Reduced order modeling of a two-dimensional flow with moving shocks. *Computers & Fluids*, Vol. **32**, No. 7, p. 917–38, 2003.
- [156] J.L. LUMLEY : *The Structures of Inhomogeneous Turbulent Flow*, pages 166–78. In : *Atmospheric Turbulence and Radio Wave Propagation*, edited by A.M. Yaglom and V.I. Tatarski, Nauka, Moscow édition, 1967.
- [157] A.J. MAHAJAN, E.H. DOWELL et D.B. BLISS : On the role of artificial viscosity in Euler and Navier-Stokes solvers. *AIAA Journal*, Vol. **29**, No. 4, p. 555–9, 1991.
- [158] M. MARION et R. TEMAM : Nonlinear Galerkin methods. *SIAM Journal on Numerical Analysis*, Vol. **26**, No. 5, p. 1139–57, 1989.
- [159] D. MARQUARDT : An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics*, Vol. **11**, p. 431–41, 1963.

- [160] O. MARSDEN, C. BOGEY et C. BAILLY : Calcul direct du bruit aérodynamique rayonné par des géométries complexes. In *17<sup>e</sup> Congrès Français de Mécanique*, Troyes (France), 29 août – 2 septembre 2005.
- [161] J.G. MARSHALL et M. IMREGUN : A review of aeroelasticity methods with emphasis on turbomachinery applications. *Journal of Fluids and Structures*, Vol. **10**, No. 3, p. 237–67, 1996.
- [162] N. MASTRONARDI, M. VAN BAREL et R. VANDEBRIL : A fast algorithm for the recursive calculation of dominant singular subspaces. *Journal of Computational and Applied Mathematics*, Vol. **218**, No. 2, p. 238–46, 2007.
- [163] G. MATHERON : *Les variables régionalisées et leur estimation*. Masson, 1965.
- [164] M. MCMULLEN, A. JAMESON et J. ALONSO : Acceleration of convergence to a periodic steady state in turbomachinery flows. In *39<sup>th</sup> Aerospace Sciences Meeting and Exhibit*, No. 0152, Reno (USA), 8–11 Janvier 2001.
- [165] M. MCMULLEN, A. JAMESON et J. ALONSO : Application of a non-linear frequency domain solver to the Euler and Navier-Stokes equations. In *40<sup>th</sup> AIAA Aerospace Sciences Meeting and Exhibit*, No. 0120, Reno (USA), 14–17 Janvier 2002.
- [166] C. MICHLER, S.J. HULSHOFF, E.H. van BRUMMELEN et R. de BORST : A monolithic approach to fluid-structure interaction. *Computers & Fluids*, Vol. **33**, No. 5–6, p. 839–48, 2004.
- [167] B.C. MOORE : Principal component analysis in linear systems : Controllability, observability, and model reduction. *IEEE Transactions on Automatic Control*, Vol. **26**, No. 1, p. 17–32, 1981.
- [168] H. J.-P. MORAND et R. OHAYON : *Interactions Fluide/Structure*. Masson, 1992.
- [169] D. MORENO, A. KROTHAPALLI, B. ALKISLAR et L.M. LOURENCO : Low-dimensional model of a supersonic rectangular jet. *Physical Review E*, Vol. **69**, No. 026304, p. 1–12, 2004.
- [170] G.-D. MORTCHÉLÉWICZ : Application of proper orthogonal decomposition to linearized Euler or Reynolds-Averaged Navier-Stokes equations. In *47<sup>th</sup> Israel Annual Conference on Aerospace Sciences*, Tel-Aviv (Israel), 2007.
- [171] B.R. NOACK, K. AFANASIEV, M. MORZYNSKI, G. TADMOR et F. THIELE : A hierarchy of low-dimensional models for the transient and post-transient cylinder wake. *Journal of Fluid Mechanics*, Vol. **497**, p. 335–63, 2003.
- [172] B.R. NOACK, P. PAPAS et P.A. MONKEWITZ : The need for a pressure-term representation in empirical Galerkin models of incompressible shear flows. *Journal of Fluid Mechanics*, Vol. **523**, p. 339–65, 2005.
- [173] A. NOUY : A generalized spectral decomposition technique to solve a class of linear stochastic partial differential equations. *Computer Methods in Applied Mechanics and Engineering*, Vol. **196**, No. 45–48, p. 4521–37, 2007.
- [174] F. OBLÉ et D. BUISINE : Instabilité et décrochage tournant en grilles d’aubes. *Comptes-Rendus de l’Académie des Sciences. Série IIb, mécanique, physique, astronomie*, Vol. **327**, No. 11, p. 1147–54, 1999.
- [175] H.M. PARK et O.Y. KIM : Reduction of modes for the control of viscous flows. *International Journal of Engineering Science*, Vol. **39**, No. 2, p. 177–200, 2001.
- [176] N. PATANKAR, P. SINGH, D. JOSEPH, R. GLOWINSKI et T.-W. PAN : A new formulation of the distributed Lagrange multiplier/fictitious domain method for particulate flows. *International Journal of Multiphase Flows*, Vol. **26**, No. 9, p. 1509–24, 2000.
- [177] L. PERRET, E. COLLIN et J. DELVILLE : Polynomial identification of POD based low-order dynamical system. *Journal of Turbulence*, Vol. **7**, No. 17, p. 1–15, 2006.
- [178] C.L. PETTIT et P.S. BERAN : Application of proper orthogonal decomposition to the discrete Euler equations. *International Journal for Numerical Methods in Engineering*, Vol. **55**, No. 4, p. 479–97, 2002.

- [179] A. PLACZEK, J.-F. SIGRIST et A. HAMDOUNI : Numerical simulation of an oscillating cylinder in a cross-flow at low Reynolds number : Forced and free oscillations. *Computers & Fluids*, Vol. **38**, No. 1, p. 80–100, 2009.
- [180] A. PLACZEK, D.-M. TRAN et R. OHAYON : Evaluation of an hybrid POD formulation for responses under prescribed displacements. In *International Conference on Noise and Vibration Engineering*, Louvain (Belgique), 15–17 septembre 2008.
- [181] A. PLACZEK, D.-M. TRAN et R. OHAYON : Hybrid proper orthogonal decomposition formulation for linear structural dynamics. *Journal of Sound and Vibration*, Vol. **318**, No. 4–5, p. 943–64, 2008.
- [182] A. PLACZEK, D.-M. TRAN et R. OHAYON : Nonlinear POD-based reduced-order models for aeroelastic compressible flows. In *International Forum on Aeroelasticity and Structural Dynamics*, Seattle (USA), 21–25 juin 2009.
- [183] F. POIRION : Modélisation temporelle des systèmes aéroélastiques. Application à l'étude des effets des retards. *La Recherche Aéronautique*, Vol. **2**, p. 103–14, 1995.
- [184] T.H. PULLIAM : Low Reynolds number numerical solutions of chaotic flows. In *27<sup>th</sup> Aerospace Sciences Meeting*, No. 89-0123, Reno (USA), 9–12 janvier 1989.
- [185] S.S. RAVINDRAN : Proper Orthogonal Decomposition in optimal control of fluids. Technical report TM-1999-209113, NASA, 1999.
- [186] D. REMPFER : *Kohärente Strukturen und Chaos beim laminar-turbulenten Grenzschichtumschlag*. Thèse de doctorat, Institut für Aerodynamik und Gasdynamik der Universität Stuttgart, 1988.
- [187] D. REMPFER : Investigations of boundary layer transition via Galerkin projections on empirical eigenfunctions. *Physics of Fluids*, Vol. **8**, No. 1, p. 175–88, 1996.
- [188] D. REMPFER : On low-dimensional Galerkin models for fluid flow. *Theoretical and Computational Fluid Dynamics*, Vol. **14**, No. 2, p. 75–88, 2000.
- [189] M.C. ROMANOWSKI : Reduced order unsteady aerodynamic and aeroelastic models using Karhunen-Loève eigenmodes. In *6<sup>th</sup> AIAA/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, No. 3981, Bellevue (Washington), September 1996.
- [190] M.C. ROMANOWSKI et E.H. DOWELL : Reduced order Euler equations for unsteady aerodynamic flows : Numerical techniques. In *34<sup>th</sup> Aerospace Sciences Meeting and Exhibit*, No. 0528, Reno (USA), Janvier 1996.
- [191] Y. ROUIZI, Y. FAVENNEC, J. VENTURA et D. PETIT : Numerical model reduction of 2D steady incompressible laminar flows : Application on the flow over a backward-facing step. *Journal of Computational Physics*, Vol. **228**, No. 6, p. 2239–55, 2009.
- [192] C.W. ROWLEY : *Modeling, Simulation and Control of Cavity Flow Oscillations*. Thèse de doctorat, California Institute of Technology, 2001.
- [193] C.W. ROWLEY : Model reduction for fluids using balanced proper orthogonal decomposition. *International Journal of Bifurcation and Chaos*, Vol. **15**, No. 3, p. 997–1013, 2005.
- [194] C.W. ROWLEY, T. COLONIUS et R.M. MURRAY : Model reduction for compressible flows using POD and Galerkin projection. *Physica D*, Vol. **189**, No. 1–2, p. 115–29, 2004.
- [195] C.W. ROWLEY et J.E. MARSDEN : Reconstruction equations and the Karhunen-Loève expansion for systems with symmetry. *Physica D*, Vol. **142**, No. 1–2, p. 1–19, 2000.
- [196] G. ROZZA, D.B.P. HUYNH et A.T. PATERA : Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. *Archives of Computational Methods in Engineering*, Vol. **15**, No. 3, p. 229–75, 2007.



## BIBLIOGRAPHIE

---

- [197] J.A. RULE, D.E. COX et R.L. CLARK : Aerodynamic model reduction through balanced realization. *AIAA Journal*, Vol. **42**, No. 5, p. 1045–8, 2004.
- [198] D. RYCKELYNCK : Réduction a priori de modèles thermomécaniques. *Comptes Rendus Mécanique*, Vol. **330**, No. 7, p. 499–505, 2002.
- [199] D. RYCKELYNCK : A priori hyperreduction method : an adaptative approach. *Journal of Computational Physics*, Vol. **202**, No. 1, p. 346–66, 2005.
- [200] R. SAMPAIO et C. SOIZE : Remarks on the efficiency of POD for model reduction in non-linear dynamics of continuous elastic systems. *International Journal for Numerical Methods in Engineering*, Vol. **72**, No. 1, p. 22–45, 2006.
- [201] S. SARKAR et K. VENKATRAMAN : Model order reduction of unsteady flow past oscillating airfoil cascades. *Journal of Fluids and Structures*, Vol. **19**, No. 2, p. 239–47, 2004.
- [202] J. SARRATE, A. HUERTA et J. DONEA : Arbitrary Lagrangian-Eulerian formulation for fluid-rigid body interaction. *Computer Methods in Applied Mechanics and Engineering*, Vol. **190**, No. 24–25, p. 3171–88, 2001.
- [203] C.V.S. SASTRY, D. ROY MAHAPATRA, S. GOPALAKRISHNAN et T.S. RAMAMURTHY : An iterative system equivalent reduction expansion process for extraction of high frequency response from reduced order finite element model. *Computer Methods in Applied Mechanics and Engineering*, Vol. **192**, No. 15, p. 1821–40, 2003.
- [204] J.-S. SCHOTTÉ et R. OHAYON : Various modelling levels to represent internal liquid behaviour in the vibration analysis of complex structures. *Computer Methods in Applied Mechanics and Engineering*, Vol. **198**, No. 21–26, p. 1913–25, 2009.
- [205] L. SCHWARTZ : *Analyse : Topologie générale et analyse fonctionnelle*. Hermann, 1970.
- [206] A.-S. SENS et A. DUGEAI : Numerical unsteady aerodynamics for turbomachinery aeroelasticity. In *10<sup>th</sup> International Symposium on Unsteady Aerodynamics, Aeroacoustics and Aeroelasticity of Turbomachines*, Durham (USA), 7–11 Septembre 2003.
- [207] H. SHAHVERDI, A.S. NOBARI, M. BEHBAHANI-NEJAD et H. HADDADPOUR : An efficient reduced-order modelling approach based on fluid eigenmodes and boundary element method. *Journal of Fluids and Structures*, Vol. **23**, No. 1, p. 143–53, 2007.
- [208] H. SHAN, L. JIANG et C. LIU : Direct numerical simulation of flow separation around a NACA0012 airfoil. *Computers & Fluids*, Vol. **34**, No. 9, p. 1096–114, 2005.
- [209] S.W. SHAW et C. PIERRE : Non-linear normal modes and invariant manifolds. *Journal of Sound and Vibration*, Vol. **150**, No. 1, p. 170–3, 1991.
- [210] I.A. SIBETHEROS et J.M. NIEDZWECKI : Analysis of single and tandem cylinder data using an orthogonal Volterra model approach. *Ocean Engineering*, Vol. **32**, No. 2, p. 135–56, 2005.
- [211] J.-F. SIGRIST, C. LAINÉ et B. PESEUX : Analyse modale d’une structure industrielle avec prise en compte du couplage fluide/structure. *Mécanique & Industrie*, Vol. **6**, No. 5, p. 553–63, 2005.
- [212] W.A. SILVA : *Discrete-time linear and nonlinear aerodynamic impulse responses for efficient CFD analyses*. Thèse de doctorat, College of William & Mary, 1997.
- [213] W.A. SILVA et R.E. BARTELS : Development of reduced-order models for aeroelastic analysis and flutter prediction using the CFL3Dv6.0 code. *Journal of Fluids and Structures*, Vol. **19**, No. 6, p. 729–45, 2004.
- [214] L.M. SILVEIRA, M. KAMON, I. ELFADEL et J. WHITE : A coordinate-transformed Arnoldi algorithm for generating guaranteed stable reduced-order models of RLC circuits. *Computer Methods in Applied Mechanics and Engineering*, Vol. **169**, No. 3–4, p. 377–89, 1999.

- [215] S. SIRISUP et G. KARNIADAKIS : Stability and accuracy of periodic flow solutions obtained by a POD-penalty method. *Physica D*, Vol. **202**, No. 3–4, p. 218–37, 2005.
- [216] S. SIRISUP et G.E. KARNIADAKIS : A spectral viscosity method for correcting the long-term behavior of POD modes. *Journal of Computational Physics*, Vol. **194**, No. 1, p. 92–116, 2004.
- [217] S. SIRISUP, G.E. KARNIADAKIS, D. XIU et G. KEVREKIDIS : Equation-free/Galerkin-free POD-assisted computation of incompressible flows. *Journal of Computational Physics*, Vol. **207**, No. 2, p. 568–87, 2005.
- [218] L. SIROVICH : Turbulence and the dynamics of coherent structures, Parts I-III. *Quarterly of Applied Mathematics*, Vol. **XLV**, No. 3, p. 561–90, 1987.
- [219] P.R. SPALART et S.R. ALLMARAS : A one equation turbulence model for aerodynamic flows. In *30<sup>th</sup> Aerospace Sciences Meeting and Exhibit*, No. 0439, Reno (USA), 6–9 janvier 1992.
- [220] A. STEINDL et H. TROGER : Methods for dimension reduction and their application in nonlinear dynamics. *International Journal of Solids and Structures*, Vol. **38**, No. 10–13, p. 2131–47, 2001.
- [221] A.V. STEPHEN, I.M. MOROZ, P.L. READ et W.-G. FRÜH : A comparison of empirical orthogonal decomposition methods in baroclinic flows. *Dynamics of Atmospheres and Oceans*, Vol. **27**, No. 1–4, p. 649–60, 1998.
- [222] E. STONE et A. CUTLER : Introduction to archetypal analysis of spatio-temporal dynamics. *Physica D*, Vol. **96**, No. 1–4, p. 110–31, 1996.
- [223] G. STRANG : *Linear Algebra and its Applications*. Academic Press, 1980.
- [224] Y. TAMURA, S. SUGANUMA, H. KIKUCHI et K. HIBI : Proper orthogonal decomposition of random wind pressure field. *Journal of Fluids and Structures*, Vol. **13**, No. 7–8, p. 1069–95, 1999.
- [225] D. TANG, D. KHOLODAR, J.-N. JUANG et E.H. DOWELL : System identification and proper orthogonal decomposition method applied to unsteady aerodynamics. *AIAA Journal*, Vol. **39**, No. 8, p. 1569–76, 2001.
- [226] K. TANG, W.R. GRAHAM et J. PERAIRE : Active flow control using a reduced order model and optimum control. In *27<sup>th</sup> Fluid Dynamics Conference*, New Orleans (USA), 17–20 juin 1996. [citeseer.ist.psu.edu/tang96active.html](http://citeseer.ist.psu.edu/tang96active.html).
- [227] J.P. THOMAS, E.H. DOWELL et K.C. HALL : Modelling viscous transonic limit cycle oscillation behavior using a harmonic balance approach. In *43<sup>rd</sup> Structures, Structural Dynamics and Material Conference*, No. 1414, Denver (USA), 22–25 avril 2002.
- [228] J.P. THOMAS, E.H. DOWELL et K.C. HALL : Using automatic differentiation to create a nonlinear reduced-order-model aerodynamic solver. *AIAA Journal*, Vol. **48**, No. 1, p. 19–24, 2010.
- [229] A.N. TIKHONOV et V.Y. ARSEININ : *Solutions of Ill-Posed Problems*. John Wiley, 1977.
- [230] D.-M. TRAN : Component mode synthesis methods using interface modes. Application to structures with cyclic symmetry. *Computers & Structures*, Vol. **79**, No. 2, p. 209–22, 2001.
- [231] D.-M. TRAN, C. LIAUZUN et C. LABASTE : Methods of fluid-structure coupling in frequency and time domains using linearized aerodynamics for turbomachinery. *Journal of Fluids and Structures*, Vol. **17**, No. 8, p. 1161–80, 2003.
- [232] D.M. TRAN : Component mode synthesis methods using partial interface modes. Application to tuned and mistuned structures with cyclic symmetry. *Computers & Structures*, Vol. **87**, No. 17–18, p. 1141–53, 2009.
- [233] T. UEDA et E.H. DOWELL : Flutter analysis using nonlinear aerodynamic forces. *Journal of Aircraft*, Vol. **21**, No. 2, p. 101–9, 1984.

## BIBLIOGRAPHIE

---

- [234] Y. UTTURKAR, B. ZHANG et W. SHYY : Reduced-order description of fluid flow with moving boundaries by proper orthogonal decomposition. *International Journal of Heat and Fluid Flow*, Vol. 26, No. 2, p. 276–88, 2005.
- [235] V. VAPNIK : *The nature of statistical learning theory*. Springer Verlag, 1995.
- [236] J. VENTURA, M. GIRAULT, D. PETIT et M.-J. ESTÈVE : Reduced order model for a forced convection case : application to a heating circular cylinder in a cross-flow at turbulent Reynolds numbers. In 6<sup>th</sup> *International Symposium on Turbulence, Heat and Mass Transfer*, Rome (Italie), 14–18 septembre 2009.
- [237] J.M. VERDON, M. BARNETT et T.C. AYER : Unsteady aerodynamic models for turbomachinery aeroelastic and aeroacoustic applications. Contractor Report No. 4698, NASA, 1995.
- [238] N. VERDON : *Un système dynamique d'ordre réduit basé sur une approche APR-POD pour l'étude de l'interaction écoulement turbulent-particules*. Thèse de doctorat, Université de La Rochelle, 2007.
- [239] G. VIGO : La décomposition orthogonale propre appliquée aux équations de Navier-Stokes compressible instationnaire. Rapport de Recherche No. 3385, INRIA, 1998.
- [240] G. VIGO : *Méthodes de décomposition orthogonale aux valeurs propres appliquées aux écoulements instationnaires compressibles complexes*. Thèse de doctorat, Université de Paris IX Dauphine, 2000.
- [241] S. VOLKWEIN : Proper Orthogonal Decomposition and Singular Value Decomposition. SFB-Preprint No. 153, Graz University, 1999. [www.uni-graz.at/imawww/volkwein/svd.ps](http://www.uni-graz.at/imawww/volkwein/svd.ps).
- [242] H. VON STORCH, G. BÜRGER, R. SCHNUR et J.-S. VON STORCH : Principal oscillation patterns : A review. *Journal of Climate*, Vol. 8, No. 3, p. 377–400, 1995.
- [243] C. WAGSCHAL : *Topologie et analyse fonctionnelle*. Hermann, 1995.
- [244] J. WELLER, E. LOMBARDI, M. BERGMANN et A. IOLLO : Numerical methods for low-order modeling of fluid flows based on POD. *International Journal for Numerical Methods in Fluids*, 2009. In press, doi 10.1002/flid.2025.
- [245] J. WELLER, E. LOMBARDI et A. IOLLO : Robust model identification of actuated vortex wakes. *Physica D*, Vol. 238, No. 4, p. 416–27, 2009.
- [246] K. WILLCOX et J. PERAIRE : Balanced model reduction via the proper orthogonal decomposition. *AIAA Journal*, Vol. 40, No. 11, p. 2323–30, 2002.
- [247] K.E. WILLCOX : *Reduced-Order Aerodynamic Models for Aeroelastic Control of Turbomachines*. Thèse de doctorat, Massachusetts Institute of Technology, 2000.
- [248] K.E. WILLCOX, J. PERAIRE et J. WHITE : An Arnoldi approach for generation of reduced-order models for turbomachinery. *Computers & Fluids*, Vol. 31, No. 3, p. 369–89, 2002.
- [249] P. XU : Truncated SVD methods for discrete linear ill-posed problems. *Geophysical Journal International*, Vol. 135, No. 2, p. 505–14, 1998.
- [250] X.-J. YU et B.-J. LIU : Stereoscopic PIV measurement of unsteady flows in an axial compressor stage. *Experimental Thermal and Fluid Science*, Vol. 31, No. 8, p. 1049–60, 2007.
- [251] Z. YU : A DLM/FD method for fluid/flexible-body interactions. *Journal of Computational Physics*, Vol. 207, No. 1, p. 1–27, 2005.
- [252] B. ZHANG, Y. LIAN et W. SHYY : Proper Orthogonal Decomposition for three-dimensional membrane wing aerodynamics. In 33<sup>rd</sup> *AIAA Fluid Dynamics Conference and Exhibit*, No. 3917, Orlando (Florida), 23–26 Juin 2003.
- [253] O.C. ZIENKIEWICZ et K. MORGAN : *Finite Elements and Approximation*. Wiley, 1972.

---

CONSTRUCTION DE MODÈLES D'ORDRE RÉDUIT NON-LINÉAIRES BASÉS SUR LA DÉCOMPOSITION  
ORTHOGONALE PROPRE POUR L'AÉROÉLASTICITÉ

---

La prédiction numérique de la réponse de systèmes aéroélastiques devient rapidement très coûteuse dès qu'il s'agit d'effectuer une analyse paramétrique. Le recours à un modèle d'ordre réduit fidèle au système dynamique initial est alors indispensable pour réduire les coûts de calcul. Ainsi, l'utilisation d'une base de vecteurs issus de la décomposition orthogonale propre (POD) combinée ensuite à une projection de Galerkin s'est imposée en mécanique des fluides. La difficulté de la construction d'un tel modèle d'ordre réduit pour les systèmes aéroélastiques est triple : tout d'abord en raison des non-linéarités inhérentes aux équations de Navier-Stokes, ensuite à cause de l'incohérence de la définition des modes POD pour un domaine en mouvement et enfin de par la nécessité d'une stabilisation.

Tout d'abord, le modèle d'ordre réduit reposant sur l'utilisation des modes POD est appliqué à un système dynamique linéaire pour lequel plusieurs formulations sont développées pour tenir compte de conditions limites spécifiques. Ensuite, le modèle d'ordre réduit est développé pour les équations de Navier-Stokes d'un fluide compressible au moyen d'un jeu de variables particulier. Ce premier modèle réduit développé pour un domaine de frontières fixes est corrigé puis validé sur l'exemple d'un profil NACA0012 fortement incliné pour provoquer l'apparition d'une allée de vortex. Finalement, le modèle d'ordre réduit est étendu au cas d'un domaine mobile sous l'hypothèse d'un mouvement de corps rigide afin d'éviter le problème de définition des modes POD. Ce modèle est alors employé pour reproduire l'écoulement transsonique autour d'un profil NACA0064 animé d'un mouvement d'oscillation.

Mots clés : modèle réduit, décomposition orthogonale propre, projection Galerkin, aéroélasticité, fluide compressible, système non-autonome.

---

CONSTRUCTION OF NONLINEAR PROPER ORTHOGONAL DECOMPOSITION BASED REDUCED-ORDER  
MODELS FOR AEROELASTIC APPLICATIONS

---

The numerical prediction of aeroelastic systems responses becomes very costly when a parametric analysis has to be performed. The need of a reduced order model close to the initial dynamical system is therefore essential to reduce the computational costs. The use of a basis composed of vectors obtained by the proper orthogonal decomposition (POD) which is then combined to a Galerkin projection has thus grown on in fluid mechanics. However there are three main difficulties to apply this methodology to aeroelastic systems : first because of the nonlinearities involved in the Navier-Stokes equations, then because of the domain motion which leads to an incoherent definition of the POD modes and finally since a stabilization is required.

Firstly the reduced order model based on the use of POD modes is applied to a linear dynamical system for which different formulations have been developed to take into account specific boundary conditions. Then the reduced order model is developed for the Navier-Stokes equations of a compressible fluid with a particular set of variables. This first reduced order model developed for a spatial domain with fixed boundaries is corrected and validated in the case of a NACA0012 airfoil with a large angle of attack such that a vortex street appears. Finally, the reduced order model is extended to the case of a moving domain by assuming a rigid body motion. In this way, the problem of definition of the POD modes is avoided. This last model is applied to the reproduction of the transonic flow around a NACA0064 airfoil oscillating around an equilibrium position.

Keywords : reduced-order model, proper orthogonal decomposition, Galerkin projection, aeroelasticity, compressible fluid, non-autonomous system.