Deep mean-field modeling for successive bifurcations exemplified for the fluidic pinball

Nan Deng

INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2021IPPAE013

Thèse de doctorat

ENSTA

IP PARIS

Deep mean-field modelling for successive bifurcations exemplified for the fluidic pinball

Modélisation en champ moyen de bifurcations successives, illustrée sur le pinball fluidique

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à l'École nationale supérieure de techniques avancées

École doctorale n°626 (EDIPP) École doctorale de l'Institut Polytechnique de Paris (EDIPP)
Spécialité de doctorat : Mécanique des fluides

Thèse présentée et soutenue à Palaiseau, le 27 septembre 2021, par

NAN DENG

Composition du Jury :

Laurette S. Tuckerman Directrice de Recherche, PMMH-ESPCI	Présidente
Dwight Barkley Professeur, University of Warwick, UK	Rapporteur
Angelo Iollo Professeur, Université de Bordeaux	Rapporteur
Steven L. Brunton Professeur, University of Washington, USA	Examinateur
Lutz Lesshafft Chargé de Recherche, LadHyX, École polytechnique	Examinateur
Themistoklis P. Sapsis Professeur, MIT, USA	Examinateur
Luc R. Pastur Professeur, IMSIA, ENSTA Paris	Directeur de thèse
Bernd R. Noack Professeur, Harbin Institute ot Technology, Shenzhen, China	Co-directeur de thèse

To my parents,
to whom I owe everything .....

Acknowledgments

I wish to express my sincere appreciation to my supervisors Prof. Luc Rémi Pastur and Prof. Bernd Rainer Noack, for their guidance, trust, and continuous support. Almost four years ago, I started my master's internship with them. Their enthusiasm and passion for research inspired me and motivated me to pursue an academic career. They provided me with the utmost help in applying for this thesis. This opened the door to my scholarly career and the first page of this exciting adventure. During the three years of this thesis, they provided me with lots of opportunities to share our results and discuss with many smart researchers all over the world. Their enthusiasm, patience, and foresight have been a great boost to my research. Without them, this work would have been impossible.

I want to thank my colleagues in our research group: Prof. Marek Morzyński, for his numerical support on the fluidic pinball, and Dr. Guy Y. Cornejo Maceda and Prof. François Lusseyran, for scientific discussions and friendship with them. The FlowCon members are also acknowledged for their insightful exchanges during the meetings. Special thanks to Prof. Jean-Christophe Loiseau. I have learned a lot from him at the beginning of my thesis. I also thank all my lab mates at IMSIA for the harmonious atmosphere they have created and for the happiness sharing with me. During my stay at LIMSI, I want to thank all the PhD students in AERO group. We have explored all the nearby cafeterias and shared joyful coffee breaks. Thanks also go to all the people at LIMSI who make it a pleasant place to be.

I would like to express my deepest gratitude to the members of the jury for their interest in my work. I would like to start by giving special thanks to Prof. Laurette S. Tuckerman. I am fortunate to work with her on the study of bifurcation. I have joined the course of Prof. Lutz Lesshafft at École Polytechnique and have learned a deal from it. I also thank Prof. Steven L. Brunton, Prof. Themis Sapsis for the precious exchanges we had during their visit at LIMSI. Their rich knowledge and academic enthusiasm have greatly inspired me, which has been of great importance to my research. Last but not least, I sincerely thank Prof. Angelo Iollo and Prof. Dwight Barkley for accepting to evaluate my work as referees.

I am deeply grateful to my girlfriend Peipei, for her constant and unconditional love, support and accompaniment. The days sharing with you are full of joy and happiness. I also want to thank my girlfriend's supervisor Prof. Dominique Lesselier and his wife for their help and care in life. I am grateful to my friends for many enjoyable times together.

My great appreciation goes to my parents, Guifang & Chuanbo, and grandparents, Kequn & Yunian, for their continuous love and support. They are my strong backing and always encourage me in everything I want to do.

Financial support by the China Scholarships Council (No.201808070123) during my thesis in the ENSTA Paris of Institut Polytechnique de Paris and numerical support from the laboratories LIMSI and IMSIA is gratefully acknowledged.

Deep mean-field modelling for successive bifurcations exemplified for the fluidic pinball

Nan Deng
UME - DFA, IMSIA, ENSTA Paris - IP Paris
828, Boulevard des Maréchaux 91120 Palaiseau, France
http://nandeng.me

Abstract
Artificial intelligence becomes increasingly important in solving problems that are difficult to handle with traditional mathematical methods. Reduced-order modelling (ROM) is no exception. In this thesis, we pave the way to automatable ROM in flow dynamics using first principles and machine learning techniques. The aim is to establish a benchmark problem for the most important dynamical features of wake flows.

The chosen benchmark configuration is the two-dimensional incompressible wake flow around the fluidic pinball, a cluster of parallel cylinders whose axis are located at the vertices of an equilateral triangle pointing upstream. At low Reynolds numbers, this configuration has a stable steady state satisfying the reflectional symmetry. With increasing Reynolds numbers, it undergoes two supercritical bifurcations of Hopf and pitchfork types, associated with the Bénard-von Kármán instability and a symmetry-breaking instability, respectively. A secondary Hopf bifurcation leads to quasi-periodic asymmetric shedding, before finally bifurcating into a chaotic regime. From numerical investigations of the periodic regimes, the asymmetric periodic shedding is characterized by three steady solutions and three limit cycles, which evidently poses a challenge to automated modelling.

Before modelling, a comprehensive understanding of the underlying mechanisms is pursued, including linear stability analysis of steady solutions, Floquet analysis of periodic solutions, and nonlinear analysis of asymptotic dynamics. These analyses, together with the Galerkin method, are the starting point of our mean-field modelling strategy. A five-dimensional least-order mean-field model is proposed, resolving the six invariant sets induced by the first two successive bifurcations. In addition, we derive an aerodynamic force model associated with the mean-field Galerkin model. Sparse calibration is applied to balance the accuracy and complexity of the model. These efforts culminate in a sparse human interpretable model for the flow dynamics and a predictive model for the unsteady forces. The mean-field models proposed above provide a challenging benchmark example for automatable ROM, combining data-driven modelling with physical constraints for a better understanding of complex flow dynamics.

For quasi-periodic and chaotic regimes, we propose a hierarchical cluster-based network modelling (HiCNM), adapting to more complex dynamics with multi-scale, multi-frequency, multi-attractor behaviours. The only assumption about the model structure is the mean-field assumption. The HiCNM enables identifying the transient and post-transient dynamics between multiple invariant sets in a self-supervised manner and steps towards automated ROM of complex dynamics.

Keywords: Mean-field modelling, stability analysis, bifurcation, Galerkin method, least-order model, Galerkin force model, hierarchical cluster-based network model, automated modelling, first principles, machine learning, fluidic pinball.

Modélisation en champ moyen de bifurcations successives, illustrée sur le pinball fluidique

Résumé
L'intelligence artificielle prend de plus en plus d'importance dans la résolution de problèmes difficiles à traiter avec les méthodes mathématiques traditionnelles. La modélisation d'ordre réduit (ROM) ne fait pas exception. Dans cette thèse, nous ouvrons la voie à une modélisation réduite automatisable en mécanique des fluides, à partir de principes premiers et de techniques d'apprentissage automatique. L'objectif est d'établir un problème de référence présentant les caractéristiques dynamiques principales des écoulements de sillage.

La configuration de référence choisie est l'écoulement de sillage incompressible bidimensionnel autour du "pinball fluidique". Ce système est constitué d'un groupe de cylindres parallèles dont les axes sont situés aux sommets d'un triangle équilatéral pointant vers l'amont. À bas nombre de Reynolds, l'écoulement stationaire, symétrique par réflexion miroir, est stable. Lorsque le nombre de Reynolds augmente, le système subit deux bifurcations supercritiques de type Hopf et fourche, associées respectivement à l'instabilité de Bénard-von Kármán et à une instabilité de brisure de symétrie. Une bifurcation de Hopf secondaire conduit à un lâcher de tourbillons asymétrique quasi-périodique, avant de bifurquer finalement vers un régime chaotique. Le régime périodique asymétrique est caractérisé par trois solutions stables et trois cycles limites, ce qui pose un défi à la modélisation automatique.

En prélude à la phase de modélisation, la compréhension des mécanismes sous-jacents est menée par analyse de stabilité linéaire des solutions stationnaires, analyse de Floquet des solutions périodiques et analyse non linéaire de la dynamique asymptotique. Ces analyses, aux côtés de la méthode de Galerkin, constituent le point de départ de notre stratégie de modélisation en champ moyen. Un modèle de champ moyen à cinq dimensions reproduit les six ensembles invariants qui résultent des deux premières bifurcations. Nous dérivons également un modèle de forces aérodynamiques à partir du modèle de Galerkin en champ moyen. Une calibration parcimonieuse permet d'obtenir un modèle facilement interprétable, équilibrant précision et complexité.

Pour les régimes quasi-périodiques et chaotiques, nous proposons un modèle de réseau hiérarchique de clusters (HiCNM), capable de s'adapter à des dynamiques plus complexes présentant un comportement multi-échelles, multi-fréquences et multi-attracteurs. La seule hypothèse sur la structure du modèle est l'hypothèse de champ moyen. Le HiCNM permet d'identifier les dynamiques transitoires et post-transitoires entre de multiples ensembles invariants d'une manière auto-supervisée, ce qui constitue une étape vers la réduction de modèle automatisée des dynamiques complexes.

Mots clés: Modèle de champ moyen, stabilité, bifurcation, modèle de Galerkin, modèle de réseau, modèle automatisé.

Contents

Chapter 1

Introduction

Contents

The study of fluid flows can go back at least 2200 years ago to the days of ancient Greece, when Archimedes formulated buoyancy force exerted on a body immersed in a fluid, which is known now as the famous Archimedes' principle. This is considered to be the first fundamental law of physics to fluid mechanics. Less progress was made for nearly 1600 years after him until Leonardo da Vinci started to study turbulence in the fifteenth century. In the next 200 years, the research on fluid flows mainly stayed on observations and experiments of many other physicists. In the late seventeenth century and the early eighteenth century, Isaac Newton formulated the laws of motion in classical mechanics, and Daniel Bernoulli published Bernoulli's principle, which was further formulated by Leonhard Euler with Bernoulli's equation. They estabished the basic of fluid mechanics and the mathematical fluid dynamics. After that, more and more mathematicians joined the study of inviscid flow, i.e. Jean le Rond d'Alembert, Joseph Louis Lagrange, Pierre-Simon Laplace, Siméon Denis Poisson. In the nineteenth century, Claude-Louis Navier and George

Gabriel Stokes used a set of partial differential equations to describe the motion of viscous fluid, which express conservation of momentum and conservation of mass for Newtonian fluids in a mathematical form, known as the Navier–Stokes equations. At this point, a precise mathematical model has been estabished.

Solutions to the Navier–Stokes equations are critically important from theoretical studies to engineering applications. However, the mathematical properties of its solutions are still challenging in this century, which is known as the Navier–Stokes existence and smoothness problem – one of the seven Millennium Prize problems in mathematics. Despite the simple mathematical form of the Navier–Stokes equations and the clear observations in experiments, a complete understanding of the turbulence is still missing. With the rapid development of computational science, we can simulate fluid flows by solving the governing equations using numerical methods. However, it is still difficult or computationally expensive to solve the Navier–Stokes equations directly, especially for the turbulent flow. Instead of solving the Navier–Stokes equations directly – Direct numerical simulation (DNS), various advancements have been focused on the turbulence models, including the Reynolds-averaged Navier–Stokes (RANS), the Large eddy simulation (LES) and the hybrid RANS-LES models. To be noted, experts are still needed to check whether the numerical solution matches the experimental results, for a good consistency.

We can numerically investigate the flow dynamics by simulating the above-mentioned mathematical models. However, the flow dynamics are characterized by high dimensionality and nonlinearity. The simulations and analyses of the flow system require a high computational cost, especially for the systems with interactions between many different scales. Reduced-order modelling (ROM) techniques have been introduced to reduce the computational cost and provides low-dimensional and human-interpretable models for flow analysis, optimization, and control design.

In this thesis, we pave the way to automatable ROM in flow dynamics using first principles and machine learning techniques. The aim is to establish a benchmark problem that shares the most important dynamical features of wake flows. The chosen benchmark configuration is the two-dimensional incompressible wake flow around the fluidic pinball, a cluster of three parallel cylinders whose axis are located at the vertices of an equilateral triangle pointing upstream. Multiple-cylinders structures are commonly found in engineering applications. The spatial organization and the size of cylinders have a significant influence on the wake flow. For the fluidic pinball, despite the simplicity of the geometry, the wake flow can present rich flow dynamics due to the wake interactions behind the cylinders. For example, at a low Reynolds number, the three cylinders work as a single bluff-body, and we can find a von Kármán vortex street downstream. At a higher Reynolds number, a base-bleeding jet appears between the back two cylinders. Keep increasing Reynolds number, the length of the jet increases. The jet undergoes a symmetry-breaking instability, with a permanent deflection of the base-bleeding jet up or down. Both the von Kármán vortex street and the jet are typical fluid dynamics examples and have been well studied for their transition from laminar to turbulent conditions. Bifurcation theory provides a very useful tool for qualitative and quantitative analysis of the behavior of complex systems, where the dynamics within systems have specific types of bifurcation. Generally, the periodic vortex shedding and the symmetry breaking result from a Hopf bifurcation and a pitchfork bifurcation, respectively. The pitchfork bifurcation originates locally in the near wake, while the von Kármán vortex street originates in the wake of the three cylinders taken as a single obstacle. The coupling relations and the interaction between these two instabilities highly depend on their spatial distribution.

The aim of the present research is to understand the transient and post-transient dynamics and the interaction of these two fundamental instabilities, and to obtain human-interpretable reduced-order models. The analysis of the linear and nonlinear dynamics is based on the linear stability analysis of the Navier-Sokes equations and the nonlinear saturation in the asymptotic regimes. Based on the physical understanding of the transition, we will introduce three kinds of reduced-order modelling strategies, which improve the performance of the Galerkin method, the force-

oriented modelling, and the cluster-based network modelling for the complex flow. The analytical works can guarantee the resulting models for the ability to capture transient dynamics and human-interpretation, but require considerable effort to understand the underlying mechanisms. However, only a part of prior knowledge is needed for a gray-box model to balance the interpretation and the cost of analytical works. The knowledge from the governing equation, the geometry, and mean-field considerations can vastly improve the performance of the model. Because of the simple geometry of the flow configuration and the generic bifurcations discussed in this work, the analysis and modelling process can be applied to other flow systems with similar dynamics.

The introductory chapter is structured as follows. § 1.1 discusses the purpose of reduced-order modelling and outlines the categories of modelling. § 1.2 introduces the flow configuration benchmarked in this thesis. The transient and post-transient dynamics with multiple attractors are briefly investigated. In § 1.3, the analytical works are introduced, including the linear stability analysis of the Navier-Stokes solutions, the instabilities and bifurcations. In § 1.4, we discuss the nonlinear saturation for the post-transient dynamics, and the mean-field deformation from the effect of the Reynolds stress. § 1.5 introduces the three modelling strategies, corresponding to the mean-field Galerkin model, the Galerkin force model and the cluster-based mean-field model, respectively. § 1.6 provides an overall conclusion of the methodologies applied and the main results obtained in this thesis, reiterates the motivation for mean-field modelling, and finishes with an outline of the following chapters.

1.1 Reduced-order modelling

The complexity of the fluid flow comes from its high-dimensionality, non-linearity, and multi-scale spatial and temporal behavior. When dealing with fluid flow problems, a massive amount of flow detail information will be generated as the resolution increases. Obviously, it is extremly hard to understand all these details of different scales, and the control optimization is impossible to be applied. Despite the nearly infinite dimensionality of a natural flow, its dynamics can usually be captured with projection onto a low-dimensional invariant manifold. Discussing and analyzing the dynamics in this low-dimensional space makes it possible to build an approximate model for the complex flow system. The purpose of reduced-order modelling is to obtain simplified, interpretable models of fluid flow, which can provide a better understanding of the underlying mechanisms. Such low-dimensional models are essential for flow dynamics prediction and effective control design.

The reduced-order modelling process for a flow system can be divided into two steps. It usually begins with a modal decomposition to achieve dimensionality reduction. Several patterns, or modes, featuring a typical behaviour in space or time are frequently enough to capture the essential flow mechanisms. The modal analysis of these modes is common practice and can help us understand complex flow dynamic (Taira *et al.*, 2017). Numerous techniques have been developed to extract those primary patterns, such as the proper orthogonal decomposition (POD) (Holmes *et al.*, 2012*a*; Berkooz *et al.*, 1993), and dynamic mode decomposition (DMD) (Schmid, 2010; Rowley *et al.*, 2009; Mezić, 2013; Tu *et al.*, 2014). As the POD modes are arranged according to the energy but not in the order of the dynamical importance, some variants of POD techniques have been developed for a better dynamics capture capability, such as the balanced proper orthogonal decomposition (balanced POD) (Willcox & Peraire, 2002; Rowley, 2005; Ilak & Rowley, 2006), and spectral proper orthogonal decomposition (SPOD) (Sieber *et al.*, 2016; Towne *et al.*, 2018; Cavalieri *et al.*, 2019; Lesshafft *et al.*, 2019).

Based on the extracted modes, we still need a nonlinear model to describe the dynamics in the low-dimensional space. Unlike the rapid development of modal decomposition or dimensionality reduction, nonlinear models depend on our understanding of nonlinear systems, and more theories and techniques need to be developed. Based on the decomposition methods, the general approaches are POD-Galerkin methods (Aubry *et al.*, 1988; Iollo *et al.*, 2000; Rowley *et al.*, 2004) and Koopman theory (Rowley *et al.*, 2009; Schmid, 2010; Mezić, 2013). The Galerkin projection-

based reduced-order model can provide a low-dimensional representation for complex flow, and then an accurate dynamical model can be derived from first principles (Noack *et al.*, 2003; Galletti *et al.*, 2004; Buffoni *et al.*, 2006). Koopman theory uses an infinite-dimensional linear operator to describe nonlinear dynamics, and DMD provides a finite-dimensional approximation to this Koopman operator. The modelling with DMD is purely data-driven and does not require any a priori assumptions or knowledge of the underlying dynamics (Kutz *et al.*, 2016).

For extreme events in fluid flows (Lesieur, 1987), POD and DMD cannot effectively capture the intermittently occurring episodes due to the stochasticity. The low-order statistical modelling will better serve the purpose to capture the relevant physics (Sapsis & Majda, 2013; Blonigan *et al.*, 2019; Sapsis, 2021).

In this thesis, we consider the flow governed by the non-dimensionalized incompressible Navier-Stokes equations:

$$\partial_t \boldsymbol{u} + \nabla \cdot \boldsymbol{u} \otimes \boldsymbol{u} = \frac{1}{Re} \triangle \boldsymbol{u} - \nabla p, \quad \nabla \cdot \boldsymbol{u} = 0, \tag{1.1}$$

where a full information of the flow state contains the velocity flow field $\boldsymbol{u}(\boldsymbol{x}, t)$ and the pressure field $p(\boldsymbol{x}, t)$, defined with $\boldsymbol{q} = (\boldsymbol{u}, p)$. The existing states and their instabilities are generally a function of control parameters. The only control parameter here is the Reynolds number $Re = U_\infty D/\nu$, where U_∞ is the velocity in the far wake, D is the cylinder diameter, and ν is the kinematic viscosity. The flow system can be rewritten as

$$\partial_t \boldsymbol{q} = F(\boldsymbol{q}, Re), \tag{1.2}$$

where F is a nonlinear function for the original Navier-Stokes equations. As shown in figure 1.1, for example, the cylinder flow $\boldsymbol{q}(\boldsymbol{x}, t)$ shows different characteristics with increasing Reynolds number Re, from the steady flow to the turbulent flow.

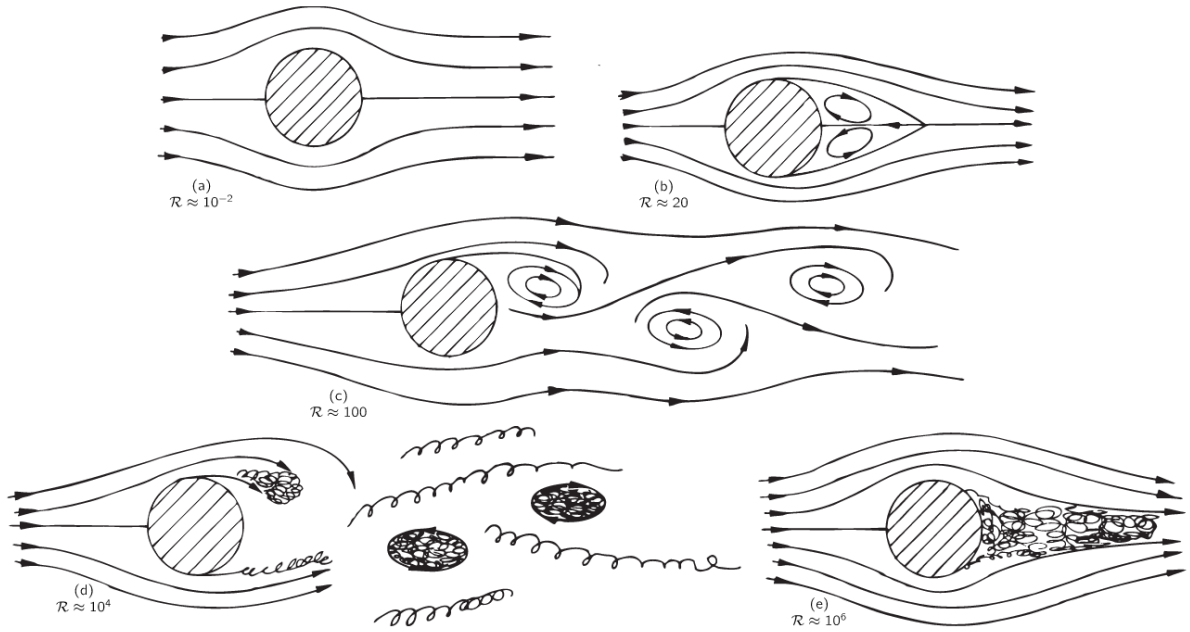


Fig. 41–6. Flow past a cylinder for various Reynolds numbers.

Figure 1.1: Cylinder flow with increasing Reynolds number Re, with characterized with multiple spatial structures with different scales: (a) the streamlines around the cylinder at a very low Re, (b) a circulation appears behind the cylinder, (c) periodic vortex shedding, known as the von Kármán streets of vortices, (d) the turbulent region appears, but still with a regular alternating motion, (e) the turbulent region reaches the cylinder, with a clear "turbulent boundary layer" and losing periodicity. From Feynman *et al.* (2011).

For real flows, the instantaneous flow state $\boldsymbol{q}(\boldsymbol{x}, t)$ is of infinite dimension in space, and varies in time. By numerical simulation, the spatial structure of the instantaneous flow state is described

on the nodes of the discretized mesh. The flow field information can be the velocity/pressure field from the numerical simulation or the data measured from an experiment. By projecting the dynamics onto a low-dimensional space, the spatio-temporal information of the flow field will usually be separated into a finite number of spatial modes $\hat{\boldsymbol{q}}_i(\boldsymbol{x})$ with amplitudes $a_i(t)$. In this case, the flow will be represented with a low-dimensional state vector $\boldsymbol{a}(t) = (a_1, \cdots, a_N)$, where N is the number of spatial modes (degrees of freedom) in the low-dimensional space. The dynamics of this low-dimensional state vector $\boldsymbol{a}$ approximates the original Navier-Stokes equations, which reads:

$$\dot{\boldsymbol{a}} = f(\boldsymbol{a}, Re), \tag{1.3}$$

where f is an approximated nonlinear function of F approximated in low-dimensional space. Note that f provides a complete description of the flow system F considering an infinite-dimensional basis $\hat{\boldsymbol{q}}_i(\boldsymbol{x})$, $i = 1, \cdots, \infty$.

1.1.1 Why reduced-order modelling?

Analytical work in infinite dimensions is intractable since we cannot consider an infinite number of patterns $\hat{\boldsymbol{q}}_i(\boldsymbol{x})$ at the same time. The nonlinearity and multi-scale problem in space and time make the flow dynamics even more complicated. A reduced-order model (ROM) of the flow system will be useful in understanding physical mechanisms, improving computational efficiency and designing control laws. In many cases, the flow dynamics usually evolves on a low-dimensional attractor, which can be charactrized with a finite number of elementary sturctures. These typical structures, featuring certain spatial and temporal scales, offer the possibility of using reduced-order models to reveal the underlying mechanisms hidden by the high-dimensionality.

Reduced-order models (ROMs) have many advantages:

1. **Simplicity:** A good feature extraction of dominant flow characteristics can promise the simplicity of the resulting ROM. New dimensionality reduction techniques are rapidly developing, and there are still many opportunities to find the best set of modes.
2. **Understanding:** The low-dimensional representation of the original nonlinear dynamics helps us to filter out the noise and unimportant terms and reveals the underlying mechanisms. Linear and nonlinear dynamics can be interpreted in terms of several degrees of freedom. Machine learning techniques provide new methods and opportunities to understand complex dynamics from the data itself.
3. **Analysis:** The interaction between the elementary degrees of freedom provides a deep insight into the nonlinear dynamics. The analytical works can figure out the fundamental mechanisms and help us to understand the nonlinear transition.
4. **Control:** ROM can figure out the dominant dynamics which contribute to the efficient design of linear and optimal control.
5. **Prediction:** A good approximation of nonlinear dynamics in low-dimensional subspaces ensures fast dynamic prediction of impending transitions and enables early warning.
6. **Speed up the computation:** A surrogate model can explore the dynamics in the low-dimensional space. The ROM is dynamically equivalent to the original system and can be mapped into the high-dimensional data to speed up the simulation.

1.1.2 White box modelling vs black-box modelling

In this subsection, we briefly categorize the different kinds of modelling. We follow Wiener (1948) and distinguish between white-box models (WBM), gray-box models (GBM), and black-box models (BBM), as shown in figure 1.2

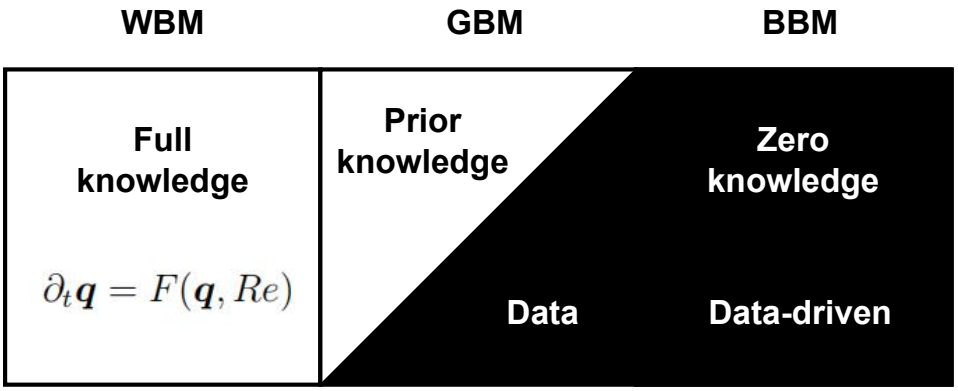


Figure 1.2: Illustration of the concept of three kinds of modelling strategies. White-box model (WBM) is a purely theoretical model, which has full knowledge of the flow system. Gray-box models (GBM) combines a partial theoretical basis with data to complete the model. Black-box models (BBM) is a statistical model without any prior knowledge of the system.

White box modelling

The Navier–Stokes equations provide a mathematical model of fluid mechanics, which can be considered as a white-box model. The modelling process is also known as clear box modelling because of having the full knowledge of the governing equations for the flow system. Complete information of the flow state can be derived from this model but requires cumbersome computation. As mentioned at the beginning of this chapter, numerous problems are still unsolved to handle with this white-box model. Some weak form of the Navier–Stokes equations has been provided to solve these problems, like RANS, LES, and some other turbulence models. They are still white-box models, as all the assumptions come from the theoretical derivation. These models are expensive to simulate numerically and therefore difficult to be used for control design, optimization, and engineering applications.

In summary, the goal of white-box modelling is to find the "right" model mathematically and physically. Therefore, the modelling process is analytical, emphasizing the cause and effect, the mathematical form, and the physical principles. It requires the human knowledge of the Navier-Stokes equations and the underlying mechanisms of the flow dynamics. The high requirements of human experience and the expensive simulation cost make it hard to apply for industrial and engineering applications.

Black box modelling

A black-box model comes from a purely data-driven modelling process, emphasizing the identification, prediction ability, and modelling speed. The goal is to find an "accurate" model for the system features, dynamics, and input-output relations. A good accuracy makes it even more useful in industrial applications and medical diagnosis.

The accuracy usually comes from the sacrifice of interpretability. As an example of neural network modelling, the internal functional relationships of multiple layers have a higher liberty and greater complexity, making it difficult to be interpreted from a physical or mathematical perspective. However, this complex neural network can build a robust model to achieve an accurate mapping of inputs and outputs.

Analogously, according to the Koopman theory (Rowley *et al.*, 2009; Schmid, 2010; Mezić, 2013), the nonlinear dynamics in a finite-dimensional space can be represented as a linear dynamics in an infinite-dimensional space. The dynamic mode decomposition (DMD) provides a linear representation of the data of a nonlinear system by an infinite-dimensional Koopman operator. The resulting models have a good ability to capture the coherent features.

Different from the above-mentioned approaches, the cluster-based reduced-order modelling (CROM) provides a new way to describe the transient dynamics. The time-related snapshots of the flow field are partitioned into clusters. Snapshots in the same cluster share similar characteristics and are represented by the cluster centroid, which is usually defined as the average of the snapshots in the cluster. The transient dynamics is described by the transitions between the clusters. Kaiser *et al.* (2014) uses a cluster-based Markov model (CMM) to describe the transient dynamics with probabilities. Nair *et al.* (2019) applied CMM to the nonlinear feedback flow control with an extension of the Markov chain, which emphasizes the non-trivial transitions between clusters. The directed network (Newman, 2018) was introduced, where clusters are considered as the nodes and transitions between clusters as the edges. Fernex *et al.* (2021) and Li *et al.* (2021) further proposed the cluster-based network model (CNM) for the time-resolved data by introducing the local interpolation between clusters with the determined transition times.

Grey box modelling

Compared to the white-box modelling and the black-box modelling, the grey box modelling can be seen as a hybrid approach, using data-driven techniques to build reduced-order models based on theoretical structures. For example, POD-Galerkin models approximate the full state by Galerkin projection of the Navier-Stokes equations on the orthogonal basis of a low-dimensional state space. The basis modes come from a data-driven POD process. The resulting Galerkin system of quadratic ordinary differential equations describes the nonlinearity of the Navier-Stokes equations.

1.1.3 Opportunities for grey box modelling

In this thesis, we focus on the grey box modelling using flow data, first principles, and machine learning techniques. The following three modelling methods have been successfully applied on the fluidic pinball, which will be further discussed in section 1.5.

Least-order mean-field modelling with sparse Galerkin regression:

Least-order mean-field models are least-order approximations of the unsteady flow field, using a minimum number of dominant modes for describing the fluctuating components and the mean-field deformation. For the transient and post-transient dynamics of the fluidic pinball, two general bifurcations in fluid mechanics are considered: one is the supercritical Hopf bifurcation, leading to the periodic release of vortices in the wake, another one is the supercritical pitchfork bifurcation, leading to the static symmetry-breaking of the unsteady fluid flow. An optimized subspace of five degrees of freedom is chosen under mean-field considerations for these two bifurcations. The resulting Galerkin system is further simplified from the constraints of frequency balance and geometry (the reflectional symmetry of the modes). This mean-field Galerkin framework provides a sparse and simplified structure based on the analytical derivation.

For the modelling process, the nonlinearities should always satisfy the above constraints and present a sparse form. Some coefficients of the model can be determined by the physical inspection of the flow dynamics: the growth rate and angular frequency of the fluctuating modes are given by the linear instability analysis of the base flow, the coefficient of the saturating term by the amplitude of the solution in the asymptotic regime, etc. The coefficients of the remaining nonlinear terms are identified by the SINDy algorithm (Brunton *et al.*, 2016*a*). This nonlinear modelling process is also known as the constrained sparse Galerkin regression (Loiseau & Brunton, 2018), with fitting the nonlinear dynamics in a given quadratic form under additional constraints. This work also shows that the cubic form can fit the dynamics better. However, there is no cubic nonlinearity in the Navier-Seokes equations explicitly. Therefore, we do not know how to constraint the model and cannot ensure that the identified model is physical. In Chapter 2, the least-order model successfully describes the underlying mechanisms of dynamics. This indicates that the optimized model basis and mean-field modelling framework are critical important for a robust Galerkin model.

The Galerkin force modelling with sparse Galerkin regression:

Galerkin force modelling aims to present a force-related model in the Galerkin framework. Generally, a Galerkin model is based on a linear decomposition of the velocity field. However, the aerodynamic forces originates from both the viscous and pressure fields, where the pressure field is a quadratic function of the velocity field from the pressure Poisson equation. This implies that the force contribution of each velocity mode cannot be added independently. Under a velocity-based Galerkin decomposition of the flow field, the body force can be derived as a constant-linear-quadratic function of the mode amplitudes either from the integral of viscous and pressure forces on the surface, or from the momentum balance of the Navier-Stokes equations. Moreover, the geometrical symmetry of the modes promotes an additional sparsity in the force model.

An accurate pressure field reconstruction from the velocity field is numerically challenging and expensive. The force contribution of each velocity mode in the Galerkin expansion can be numerically determined by solving the pressure Poisson equation with the homogeneous Neumann boundary conditions. As each mode carries its own contribution to the force balance, an accurate reconstruction of the force dynamics requires many leading modes. The resulting projection-based force model only provides the force contribution of each mode without any relevance for the underlying dynamics. In Chapter 4, we exemplify the Galerkin force modelling for the unforced fluidic pinball based on the least-order mean-field model. The nonlinear modelling of the force dynamics is based on the elementary degrees of freedom associated with the instabilities undergone by the flow, and is implemented with a sparse Galerkin regression. The unsteady force evolution of the transient and post-transient dynamics can be successfully reproduced with only seven velocity modes under the mean-field consideration. The sparse drag and lift force formulae indictes the elementary drag- and lift-producing modes.

Hierarchical cluster-based network modelling:

The above two strategies are based on a low-dimensional approximation of the system, with the nonlinear dynamics modelled in an optimal subspace. The cluster-based reduced-order modelling (CROM) provides a new perspective to decribe the complex dynamics, by coarse-graining hundreds of thousands time-resolved snapshots into a few number of centroids with clustering. This data-driven modelling strategy can be fully automated by black-box modelling parametrized by the number of clusters, and provides a statistical description of a high-dimensional flow system.

The Reynolds decomposition under mean-field considerations divides the flow field into a hierarchy of three components: a slowly-varying mean-flow field, coherent components and non-coherent small scale fluctuations. Moreover, these three components have different frequency characteristics. The hierarchical strategy can be easily applied to the data-driven modelling process, especially for the transient flow dynamics from the onset of the instability to the saturated state. With a clear frequency separation, the mean-flow distortion during the transient dynamics can be accurately modelled based on the low-pass filtered data, removing the small scales and harmonic components of the fluctuating dynamics. Therefore, CROM can be generalized for multi-scale and multi-frequency dynamics by incorporating the hierarchical structure. The resulting hierarchical cluster-based reduced-order model (HiCROM) is physically consistent with the Reynolds decomposition, and able to identify transient and post-transient dynamics in a self-supervised manner. The cluster distributions can be dynamically optimized according to the resolution of the transient dynamics. In Chapter 5, we will apply HiCROM to the transient and post-transient dynamics of the unforced fluidic pinball at three different Reynolds numbers, respectively associated with a periodic, a quasi-periodic and a chaotic dynamics.

1.2 Challenges with multiple invariant sets

Reduced-order modelling most usually focuses on the post-transient dynamics evolving on the attractor. The modelling of the transient and post-transient dynamics from the fixed point to the asymptotic regime of the attractor needs more effort as it has to feature the mean-field distortion, the local dynamics, and the nonlinear saturation. Even more challenging is to build a universal

reduced-order model for multiple transient trajectories with multiple invariant sets. The flow dynamics considered in this thesis involves several transient and post-transient regimes, three fixed points, three limit cycles, two quasi-periodic attractors and a chaotic attractor, at three different Reynolds numbers.

1.2.1 A benchmark configuration – the fluidic pinball

Fluid flow around a circular cylinder is one of the most famous fundamental fluid mechanics problem, which has been well studied over decades. Compared to the single-cylinder, the multiple-cylinders structures are more frequently used in industry, i.e. in heat exchangers, undersea pipelines, battery pack assembly, etc. The configurations with multiple cylinders offer more possibilities for different kinds of instabilities. The interaction between cylinder wakes produces more complex flow dynamics.

In this work, we are concerned with the flow over a cluster of three parallel cylinders, which has been experimentally studied in a context of heat transfer, fluid-structure interactions and multiple frequencies interactions over the past few decades (Price & Paidoussis, 1984; Sayers, 1987; Lam & Cheung, 1988; Tatsuno *et al.*, 1998; Bansal & Yarusevych, 2017). Five different flow patterns have been identified with varying the spacing ratios and Reynolds numbers by the numerical simulation (Bao *et al.*, 2010; Zheng *et al.*, 2016; Gao *et al.*, 2019; Chen *et al.*, 2020).

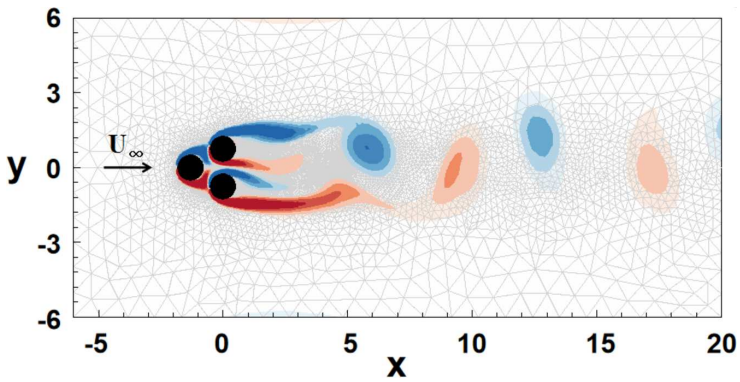


Figure 1.3: Configuration of the fluidic pinball and dimensions of the simulated domain. A typical field of vorticity is represented in color with $[-1.5, 1.5]$. The upstream velocity is denoted U_∞.

For the fluidic pinball, three fixed cylinders of diameter D whose axes are located at the vertices of an equilateral triangle of side $3D/2$ are placed in the (x, y) plane and perpendicularly to this plane. The gap distance between the cylinders is one radius R and the formed triangle points upstream, as shown in figure 1.3. They are placed in a viscous incompressible uniform flow at speed U_∞. The cylinders can rotate at different speeds, which allows one to change the paths of the incoming fluid just as flippers manipulate the ball of a conventional pinball machine. This configuration is a great sandbox to test different flow control strategies and has been used for the evaluation of flow controllers for multiple-input multiple-output(MIMO) control dynamics (Ishar *et al.*, 2019; Cornejo Maceda *et al.*, 2021). In this study, all three cylinders remain static as we are interested in the natural dynamics of the flow as the Reynolds number is increased. The unforced fluidic pinball has already shown a surprisingly rich dynamics.

The flow is considered in the Cartesian coordinate system bounded in a rectangular domain $[-6, 20] \times [-6, 6]$. The origin of the Cartesian coordinate system is placed in the middle of the back two cylinders. The computational domain is discretized on an unstructured grid. The pinball configuration uses a grid with 4 225 triangles and 8 633 vertices. The boundary conditions comprise a no-slip condition on the cylinders and a unit velocity in the far field:

$$U_r = 0 \text{ on the cylinders and } U_\infty = \boldsymbol{e}_x \text{ at infinity.} \tag{1.4}$$

The far-field boundary conditions are exerted on the inflow, upper and lower boundaries, while the outflow boundary is assumed to be stress-free, transparent for the outgoing fluid structures. To test the grid dependency of the solution, we have compared the result of simulations with a refined grid in Chapter 2. Both simulations prove grid independence and yield dynamically consistent results.

The instantaneous flow field is calculated by two-dimensional direct numerical simulations (DNS). The unsteady Navier-Stokes solver is based on fully implicit time integration and Finite-Element Method discretization (Noack & Morzyński, 2017; Noack *et al.*, 2003, 2016). The time integration is third-order accurate while FEM discretization employs second-order Taylor-Hood finite elements (Taylor & Hood, 1973). The solution is obtained iteratively, with the Newton-Raphson type approach. The tangent matrix is updated on each iteration, and computations are carried out until the residual is under a prescribed tolerance. The steady solution is obtained in a similar Newton-Raphson iteration for the steady Navier-Stokes equations. For $Re > 68$, there exist three unstable steady solutions, one characterized by a symmetric base-bleeding jet $\boldsymbol{U}_s(\boldsymbol{x})$, and two by a top or bottom deflected jet $\boldsymbol{U}_s^{\pm}(\boldsymbol{x})$. The choice of initial conditions in the iteration triggers the convergence to one of the three steady solutions. The solver quickly converges to one of the steady states, and a final near-zero residual confirms that this is indeed the steady flow solution sought. Analogously, for the unsteady Navier-Stokes solver, the transient dynamics depends on the initial condition. Some typical initial conditions are the unstable steady solutions $\boldsymbol{U}_s(\boldsymbol{x})$, $\boldsymbol{U}_s^{\pm}(\boldsymbol{x})$. It requires an initial perturbation to destabilize the flow from the given steady solution.

1.2.2 Transient and post-transient dynamics

The flow dynamics is divided into a transient regime, when starting close by an unstable solution, either steady or periodic, and a post-transient regime, when evolving on the attractor. Due to the successive bifurcations (see § 1.3.2) occurring in the flow system, the unforced fluidic pinball possesses multiple invariant sets. In this subsection, the transient and post-transient dynamics with multiple invariant sets will be described with the time evolution of the drag and lift forces on the cylinders. The vorticity field of flow states explored during the transient dynamics will be shown. In the end, we will present the 3D phase space with the force dynamics.

Flow states during the transient and post-transient dynamics

We apply DNS starting from the unstable steady symmetric solution $\mathbf{u}_s$ to the asymptotic regime, for different values of the Reynolds number Re. Theoretically, the system should never leave the initial state without perturbations as it is a fixed point of the system. In practice, numerical discretization errors make the solution imperfect. Despite an extremely long transient (several hundreds of convective times), the trajectory of the system eventually escapes from $\mathbf{U}_s$ and asymptotically reaches the attractor. The flow states explored during the transient and post-transient dynamics are illustrated with the vorticity field in figure 1.4.

We notice that the reflectional symmetry in y proves to be an important part of this scenario. For a velocity field $\mathbf{u} = (u, v)$, we define the y-reflection operator R via $R(u, v)(x, y) \equiv (u, -v)(x, -y)$. For a symmetric field, u is even and v is odd in y. The spanwise vorticity $\partial_x v - \partial_y u$ used to represent the flows in our visualizations, manifests y-reflection symmetry by being odd in y.

The symmetric steady solution at different Reynolds numbers is always reflection-symmetric with respect to $y = 0$, with $R\mathbf{U}_s = \mathbf{U}_s$. This solution undergoes a supercritical Hopf bifurcation at $Re_1 \approx 18$ with a pair of y-antisymmetric eigenmodes, leading to cyclic vortex shedding in the wake flow, see § 1.3.2 for detail. For the von Kármán vortex street formed at $Re = 30$, it has the spatio-temporal symmetry $\mathbf{u}_{\mathrm{vk}}(t + T/2) = R\mathbf{u}_{\mathrm{vk}}(t)$ (Barkley, 2006). Moreover, its one-period time-average satisfies the y-reflection symmetry. For this reason, we will call this limit cycle (or the vortex shedding) symmetric despite the fact that the instantaneous flows are not symmetric. We

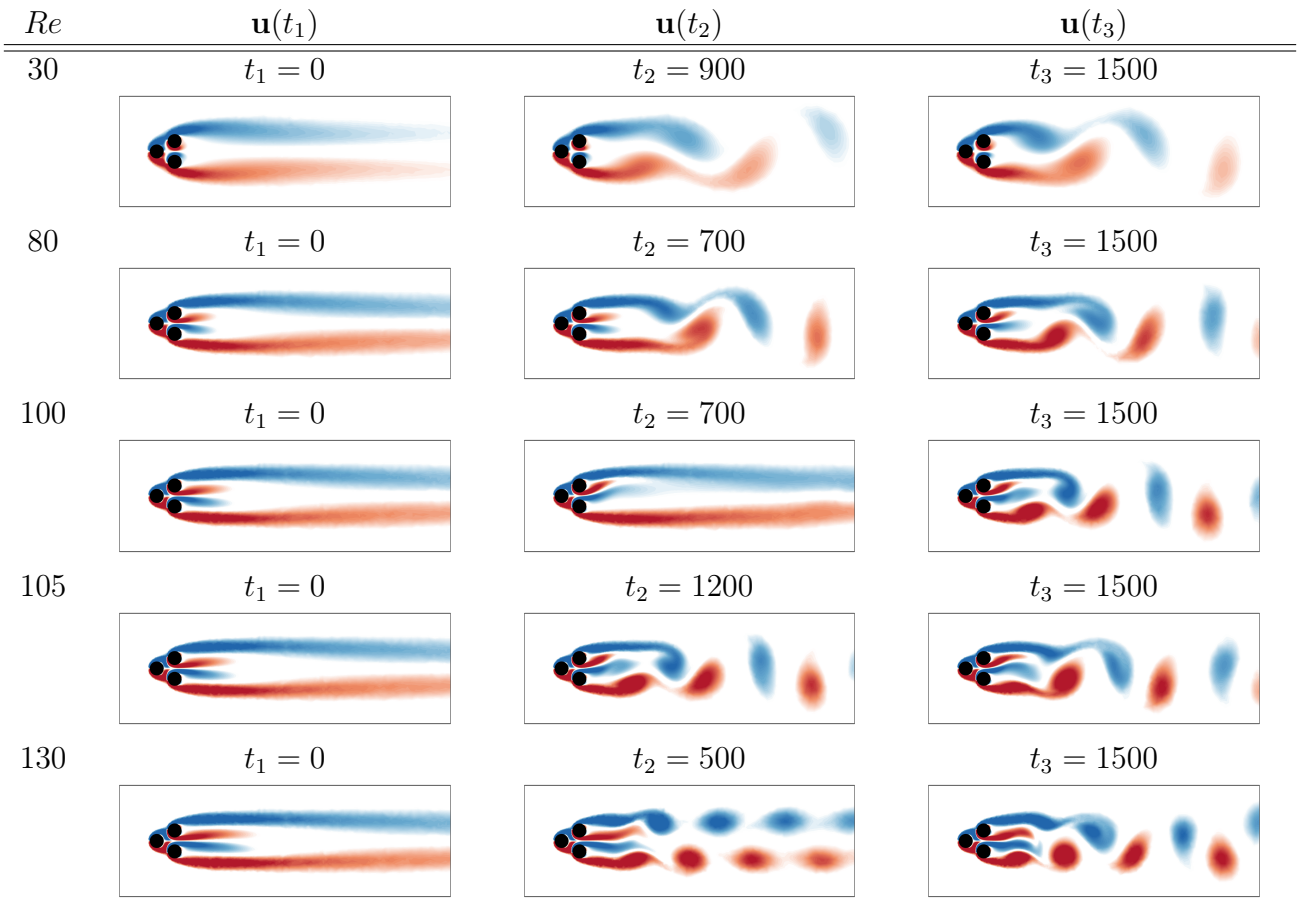


Figure 1.4: Flow states during the transient dynamics at different values of the Reynolds numbers, resulting from DNS starting with the corresponding symmetric steady solution. Three snapshots were selected for their typical flow characteristics. Vorticity fields are plotted in color with the range $[-1.5, 1.5]$.

also notice that the destabilization starts from the far wake for the vortex shedding. The length of the recirculation zone behind the cylinders will get shorter from the steady solution to the final stable vortex street.

The system undergoes a supercritical pitchfork bifurcation at $Re = 68$, which breaks the reflectional symmetry by deflecting the base-bleeding jet to the top or the bottom, see § 1.3.2 for detail. At $Re = 80$, the von Kármán vortex street preserves the spatio-temporal symmetry until $t \approx 700$, and the base-bleeding jet behind the back two cylinders is almost unaffected by the development of the vortex shedding. However, this state is unstable, and the base-bleeding jet will eventually deflect to the top or the bottom.

At $Re = 100$, the transient dynamics exhibits a reversed order. The base-bleeding jet deflects first to the top, before the formation of the von Kármán vortex street. To be noted, the base-bleeding jet will keep deflected to the top even in the fully developed vortex street. The von Kármán vortex street is affected by the base-bleeding jet and slightly deflected downside. The coupling relationship between them is still weak in this case.

At $Re = 105$, the transient dynamics is similar to $Re = 100$. However, the base-bleeding jet will oscillate around the upwards deflected position at a lower frequency f_{JET} and modulate the vortex shedding in the near wake. The coupling relationship between them cannot be ignored anymore.

At $Re = 130$, the vortex shedding comes out first. The initial symmetric von Kármán vortex street characterized by a staggered arrangement is separated into two sub-streets of positive and negative vorticity. The length of the recirculation zone is very short now, and the base-bleeding will also oscillate and mix the two sub-streets up. The base-bleeding jet and the von Kármán

vortex street are highly coupled and interact with each other.

Dynamics of the drag and lift forces

In figure 1.5 are represented the transient and post-transient dynamics of the lift C_L and drag C_D coefficients from the unstable symmetric steady solution $\mathbf{U}_s$ to the asymptotic regime, for different values of the Reynolds number Re. The transient dynamics shown in figure 1.5 reveals the following features:

- For $Re = 30$, the system stays in the vicinity of the unstable steady solution $\mathbf{U}_s$ for an extremely long time. An obvious transition is observed over the time range $t \in [800, 1000]$. C_D increases quickly, and C_L starts to oscillate around a vanishing mean value with an increasing amplitude. The flow eventually goes into an asymptotic regime. In the vicinity of the attractor, the oscillation frequency of C_L is the same as the vortex shedding frequency f_{VS}, while C_D oscillates at twice the frequency.
- For $Re = 80$, during the transient dynamics, one can see that C_L starts oscillating around a vanishing mean value before reaching its asymptotic non-zero mean value. Henceforth, the symmetry of the mean flow is broken in a second step only, hundreds of convective time units after the wake have started to oscillate. The lowest C_D is associated with the symmetric steady solution $\mathbf{U}_s$. The transient state, observed over the time range $t \in [700, 800]$, has a lower drag coefficient than that of the final regime. The amplitude of oscillation of C_D is very small, and the frequency is identical to the vortex shedding frequency f_{VS}.
- For $Re = 100$, C_L first reaches a non-zero value before it starts to oscillate, and C_D decreases to the lowest value before it increases to the highest value. From figure 1.4, we know the symmetry breaking occurs first and the resulting state is the asymmetric steady solution with base-bleeding jet deflected to the top. In the asymptotic regime, the amplitude of oscillation of C_D is very small, but C_L oscillates with a much larger amplitude. Their oscillation frequencies are identical to the vortex shedding frequency f_{VS}. We notice that a low frequency appears in the beginning of the saturation but quickly damps.
- For $Re = 105$ & 110, C_D first reaches the lowest value, then oscillates with larger amplitude, at the frequency of the jet modulation, which is one order of magnitude lower than the frequency of the vortex shedding, $f_{\text{JET}} \approx f_{\text{VS}}/12$. The oscillation at f_{VS} also exists but with a very small amplitude. C_L first breaks the symmetry before it starts to oscillate at f_{VS} and is finally modulated by the frequency of the jet.
- For $Re = 120$, before entering into the fully chaotic regime, the initial part of the transient dynamics of C_D and C_L indicates the flow first reaches the asymmetric steady solution with base-bleeding jet deflected to the top. C_L starts to oscillate around the non-zero mean value, but will finally trapped into the chaotic regime with vanishing mean value.
- For $Re = 130$, C_L starts to oscillate around zero, before entering into the chaotic regime with vanishing mean value. This corresponds to the new kind of initial transient dynamics preserving the symmetry observed in figure 1.4. C_D begins to increase monotonically, but soon shows a chaotic dynamics with a wider range of variation than $Re = 120$.

Force dynamics with different initial conditions

The transient and post-transient dynamics depend on the initial conditions. Except for $Re = 30$, we apply DNS respectively with choosing one of the three steady solutions as the initial conditions. We apply the reflectional symmetry operator to the trajectory issued from the symmetric steady solution to obtain its mirror-conjugated trajectory. Four trajectories are plotted in an embedding space of the drag coefficient $C_D(t)$, lift coefficient $C_L(t)$ and time-delayed lift coefficient $C_L(t-\tau)$, with $\tau = 2$ units of time, as shown in figure 1.6.

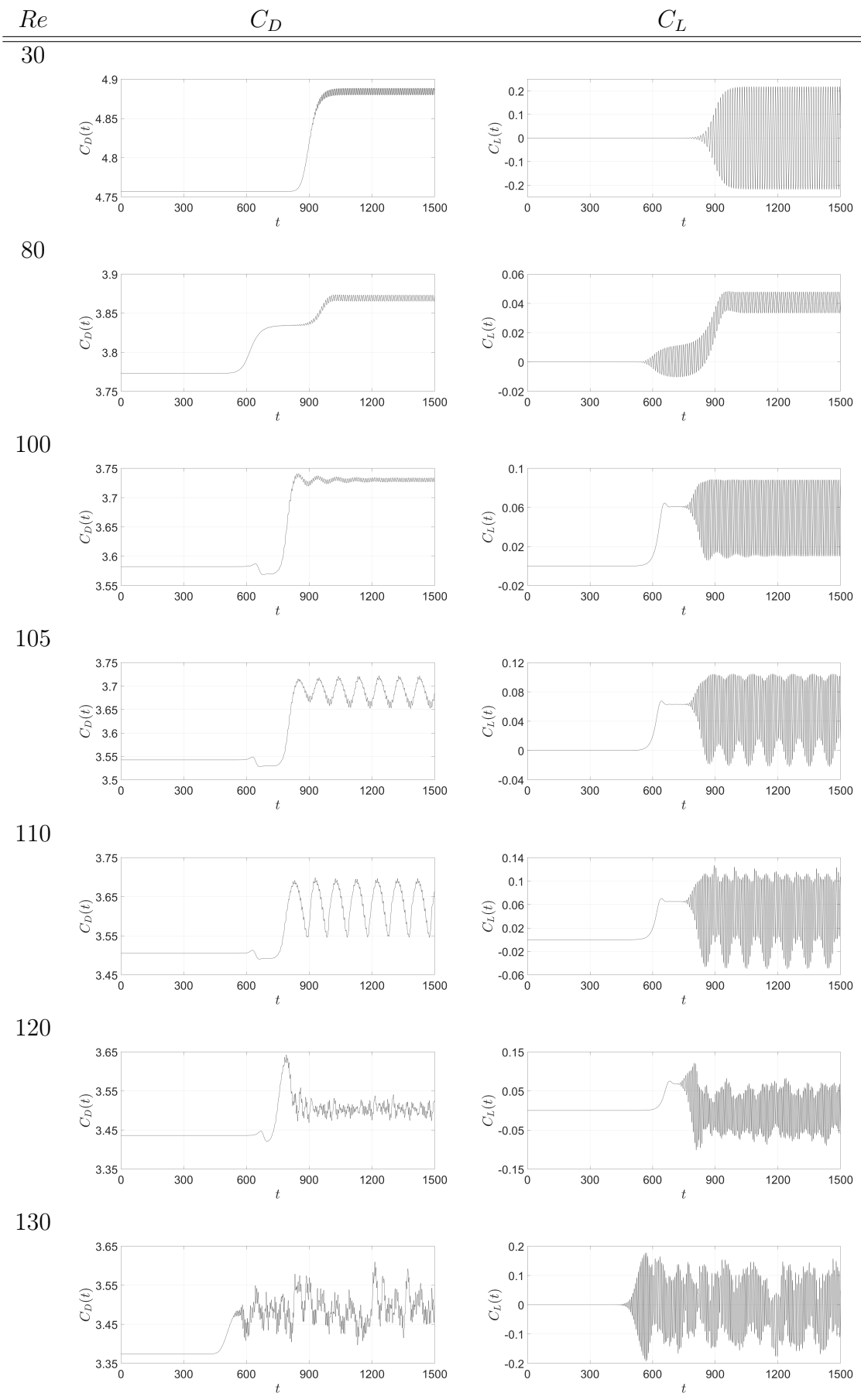


Figure 1.5: Transient dynamics of the drag C_D (left) and lift C_L (right) coefficients for increasing values of the Reynolds numbers, resulting from DNS starting with the corresponding symmetric steady solution.

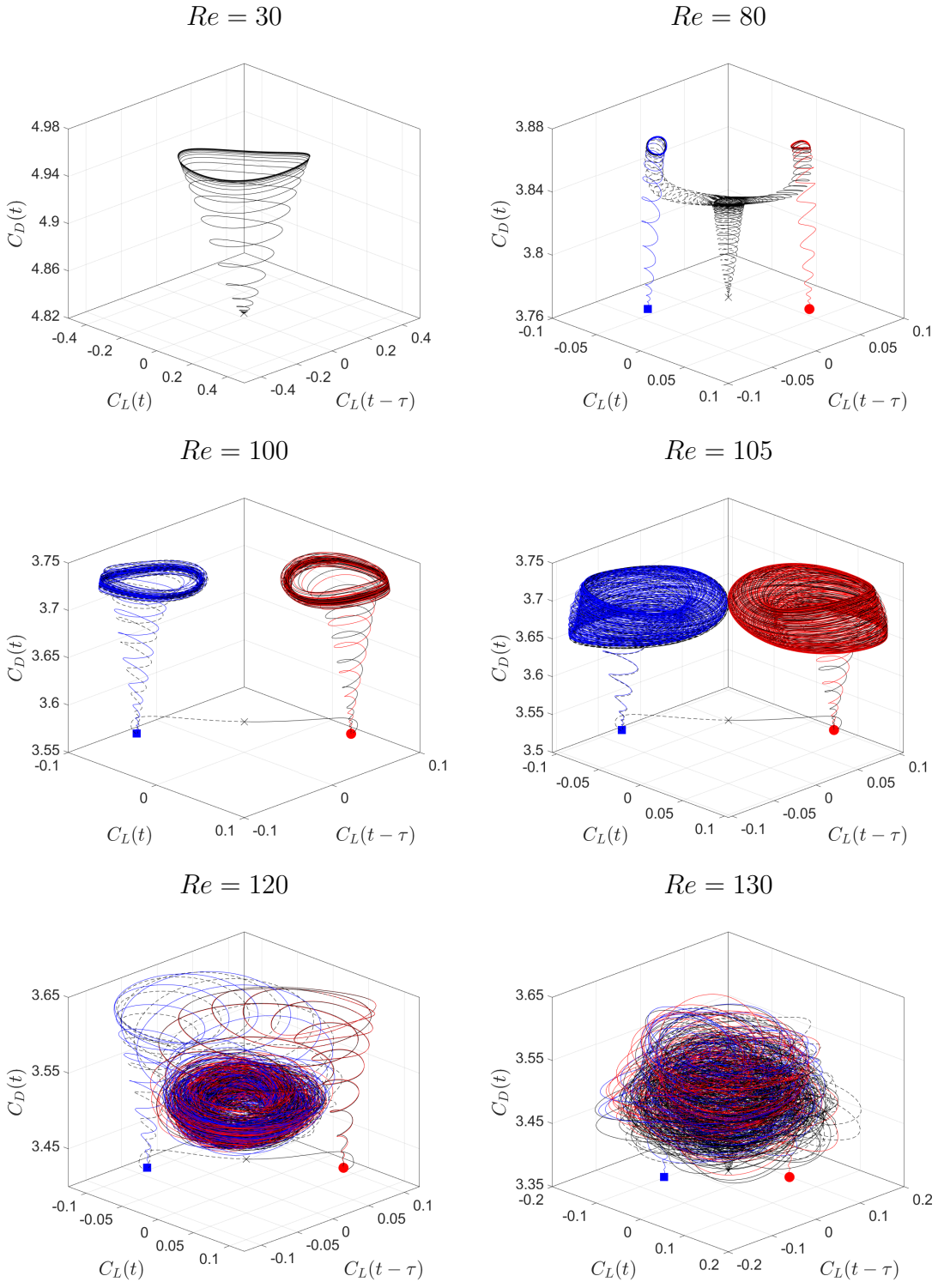


Figure 1.6: Scenarios of transient and post-transient dynamics, based on the drag coefficient $C_D(t)$, lift coefficient $C_L(t)$ and time-delayed lift coefficient $C_L(t-\tau)$, resulting from the DNS starting with the following unstable symmetric/asymmetric steady solutions at different Reynolds numbers: the black curve from the symmetric steady solution $\boldsymbol{u}_s(\times)$, the black dashed curve is its mirror-conjugated trajectory, the red curve from the asymmetric steady solution $\boldsymbol{u}_s^-(\bullet)$, and the blue curve from the asymmetric steady solution $\boldsymbol{u}_s^+(\blacksquare)$.

For a typical trajectory in the phase space, the dynamics is first linear in the neighbourhood of the steady solution, and becomes progressively nonlinear as the amplitude of the perturbation grows, until the asymptotic regime is reached on the attractor. Especially for the trajectories from the symmetric steady solution at $Re > 68$ beyond the critical value of the pitchfork bifurcation, they will first approach a transient state, for instance, an unstable limit cycle, like $Re = 80$, or the unstable asymmetric steady solutions, like $Re = 100, 105, 120$, before entering into the permanent state.

At $Re = 30$, there is only the von Kármán instability and only one trajectory from the fixed point to the limit cycle. For $80 \leq Re \leq 105$, there exist multiple trajectories from the three steady solutions to the two attractors. For $Re = 80$, there exists three unstable fixed points, an unstable limit cycle and two stable limit cycles. For $Re = 100$, the trajectories travel through three unstable fixed points and two stable limit cycles. For $Re = 105$, the two stable limit cycles turn into two attracting torus with quasi-periodic dynamics. For $Re \geq 120$, The flow will inherit the symmetry of the steady solutions before being trapped by the chaotic regime. The size of the chaotic area in the phase space increases from $Re = 120$ to 130. We also note there is a random increase in drag at $Re = 130$, while the chaotic dynamics of the drag coefficient is limited to a small range at $Re = 120$.

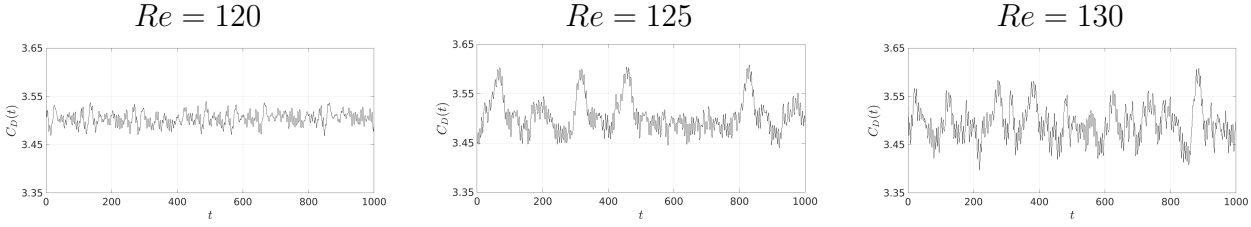


Figure 1.7: Temporal evolution of the drag coefficient $C_D(t)$ in the chaotic regimes at different values of the Reynolds number.

In figure 1.7, we observe a more apparent random increase at $Re = 125$, with a lower occurring probability within the same simulating duration comparing to the case at $Re = 130$. The fluidic pinball exhibits a stochastic character, involving the sudden transitions between chaotic and quasi-periodic regimes, with no specific frequency associated with this seldom event (Sapsis, 2021).

In summary, the unforced fluidic pinball provides rich flow dynamics. We are concerned with building a mean-field model at different Reynolds numbers, benchmarking the transient and post-transient dynamics for three cases: with six invariant sets, with the quasi-periodic flow regime, and with the chaotic flow state.

1.3 Linear stability analysis

In this section, we are interested in the dynamics of small perturbations in the vicinity of the base flow, which can be a steady solution or a periodic solution of the Navier-Stokes equations. Under this consideration, the time evolution of a tiny perturbation can be considered as a linear dynamics in a finite transient time. The unsteady flow system in Eq. (1.1) has at least one steady solution $\boldsymbol{Q}_s = (\boldsymbol{U}_s, P_s)$, satisfying the steady Navier-Stokes equations

$$0 = -\nabla \cdot (\boldsymbol{U}_s \otimes \boldsymbol{U}_s) + \nu \triangle \boldsymbol{U}_s - \nabla P_s, \tag{1.5a}$$

$$0 = \nabla \cdot \boldsymbol{U}_s. \tag{1.5b}$$

The boundary conditions are the same as the unsteady Navier-Stokes equations (1.1). The Newton's method, also known as the Newton–Raphson method, is used to solve the steady Navier-Stokes equations. In the form of Eq. (1.2), $F(\boldsymbol{Q}_s, Re) = 0$ indicates that the steady solution is a flow field depending on the Reynolds number. We note that an unstable steady solution can be

numerically computed by solving the steady Navier-Stokes equations, but is difficult to obtain in experiments or can only be observed transiently.

1.3.1 Linear stability of the base flow

To determine the stability of the base flow, one often employs the linear stability analysis, assuming linear dynamics of small perturbations of the base flow. The linear stability analysis of the base flow reveals the underlying mechanisms of the instability and predicts the early stages of the transient dynamics.

Linear stability analysis for the steady solution

Around the steady state $\boldsymbol{Q}_s(\boldsymbol{x})$, also known as the equilibrium of the governing equations, we consider an infinitesimal perturbation $\boldsymbol{q}'(\boldsymbol{x}, t)$:

$$\boldsymbol{q}(\boldsymbol{x}, t) = \boldsymbol{Q}_s(\boldsymbol{x}) + \boldsymbol{q}'(\boldsymbol{x}, t). \tag{1.6}$$

Substitute Eq. (1.6) into the unsteady Navier-Stokes equations (1.1), and neglect the nonlinear term $\boldsymbol{u}' \otimes \boldsymbol{u}'$ with second-order smallness, the dynamics of the perturbation is governed by the linearized Navier-Stokes equations:

$$\begin{aligned} \partial_t \boldsymbol{u}' &= -\nabla \cdot (\boldsymbol{u}' \otimes \boldsymbol{U}_s + \boldsymbol{U}_s \otimes \boldsymbol{u}') + \nu \triangle \boldsymbol{u}' - \nabla p', && \text{(1.7a)} \\ 0 &= \nabla \cdot \boldsymbol{u}', && \text{(1.7b)} \end{aligned}$$

with the homogeneous boundary conditions.

By definining the linearized Navier-Stokes operator around the steady solution $\boldsymbol{U}_s$ as $\mathcal{L}_{\boldsymbol{U}_s}$ and the perturbation state vector $\boldsymbol{q}' = (\boldsymbol{u}', p')$, the linearized system can be written as:

$$\partial_t \boldsymbol{q}' = \mathcal{L}_{\boldsymbol{U}_s} \boldsymbol{q}'. \tag{1.8}$$

The modal analysis of stability leads to an eigenvalue problem $(\sigma + i\omega)\,\hat{\boldsymbol{q}} = \mathcal{L}_{\boldsymbol{U}_s}\hat{\boldsymbol{q}}$, with perturbation in the form $\boldsymbol{q}'(\boldsymbol{x}, t) = \hat{\boldsymbol{q}}(\boldsymbol{x})e^{(\sigma+i\omega)t}$, which reads

$$\begin{aligned} \boldsymbol{u}'(\boldsymbol{x}, t) &= \hat{\boldsymbol{u}}(\boldsymbol{x})e^{(\sigma+i\omega)t} && \text{(1.9a)} \\ p'(\boldsymbol{x}, t) &= \hat{p}(\boldsymbol{x})e^{(\sigma+i\omega)t}. && \text{(1.9b)} \end{aligned}$$

To solve the eigenvalue problem, the simplest method is the power iteration method by repeatedly performing $\mathcal{L}_{\boldsymbol{U}_s}$ on an arbitrary initial perturbations $\hat{\boldsymbol{q}}_0$. Multiple eigenvalues can be computed by the Arnoldi method, or subspace iteration methods (Tuckerman & Barkley, 2000). We apply subspace iteration to obtain the leading eigenvalues of $\mathcal{L}_{\boldsymbol{U}_s}$, which is performed on a Krylov subspace of dimension 9-20. A maximum number of 100 iterations is preset, and the iterations will stop when the residual of the eigenvalue problem is less than 10^{-5} (Morzyński *et al.*, 1999).

The real part of the eigenvalue σ, corresponding to the growth rate of the perturbation, indicates the stability of the considered base flow $\boldsymbol{U}_s$. If all the σ are negative, the base flow is stable, indicating that all the perturbations will decay to zero, otherwise, the base flow is unstable. As the base flow depends on the Reynolds number, the eigenvalues will vary based on the Reynolds number. Generally, a stable flow tends to become unstable with increasing Reynolds number, when an eigenvalue crosses zero. The imaginary part ω determines the frequency of the instability at onset, which is non-zero for a Hopf bifurcation (Strogatz *et al.*, 1994).

Floquet stability analysis for time periodic flow

For time-periodic flows, we consider a T-periodic solution $\boldsymbol{Q}_p = (\boldsymbol{U}_p, P_p)$ of the unsteady Navier-Stokes equations (1.1), satisfying

$$\partial_t \boldsymbol{U}_p = -\nabla \cdot (\boldsymbol{U}_p \otimes \boldsymbol{U}_p) + \nu \triangle \boldsymbol{U}_p - \nabla P_p, \tag{1.10a}$$

$$0 = \nabla \cdot \boldsymbol{U}_p. \tag{1.10b}$$

The boundary conditions are the same as the unsteady Navier-Stokes equations.

Similar to the linear stability analysis for the base flow $\boldsymbol{U}_s$, the Floquet stability problem deals with a T-periodic base flow $\boldsymbol{U}_p(\boldsymbol{x}, t) = \boldsymbol{U}_p(\boldsymbol{x}, t + T)$ (Barkley & Henderson, 1996*a*; Schatz *et al.*, 1995). The linearized system around the T-periodic solution reads:

$$\partial_t \boldsymbol{q}' = \mathcal{L}_{\boldsymbol{U}_p(t)} \boldsymbol{q}' \tag{1.11}$$

The linear operator $\mathcal{L}_{\boldsymbol{U}_p(t)}$ is T-periodic because of the base flow $\boldsymbol{U}_p(\boldsymbol{x}, t)$. The solutions to Eq. (1.11) are seeked as:

$$\boldsymbol{q}'(\boldsymbol{x}, t) = \hat{\boldsymbol{q}}(\boldsymbol{x}, t) e^{(\sigma + i\omega)t}, \tag{1.12}$$

with the T-periodic Floquet modes $\hat{\boldsymbol{q}}(\boldsymbol{x}, t)$ and the corresponding Floquet exponents $\sigma + i\omega$. The Floquet operator is defined as the time-integrated $\mathcal{L}_{\boldsymbol{U}_p(t)}$, starting from a perturbation at t_0 and considering its evolution on the linearized system over one period T, which reads:

$$A_F = \exp\left(\int_{t_0}^{t_0+T} \mathcal{L}_{\boldsymbol{U}_p(t)} dt \right). \tag{1.13}$$

The Floquet multipliers of A_F can be written as $\lambda_F = e^{(\sigma + i\omega)T}$, with the relation to the Floquet exponents $\sigma + i\omega = (\ln \lambda_F)/T$. The eigenproblem can be solved by the same methods as for the linear stability analysis of the steady solution. In our case, at the critical value of the pitchfork bifurcation, the dominant Floquet multiplier is of interest for the stability analysis of the periodic solution preserving the symmetry. The resulting unstable eigenvector is associated with a real eigenvalue and can be obtained by a simple power iteration method. Analogously, the T-periodic solution is stable if the real part σ is negative for all the Floquet exponents, otherwise, the periodic solution is unstable. The eigenvectors are the Floquet modes at instant t_0, and one can integrate Eq. (1.11) to obtain the periodic Floquet modes at other instants.

1.3.2 Instabilities and bifurcations

In this section, we briefly introduce the bifurcations detected from the linear stability analysis of the steady solutions and from the Floquet analysis of the periodic solutions. Before the model analysis of the eigenvectors, we first discuss the reflection symmetry properties existing in this flow system, together with a relevant symmetry-based decomposition.

Z_2-group with reflection symmetry

Since the flow configuration of the fluidic pinball has reflection symmetry (sometimes called mirror symmetry) with respect to the x-axis, an important feature of the modal analysis is the reflection symmetry properties of the modes. For a flow state $\boldsymbol{q} = (u, v, p)$, a symmetry-based decomposition of the flow leads to a symmetric component $\boldsymbol{q}^s = (u^s, v^s, p^s) \in \mathcal{U}^s$ with

$$u^s(x, -y) = u^s(x, y), \quad v^s(x, -y) = -v^s(x, y), \quad p^s(x, -y) = p^s(x, y), \tag{1.14}$$

and an antisymmetric component $\boldsymbol{q}^a = (u^a, v^a, p^a) \in \mathcal{U}^a$ satisfying

$$u^a(x, -y) = -u^a(x, y), \quad v^a(x, -y) = v^a(x, y), \quad p^a(x, -y) = -p^a(x, y), \tag{1.15}$$

where $\mathcal{U}^s$ and $\mathcal{U}^a$ are the set of symmetric and antisymmetric vector fields, respectively. Analogously, we can derive the spanwise vorticity $\omega = \partial_x v - \partial_y u$ with the following symmetry properties:

$$\omega^s(x,-y) = -\omega^s(x,y), \quad \omega^a(x,-y) = \omega^a(x,y). \tag{1.16}$$

In summary, for the symmetric field $\boldsymbol{u}^s$, u, p are even and v, ω are odd in y, and opposite odd-even rules for the antisymmetric field $\boldsymbol{u}^a$.

This symmetry-based decomposition can be applied to the instantaneous flow field with:

$$\boldsymbol{u}(\boldsymbol{x},t) = \boldsymbol{u}^s(\boldsymbol{x},t) + \boldsymbol{u}^a(\boldsymbol{x},t). \tag{1.17}$$

When considering the modal interactions, for example the convective term, the group of symmetric and antisymmetric vector fields $(\boldsymbol{u}^s, \boldsymbol{u}^a)$ is shown to form a Z_2-group, which has the following symmetry properties:

$$\nabla\cdot\boldsymbol{u}^s\otimes\boldsymbol{u}^s, \quad \nabla\cdot\boldsymbol{u}^a\otimes\boldsymbol{u}^a \in \mathcal{U}^s, \tag{1.18a}$$
$$\nabla\cdot\boldsymbol{u}^s\otimes\boldsymbol{u}^a, \quad \nabla\cdot\boldsymbol{u}^a\otimes\boldsymbol{u}^s \in \mathcal{U}^a. \tag{1.18b}$$

We define a y-reflection operator R via $R(u,v,p,\omega)(x,y) \equiv (u,-v,p,-\omega)(x,-y)$. Eigenvectors obtained by linearizing about a reflection-symmetric state are necessarily either symmetric or antisymmetric. A pitchfork bifurcation occurs, if the eigenvalue is real and the eigenvector is antisymmetric, resulting in two symmetrically related asymmetric branches. If the eigenvalue is real and the eigenvector symmetric, a transcritical bifurcation occurs. A Hopf bifurcation comes along with a pair of complex conjugate eigenvalues. When the eigenvectors are antisymmetric, the resulting limit cycle satisfies the spatio-temporal symmetry $R\mathbf{u}(t) = \mathbf{u}(t+T/2)$ (Barkley, 2006). When the eigenvectors are symmetric, the limit cycle remains symmetric throughout $R\mathbf{u}(t) = \mathbf{u}(t)$.

Instabilities of the steady solutions

The linear stability analysis of the symmetric steady solutions $\boldsymbol{U}_s$ at different Reynolds numbers has been performed on a Krylov subspace of dimension 9 with 100 iterations. Two pairs of complex-conjugated eigenvalues, and a real eigenvalue can be found with positive real parts with increasing Reynolds numbers. The first pair of complex conjugated eigenvalues crosses the vertical axis ($\sigma = 0$) when the Reynolds number is changing from 18 to 19, see figure 1.8(left). It is associated with a Hopf bifurcation at $Re_1 = 18$. The second pair of complex conjugated eigenvalues crosses the vertical axis as the Reynolds number is changing from 64 to 65, see figure 1.8(middle). It is associated with a Hopf bifurcation at $Re_2 = 64$. The real eigenvalue becomes positive at $Re_3 = 68$, see figure 1.8(right) and is associated with a pitchfork bifurcation.

We plot the growth rate of the corresponding eigenmodes in figure 1.9 as functions of the Reynolds number. We notice the first complex-conjugated pair has the largest growth rate for $Re \leq 90$, and the real eigenvalue has the largest growth rate for $Re \geq 95$. For the two complex-conjugated pairs, the first pair has a larger growth rate for $Re \leq 105$. We recall that these growth rates only indicate the linear dynamics in the neighbourhood of the steady solution $\boldsymbol{U}_s$. As the instability develops, the linear dynamics no longer applies. The growth rate will decay to zero when reaching saturation, and the frequency can be also changed by the nonlinear interaction. We will discuss the nonlinear dynamics in the asymptotic regimes in the next section.

To discuss the eigenmodes associated with two of the three unstable steady solutions, we consider the case at $Re = 80$. As a result of the pitchfork bifurcation, there exist three steady solutions beyond $Re_3 \approx 68$: the symmetric steady solution $\boldsymbol{U}_s$, and two mirror-conjugated asymmetric steady solutions $\boldsymbol{U}_s^{\pm}$ with the base-bleeding jet deflected up or down, like in figure 1.18 for $Re = 100$. They are unstable with respect to the periodic vortex shedding due to the Hopf bifurcation beyond $Re_1 \approx 18$.

Figure 1.10(a) shows the eigenmodes associated with the two pairs of complex-conjugated eigenvalues (mode A and mode B) and with the real eigenvalue (mode C) from the linear stability

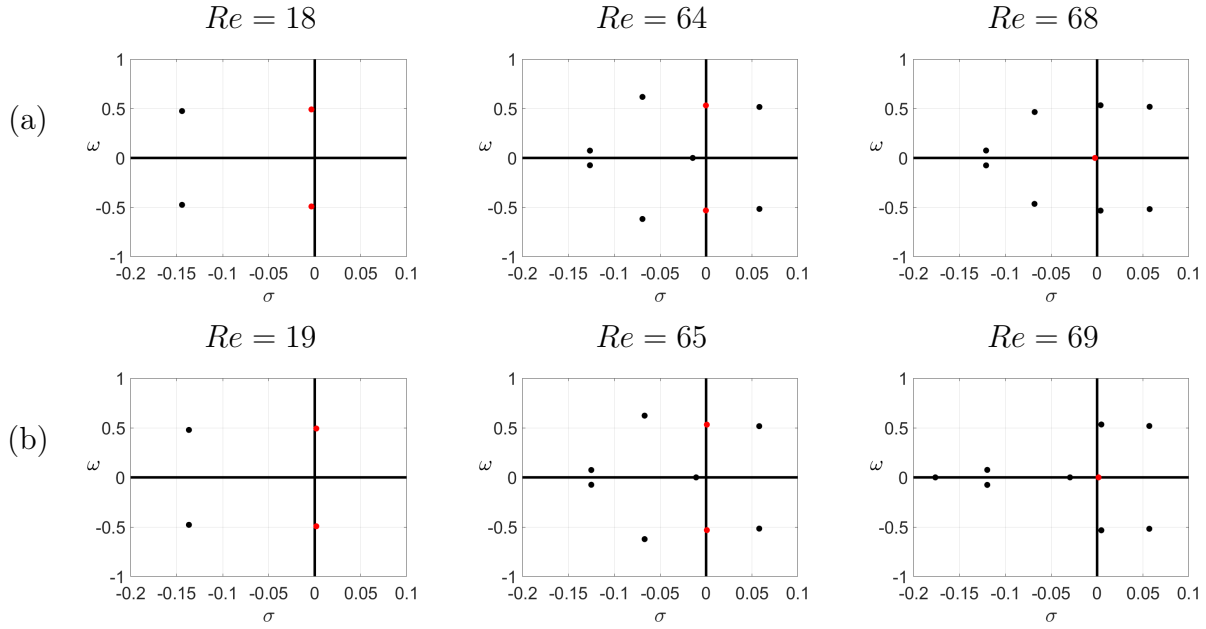


Figure 1.8: Eigenspectrum resulting from the linear stability analysis of the symmetric steady solution $\boldsymbol{U}_s$. With increasing Re, the eigenvalues cross the imaginary axis, as illustrated with the red points from (a) to (b). The first pair of complex-conjugated eigenvalues crosses the imaginary axis at the critical value of the first Hopf bifurcation $Re_1 \approx 18$. The second pair of complex-conjugated eigenvalues crosses the imaginary axis at the critical value of the second Hopf bifurcation $Re_2 \approx 64$. The real eigenvalue crosses the imaginary axis at the critical value of the pitchfork bifurcation $Re_3 \approx 68$.

analysis of the symmetric steady solution $\boldsymbol{U}_s$. As defined in Eq. (1.16), these three eigenmodes are reflection-antisymmetric. Mode C concentrates in the near wake behind the back two cylinders, which is associated with the symmetry-breaking instability of the base-bleeding jet. The super-critical pitchfork bifurcation breaks the reflection symmetry by deflecting the base-bleeding jet up or down. In contrast, Mode A and mode B originate in the far wake of the three cylinders, which can be taken as a whole and approximated by a single obstacle. These two modes are associated with the von Kármán instability of the vortex street.

These two pairs of complex-conjugated eigenvectors have very close angular frequencies. However, only one frequency can be observed in the asymptotic regime of a stable limit cycle for $Re \leq 105$. Mode B will quickly cancel out as mode A owns a larger growth rate. The far wake region will be dominated by the development of mode A. At $Re \geq 110$, mode B has a larger growth rate. The initial transient dyanmics is different, as exemplified in figure 1.4. The symmetric von Kármán vortex street at $Re = 80$, characterized by a staggered arrangement of the vortices, evolves into two spatially well-separated sub-streets of positive and negative vorticity at $Re = 130$.

Considering the linear stability analysis of the asymmetric steady solution $\boldsymbol{U}_s^-$, we can only find two pairs of complex-conjugated eigenvalues, associated with mode A^- and mode B^-, as shown in figure 1.10(b). Because the base flow $\boldsymbol{U}_s^-$ already breaks the reflection symmetry, all the eigenmodes are slightly deformed and asymmetric. In addition, the symmetry breaking of the base-bleeding jet leads to a larger growth rate for mode A and mode B.

Instabilities of the periodic solutions

The Floquet stability analysis integrates a small perturbation vector $\boldsymbol{q}$ over a given T-periodic base flow with the linearized governing equations (1.11). The starting point is the computation of the periodic solutions.

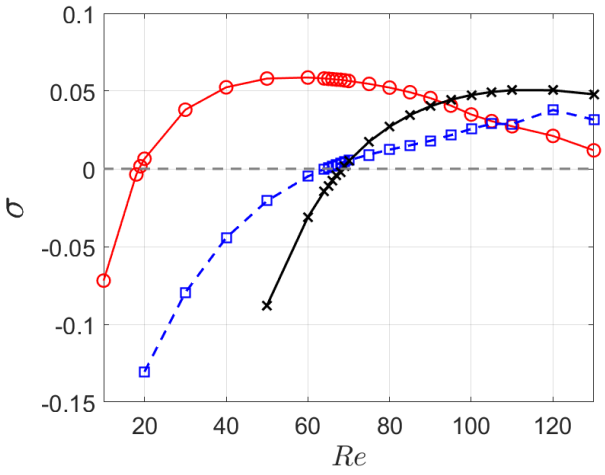


Figure 1.9: Evolution of the real part of the corresponding eigenvalues of the symmetric steady solutions $\boldsymbol{u}_s$ with the Reynolds number changing from 10 to 130. The red curve marked with circles corresponds to the first pair of complex-conjugated eigenvalues, the blue dashed curve marked with blocks corresponds to the second pair of complex-conjugated eigenvalues, and the black curve marked with crosses corresponds to the real eigenvalue.

For the periodic regime, a stable vortex shedding state is easy to obtain from the asymptotic state by direct numerical simulation. A large initial perturbation can accelerate the convergence to the attracting set. In our case, an unstable periodic solution preserving the symmetry is transiently observed in the DNS at $Re = 80$ starting close to the symmetric steady solution. Symmetry breaking occurs in a second stage and leads to a stable asymmetric periodic solution. To stabilize the symmetric periodic flow, we introduced a symmetric constraint on the base-bleeding jet. Details can be found in appendix 2. In our case, we directly obtain the base flow as a series of continuous snapshots from DNS by the same time interval. We note that a Fourier interpolation can be applied to reproduce the base flow at each time step by using a limited number of Fourier modes.

As an example, at $Re = 80$, we perform the Floquet stability analysis for both the unstable symmetric and the stable asymmetric periodic solutions, as detailed in Chapter 2. A simple power iteration method on a single vector can be also applied as long as the unstable eigenvector is unique and real. Here, we apply the block-Arnoldi method with random initial perturbation vectors (Shaabani-Ardali *et al.*, 2019), and compute the eigenvalues and eigenvectors in a high-dimensional Krylov subspace of 200-600 basis vectors. The same leading Floquet mode can be found with above mentioned methods, with a very close eigenvalue. Compared with the standard Arnoldi method, this algorithm can compute 20 vectors at each iteration, improving the construction efficiency of the Krylov subspace. The leading eigenvalue can converge after four iterations within a Krylov subspace with 80 vectors. The resulting Floquet multipliers are shown in figure 1.11. All the multipliers are inside of the unit circle for the stable asymmetric periodic solution, while only one real multiplier larger than 1 is found for the unstable symmetric periodic solution.

The resulting eigenvectors are the Floquet modes at instant t_0. We integrate Eq. (1.11) to obtain the periodic Floquet mode associated with the unstable multiplier at other instants, as shown in figure 1.12.

For quasi-periodic or chaotic regimes, stabilizing the flow on an unstable periodic solution is much more difficult. As shown in figure 1.5, the new emerging frequency first appears at $Re = 100$, but eventually dissipate and leads to a periodic solution. At $Re = 105$, this low-frequency does not resolve, indicating a stable quasi-periodic solution. Close to the critical value of the secondary Hopf bifurcation, there is a long transient regime of quasi-periodic dynamics before convergence to the periodic solution, as shown in figure 1.13. The low-frequency modulation slowly decreases with

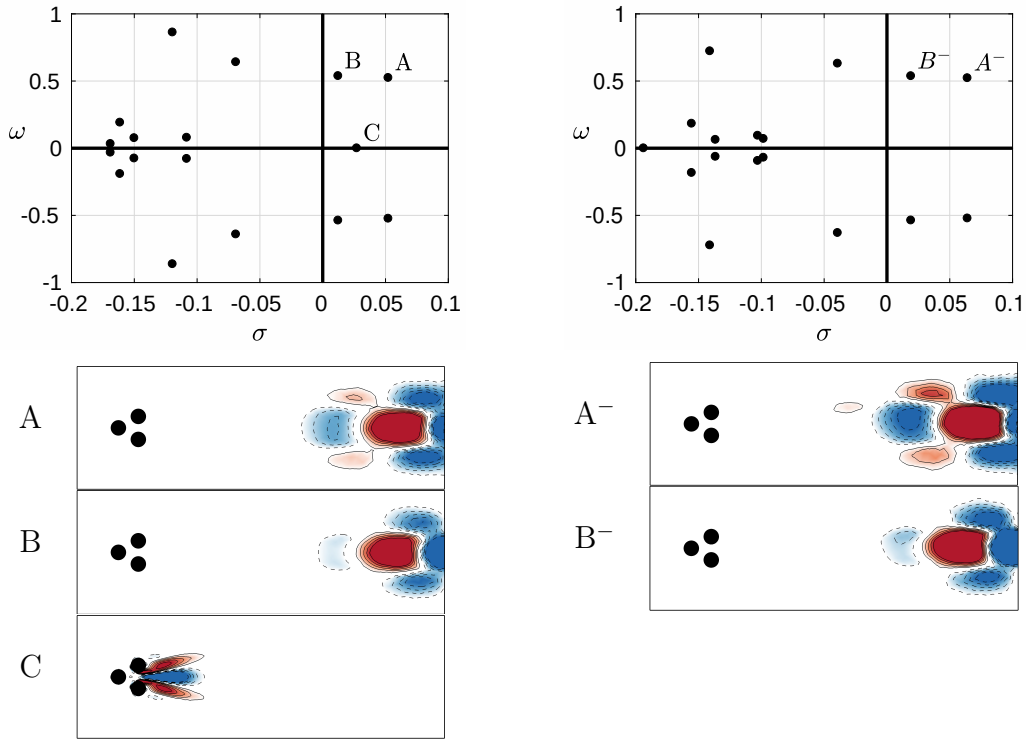


Figure 1.10: Eigenspectrum (top) and real part of the eigenvectors (bottom) of the symmetric steady solution $\boldsymbol{U}_s$ (left), of the asymmetric steady solution $\boldsymbol{U}_s^-$ (right), both at $Re = 80$. The red color and solid contours in the eigenvectors are positive values of the vorticity, blue color and dashed contours are negative values.

time for $100 \leq Re < 104$, indicating that the transient quasi-periodic solution is only transient and will become stable at $Re \geq 104$.

We perform the Floquet analysis on one asymmetric limit cycle at $Re = 100$. The leading Floquet multipliers form a complex-conjugate pair of eigenvalues close to the unit cycle, as illustrated in figure 1.14(a). Based on the same periodic base-flow, we slightly increase the Reynolds number to see the evolution of this pair of eigenvalues. Although it is not the exact periodic solution for the Reynolds number under consideration, the Floquet analysis can however predict the most dangerous modes. As shown in figures 1.14(b, c), the complex eigenvalues cross the unit cycle as the Reynolds number changes from 105 to 110.

The associated eigenvectors are also complex conjugated. We consider the real part of the eigenvector of the leading pair of complex-conjugate eigenvalues to determine the evolution of the

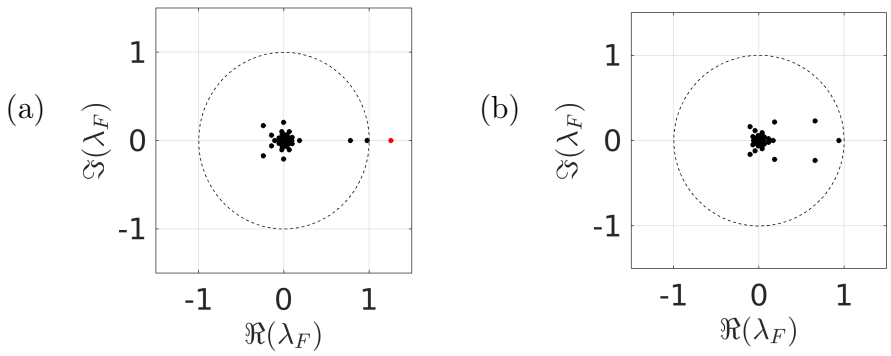


Figure 1.11: Floquet multipliers for the (a) symmetric periodic solution and (b) asymmetric periodic solutions at $Re = 80$. The red point indicates the leading Floquet multiplier associated with the unstable eigenvectors breaking the symmetry.

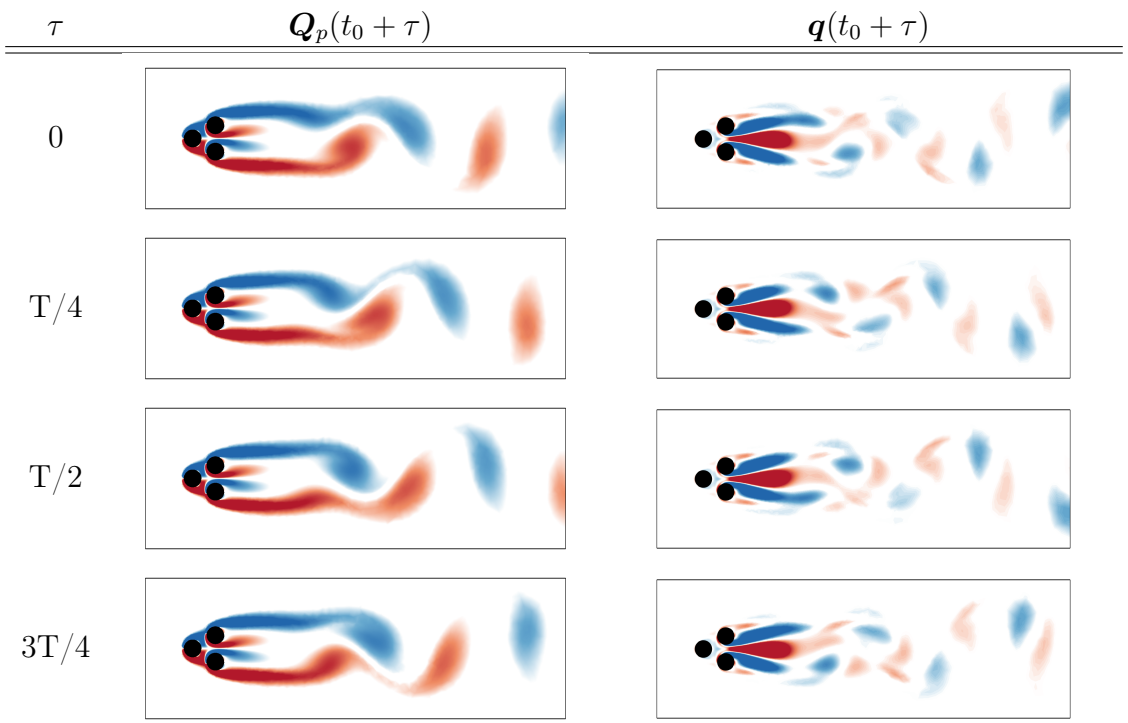


Figure 1.12: The T-periodic base flow $\boldsymbol{Q}_p(t_0+\tau)$ of the unstable symmetric periodic solution at $Re = 80$, together with the real part of the unstable eigenvector $\boldsymbol{q}(t_0+\tau)$ associated with the leading Floquet multiplier of figure 1.11(a). The spatio-temporal symmetry $\boldsymbol{u}(t+T/2) = R\boldsymbol{u}(t)$ implies that the next two instants are obtained by y-reflection and sign change (color reversal) of these vorticity fields.

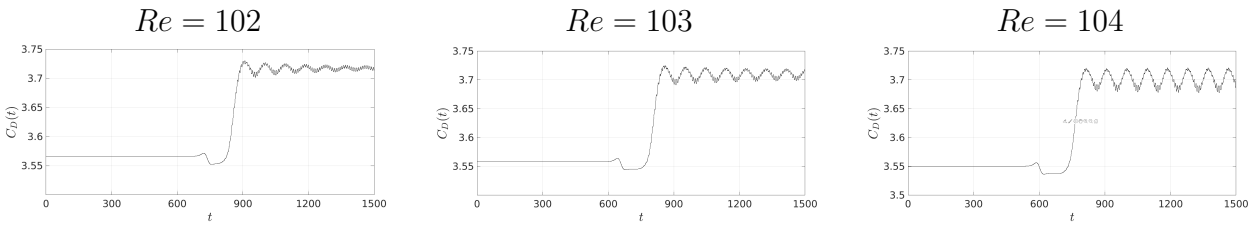


Figure 1.13: Transient dynamics of the drag C_D coefficients for the Reynolds numbers close to the critical value of the secondary Hopf bifurcation.

optimal perturbation on the periodic linear operator. We define the kinetic energy gain over a period as

$$G(n) = \frac{||e^{\mathcal{L}_{\boldsymbol{U}_p} nT}\boldsymbol{q}_0'||^2}{||e^{\mathcal{L}_{\boldsymbol{U}_p}(n-1)T}\boldsymbol{q}_0'||^2}, \tag{1.19}$$

where the norm is the inner product of the velocity field indicating the kinetic energy of the perturbation. The initial perturbation $\boldsymbol{q}_0'$ is optimal when considering the leading eigenmode from the Floquet analysis. The evolution of the optimal energy gain, over an integer number of periods, are illustrated in figure 1.15. A new period appears in the optimal gain with 11 periods of the base-flow. As shown in figure 1.16, the spatial structure of the optimal perturbation at instant t_0 will recover every 11 iterations with one periodic stepping on the base-flow. The vortex pairs are fixed with the phase of the base flow at instant t_0, but rotate with the new frequency.

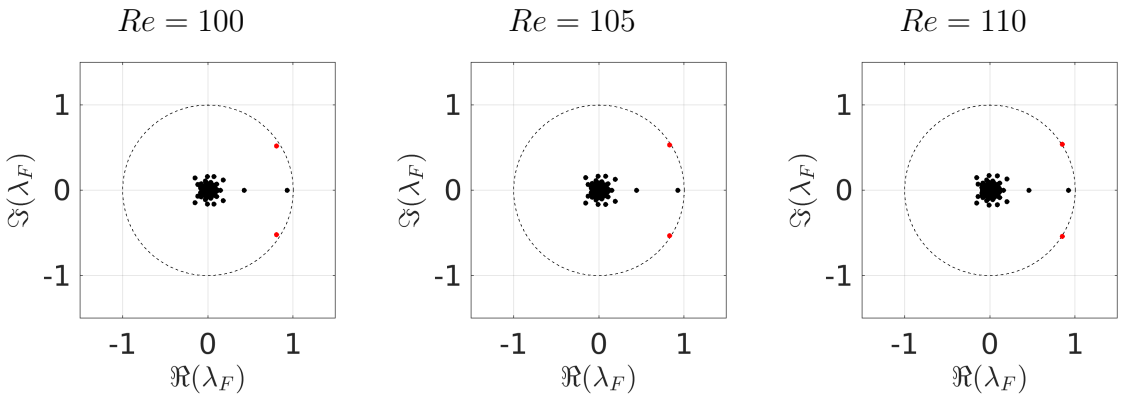


Figure 1.14: Evolution of the Floquet multipliers as the Reynolds number changes from 105 to 110, with the base-flow from the asymmetric periodic solution at $Re = 100$. The red points indicate the leading complex pair of the Floquet multipliers associated with the secondary Hopf bifurcation.

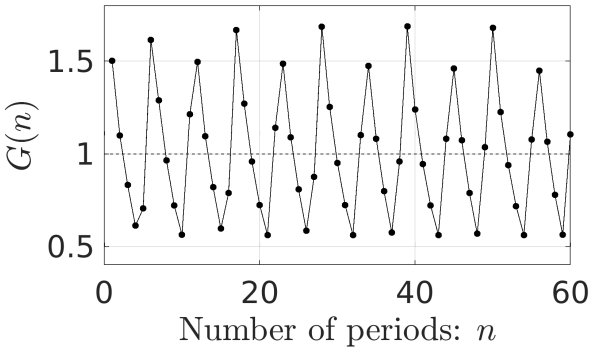


Figure 1.15: Evolution of the optimal energy gain over an integer number of periods of the asymmetric periodic solution at $Re = 100$.

1.4 Weakly nonlinear analysis

The linear dynamics only involve a minimal neighbourhood of the steady solution, as the perturbation $\boldsymbol{u}'$ is tiny at the onset of the instability. As the perturbation develops, the nonlinear term $\boldsymbol{u}' \otimes \boldsymbol{u}'$ cannot be ignored, and will start to modify the base flow through the effect of the Reynolds stress $\overline{(\boldsymbol{u}' \cdot \nabla)\boldsymbol{u}'}$. Considering the mean-field distortion from the steady solution $\boldsymbol{U}_s$ to the mean flow $\overline{\boldsymbol{U}}$, the original linear dynamics is no longer valid, and the nonlinear interactions will drive the transient dynamics until saturation.

Figure 1.17 shows the transient and post-transient dynamics starting from the steady solution of the cylinder flow. This oscillatory instability corresponds to a simple supercritical Hopf bifurcation, where a pair of conjugated oscillatory modes is enough to describe the fluctuating components. The nonlinear dynamics of this system can be projected into a three-dimensional space, and described by the temporal evolution of two oscillatory modes a_1, a_2 and a shift mode a_Δ associated with the mean-field deformation. During the evolution from the fixed point to the limit cycle, the conjugated oscillatory modes is constantly deformed and gets closer to the cylinder. The length of the recirculation bubble decreases continuously from the steady solution to the mean flow in the asymptotic regime. This deformation of the mean flow is the result of Reynolds stresses generated by the fluctuating field (Barkley, 2006).

In the following subsections, we will discuss the nonlinear interaction during the transient and post-transient dynamics.

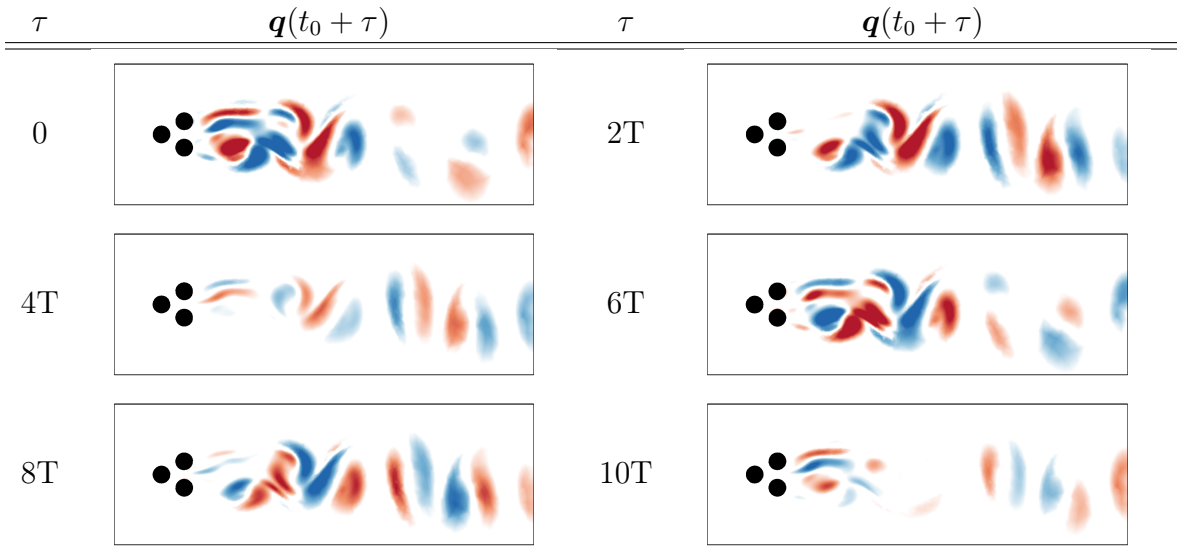


Figure 1.16: The real part of the unstable eigenvector $\boldsymbol{q}(t_0+\tau)$ associated with the leading pair of complex Floquet multipliers of figure 1.14(a). The eigenmode at instant t_0 presents rotating vortex pairs on a time scale of 11 base periods.

1.4.1 Symmetry breaking of the base flow

The supercritical pitchfork bifurcation at the critical value $Re_{\mathrm{PF}} \approx 68$ breaks the reflection symmetry of the base flow, the symmetric steady solution $\boldsymbol{U}_s(\boldsymbol{x})$, by deflecting the base-bleeding jet up or down. Two additional asymmetric steady solutions appear, namely $\boldsymbol{U}_s^+(\boldsymbol{x})$ and $\boldsymbol{U}_s^-(\boldsymbol{x})$, as shown in figure 1.18(right). Initialized with the three steady solutions at $Re = 130$, we gradually decrease the Reynolds number and determine the steady solutions separately. The three steady solutions converge to a unique solution at $Re = 68$, as shown with the lift coefficients of the steady solutions in figure 1.18(left).

To obtain three steady solutions

In our case with a steady solver, it is straightforward to compute the symmetric steady solution by solving the steady Navier-Stokes equations with a Newton-Raphson iteration method. One can start with a low Reynolds number $Re \ll Re_{\mathrm{PF}}$ first to ensure fast convergence to a symmetric steady solution, then use this steady solution as an initial condition to calculate the steady solution at an higher Reynolds number.

At Reynolds number $Re > Re_{\mathrm{PF}}$, the initial condition is critically important and decides which steady solution the steady solver converges to. Theoretically, if we have three solutions with different states of the base-bleeding jet, we can use them as the initial condition to obtain the corresponding three steady solutions. However, an initial condition closer to the target steady solution can ensure correct convergence. An option is to start with an asymmetric steady solution at another Reynolds number. It will quickly converge to the asymmetric steady solution at the given Reynolds number, sharing with the initial condition the same deflecting direction of the base-bleeding jet.

Another option is to start with an approximate steady state at the same Reynolds number, which is the closest solution to the target steady solution. We apply a numerical trick to obtain the approximate steady state by applying the unsteady solver with an over-large time stepping scale. When the time step is larger than the vortex shedding period, the coherent structure (the vortex shedding) is quickly damped and the flow converges to an approximation of the steady state. However, this artificially large time step will not affect the symmetry breaking of the base-bleeding jet. Taking this approximate state from the unsteady solver as an initial condition for the steady solver, an accurate asymmetric steady solution can be obtained. This method is still applicable

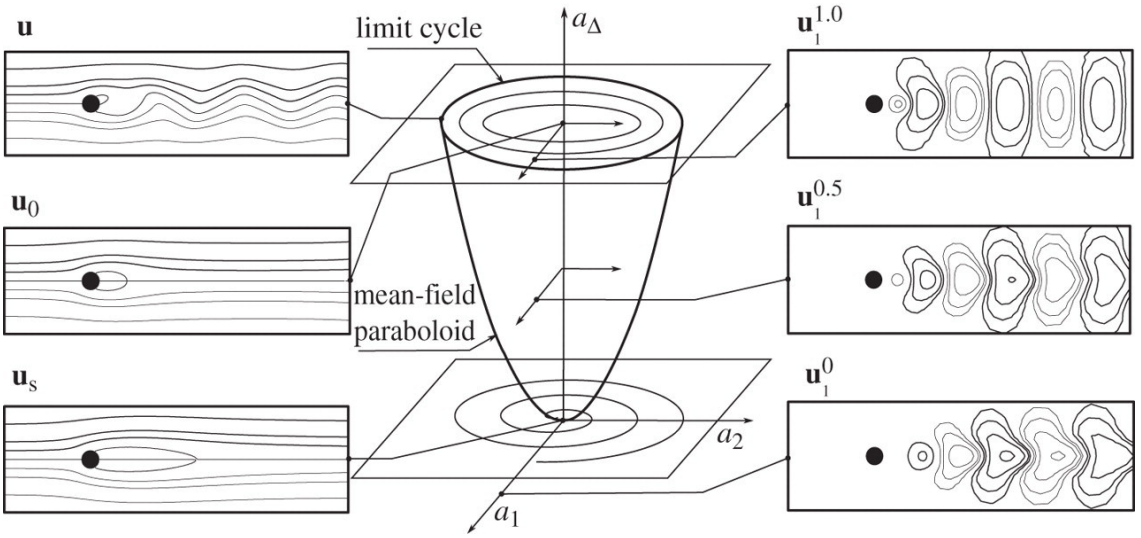


Figure 1.17: Nonlinear saturation during the transient and post-transient dynamics of the cylinder wake. On the left side: the streamline patterns of the steady solutions $\boldsymbol{u}_s(\boldsymbol{x})$, the time-averaged mean flow field $\boldsymbol{u}_0(\boldsymbol{x})$ and a snapshot of the vortex shedding on the limit cycle $\boldsymbol{u}(\boldsymbol{x},t)$. A pair of conjugated oscillatory modes are used to describe the local dynamics, which are varying along the manifold a_Δ. From Tadmor *et al.* (2011).

for cases with no prior knowledge of solutions at other Reynolds numbers.

Pitchfork bifurcation and symmetry breaking

At $Re > Re_{\mathrm{PF}}$, there exist three steady solutions satisfying the steady Navier-Stokes equations (1.5). Considering the following decomposition for a simple pitchfork bifurcation:

$$\boldsymbol{U}(\boldsymbol{x},t) = \langle \boldsymbol{U} \rangle (\boldsymbol{x},t) + \boldsymbol{u}_4(\boldsymbol{x},t), \tag{1.20}$$

where $\boldsymbol{u}_4(\boldsymbol{x},t)$ is the reflectional anti-symmetric component and $\langle \boldsymbol{U} \rangle (\boldsymbol{x},t)$ the symmetric component of the unsteady flow field. The pitchfork bifurcation breaks the symmetry of the flow and results in two asymmetric branches of opposite anti-symmetric component: one solution with $\boldsymbol{U}^+ = \langle \boldsymbol{U} \rangle + \boldsymbol{u}_4$, and another one with $\boldsymbol{U}^- = \langle \boldsymbol{U} \rangle - \boldsymbol{u}_4$. Hence, $\langle \boldsymbol{U} \rangle$ can be seen as the phase average of the two asymmetric solutions with $\langle \boldsymbol{U} \rangle = (\boldsymbol{U}^+ + \boldsymbol{U}^-)/2$.

Introducing a shift component $\boldsymbol{u}_5(\boldsymbol{x},t)$ for the slowly-varying dynamics of the mean flow, we can rewrite $\langle \boldsymbol{U} \rangle (\boldsymbol{x},t) \equiv \boldsymbol{U}_s(\boldsymbol{x}) + \boldsymbol{u}_5(\boldsymbol{x},t)$. We note that the subscripts of $\boldsymbol{u}_4(\boldsymbol{x},t)$ and $\boldsymbol{u}_5(\boldsymbol{x},t)$ are consistent with the index used in the least-order mean-field model introduced in 2. However, $\boldsymbol{u}_4(\boldsymbol{x},t)$ and $\boldsymbol{u}_5(\boldsymbol{x},t)$ represent here the time-resolved flow components instead of the spatial modes.

Introducing Eq. (1.20) into the Navier-Stokes equations (1.1), the equations can be separated into the anti-symmetric part and the symmetric part:

$$\partial_t \boldsymbol{u}_4 = -(\langle \boldsymbol{U} \rangle \cdot \nabla)\boldsymbol{u}_4 - (\boldsymbol{u}_4 \cdot \nabla) \langle \boldsymbol{U} \rangle + \nu \triangle \boldsymbol{u}_4 - \nabla p_4, \tag{1.21a}$$

$$(\boldsymbol{u}_4 \cdot \nabla)\boldsymbol{u}_4 = -(\langle \boldsymbol{U} \rangle \cdot \nabla) \langle \boldsymbol{U} \rangle + \nu \triangle \langle \boldsymbol{U} \rangle - \nabla \langle P \rangle , \tag{1.21b}$$

subject to the incompressibility condition. The original boundary conditions apply to $\boldsymbol{U}_s(\boldsymbol{x})$, and $\boldsymbol{u}_4(\boldsymbol{x},t)$, $\boldsymbol{u}_5(\boldsymbol{x},t)$ satisfy the homogeneous boundary conditions. Substituting $\langle \boldsymbol{U} \rangle (\boldsymbol{x},t) \equiv \boldsymbol{U}_s(\boldsymbol{x}) + \boldsymbol{u}_5(\boldsymbol{x},t)$ into the symmetric part of Eq. (1.21), the mean-field distortion $\boldsymbol{u}_5(\boldsymbol{x},t)$ and the fluctuating component $\boldsymbol{u}_4(\boldsymbol{x},t)$ must obey

$$(\boldsymbol{u}_4 \cdot \nabla)\boldsymbol{u}_4 = -(\boldsymbol{U}_s \cdot \nabla)\boldsymbol{u}_5 - (\boldsymbol{u}_5 \cdot \nabla)\boldsymbol{U}_s + \nu \triangle \boldsymbol{u}_5 - \nabla p_5. \tag{1.22}$$

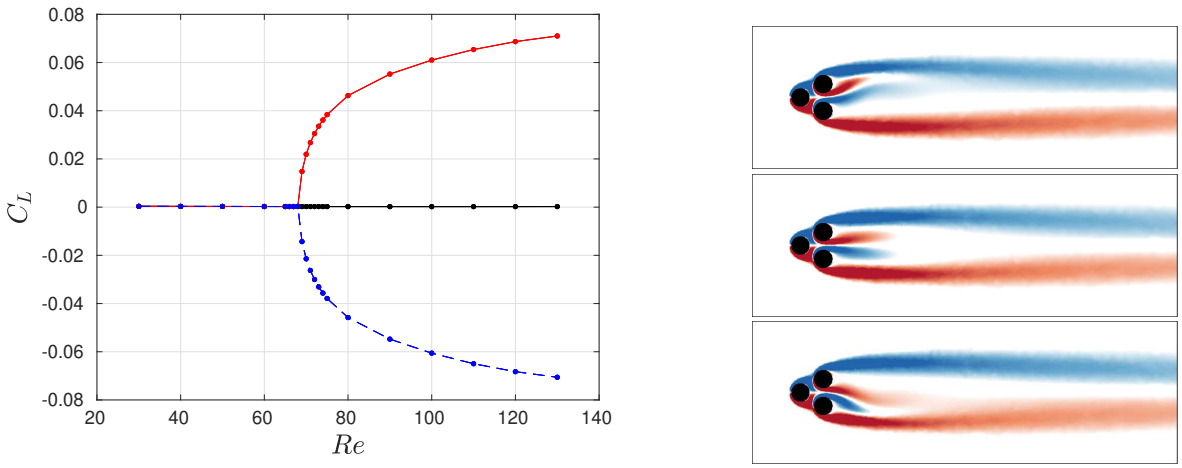


Figure 1.18: Lift coefficients at different Reynolds numbers (a) of the symmetric steady solutions $\boldsymbol{U}_s$ (black curve),the asymmetric steady solutions $\boldsymbol{U}_s^-$ (blue curve), the asymmetric steady solutions $\boldsymbol{U}_s^+$ (red curve), exemplified with the vorticity field of $\boldsymbol{U}_s^+$, $\boldsymbol{U}_s$, $\boldsymbol{U}_s^-$ at $Re = 100$ from top to bottom (b).

The nonlinear term $(\boldsymbol{u}_4 \cdot \nabla)\boldsymbol{u}_4$ corresponds to the Reynolds stress of the anti-symmetric part of the perturbations.

When $\partial_t \boldsymbol{u}_4 = 0$, the system reaches one of the two asymmetric steady solutions ($\boldsymbol{u}_4 \neq 0$). The nonlinear saturation yields:

$$(\boldsymbol{u}_5 \cdot \nabla)\boldsymbol{u}_4 + (\boldsymbol{u}_4 \cdot \nabla)\boldsymbol{u}_5 = -(\boldsymbol{U}_s \cdot \nabla)\boldsymbol{u}_4 - (\boldsymbol{u}_4 \cdot \nabla)\boldsymbol{U}_s + \nu \triangle \boldsymbol{u}_4 - \nabla p_4. \tag{1.23}$$

The right hand side of Eq. (1.23) can be written as $\mathcal{L}_{\boldsymbol{U}_s}\boldsymbol{u}_4(\boldsymbol{x}, t)$, which contributes to the linear instability of the symmetric steady solution. In this case, the growth of $\boldsymbol{u}_4(\boldsymbol{x}, t)$ has been saturated by the nonlinear interaction of $\boldsymbol{u}_4(\boldsymbol{x}, t)$ and $\boldsymbol{u}_5(\boldsymbol{x}, t)$ on the left-hand side.

1.4.2 Mean flow stability analysis

Strictly speaking, the time-averaged mean flow of a periodic state is not an exact solution of the Navier-Stokes equations. It is a statistical solution only and cannot reveal the stability of the physical system. The mean-field distortion is a direct consequence of the nonlinear saturation process. For the periodic flow resulting from the Hopf bifurcation, like in figure 1.17, the mean-flow field, initially merged with the steady solution, gets distorted due to the Reynolds stress associated with the fluctuations during the transient and post-transient dynamics. The linear stability analysis of the steady solution can only predict the instability at the onset of the instability. When the dynamics has saturated, the mean flow becomes marginally stable, with the growth rate close to zero. The frequency after nonlinear saturation is also different from that at the onset of the instability. However, the nonlinear frequency could be correctly captured by the stability analysis of the mean flow in the case of the cylinder wake (Barkley, 2006). Sipp & Lebedev (2007) further apply mean flow stability analysis to both cylindrical and open cavity flows and demonstrate that two precise conditions must be satisfied to obtain relevant and useful results. The theorical conditions for the use and meaning of this method is also discussed in Beneddine *et al.* (2016) on a turbulent backward facing step flow.

The Reynolds decomposition reads:

$$\boldsymbol{U}(\boldsymbol{x}, t) = \overline{\boldsymbol{U}}(\boldsymbol{x}) + \boldsymbol{u}'(\boldsymbol{x}, t), \tag{1.24}$$

with zero-mean fluctuating components $\boldsymbol{u}'$ and the mean-flow $\overline{\boldsymbol{U}}$ from time-averaging:

$$\overline{\boldsymbol{U}}(\boldsymbol{x}) = \frac{1}{T}\int_0^T \boldsymbol{U}_p(\boldsymbol{x}, t)dt. \tag{1.25}$$

Introducing Eq. (1.24) into Eq. (1.1) yields:

$$\begin{aligned} \partial_t \boldsymbol{u}' &= -(\boldsymbol{u}'\cdot\nabla)\boldsymbol{u}' - (\overline{\boldsymbol{U}}\cdot\nabla)\boldsymbol{u}' - (\boldsymbol{u}'\cdot\nabla)\overline{\boldsymbol{U}} - (\overline{\boldsymbol{U}}\cdot\nabla)\overline{\boldsymbol{U}} \\ &\quad +\nu\triangle\boldsymbol{U} + \nu\triangle\boldsymbol{u}' - \nabla\overline{P} - \nabla p'. \qquad (1.26a) \\ 0 &= \nabla\cdot\boldsymbol{u}', \qquad (1.26b) \end{aligned}$$

with the homogeneous boundary condition. Reynolds-averaged Navier-Stokes (RANS) equations read:

$$\overline{(\boldsymbol{u}'\cdot\nabla)\boldsymbol{u}'} = -(\overline{\boldsymbol{U}}\cdot\nabla)\overline{\boldsymbol{U}} + \nu\triangle\overline{\boldsymbol{U}} - \nabla\overline{P}, \qquad (1.27)$$

which indicates the Reynolds stress changes the time-averaged mean flow. Introduce Eq. 1.27 into Eq. 1.26:

$$\partial_t\boldsymbol{u}' = -(\overline{\boldsymbol{U}}\cdot\nabla)\boldsymbol{u}' - (\boldsymbol{u}'\cdot\nabla)\overline{\boldsymbol{U}} + \nu\triangle\boldsymbol{u}' - \nabla p' \underbrace{-(\boldsymbol{u}'\cdot\nabla)\boldsymbol{u}' + \overline{(\boldsymbol{u}'\cdot\nabla)\boldsymbol{u}'}}_{\boldsymbol{f}}. \qquad (1.28)$$

Only consider the first-order term and Eq. (1.28) writes

$$\partial_t\boldsymbol{q}' = \mathcal{L}_{\overline{\boldsymbol{U}}}\boldsymbol{q}', \qquad (1.29)$$

with $\mathcal{L}_{\overline{\boldsymbol{U}}}$ the linearized Navier-Stokes operator around the time-averaged mean flow and the solution in the form $\boldsymbol{q}'(\boldsymbol{x},t) = \hat{\boldsymbol{q}}(\boldsymbol{x})e^{(\sigma+i\omega)t}$.

The result owns a so-called RZIF property: the real part is zero, and the imaginary part is the frequency of the limit cycle (Turton *et al.*, 2015). For a stable limit cycle with angular frequency ω, the flow field can be decomposed into the temporal Fourier decomposition:

$$\boldsymbol{U}(\boldsymbol{x},t) = \overline{\boldsymbol{U}}(\boldsymbol{x}) + \sum_{n\neq 0}\boldsymbol{u}_n e^{in\omega t}. \qquad (1.30)$$

with $\overline{\boldsymbol{U}} > \ \boldsymbol{u}_{\pm n}$, $\boldsymbol{u}_{\pm n} \sim O(\epsilon^n)$ and $\boldsymbol{u}_{\pm n}$ a complex conjugate pair of n-th order harmonic. Introducing Eq. (1.30) into Eq. (1.1), the nonlinear terms are divided into different frequencies $\mathcal{N}_i = \sum_{n\neq 0}(\boldsymbol{u}_n\cdot\nabla)\boldsymbol{u}_{-n+i}$:

$$\begin{aligned} 0 &= -(\overline{\boldsymbol{U}}\cdot\nabla)\overline{\boldsymbol{U}} - \mathcal{N}_0 + \nu\triangle\overline{\boldsymbol{U}} - \nabla\overline{P}, \qquad (1.31a) \\ i\omega\boldsymbol{u}_1 &= -\mathcal{N}_1 + \nu\triangle\boldsymbol{u}_1 - \nabla p_1, \qquad (1.31b) \\ i2\omega\boldsymbol{u}_2 &= -\mathcal{N}_2 + \nu\triangle\boldsymbol{u}_2 - \nabla p_2, \qquad (1.31c) \\ &\cdots \qquad (1.31d) \end{aligned}$$

The non-harmonic term $\mathcal{N}_0$ can be seen as the decomposition of the Reynolds stress from different harmonic terms, which contributes to the mean-field distortion. For the first harmonic terms, assuming:

$$-(\overline{\boldsymbol{U}}\cdot\nabla)\boldsymbol{u}_1 - (\boldsymbol{u}_1\cdot\nabla)\overline{\boldsymbol{U}} \gg -(\boldsymbol{u}_2\cdot\nabla)\boldsymbol{u}_{-1} - (\boldsymbol{u}_{-1}\cdot\nabla)\boldsymbol{u}_2, \qquad (1.32)$$

Eq.(1.31b) can be written as

$$i\omega\boldsymbol{u}_1 = -(\overline{\boldsymbol{U}}\cdot\nabla)\boldsymbol{u}_1 - (\boldsymbol{u}_1\cdot\nabla)\overline{\boldsymbol{U}} + \nu\triangle\boldsymbol{u}_1 - \nabla p_1, \qquad (1.33)$$

with intriguingly $\sigma = 0$ and ω the angular frequency of the limit cycle. This property is a direct result of the harmonic balance (Dušek *et al.*, 1994). From the same idea, Mantič-Lugo *et al.* (2014) proposed the self-consistent model to predict the nonlinear saturation from $\boldsymbol{U}_s$ to $\overline{\boldsymbol{U}}$ for the supercritical Hopf bifurcation of the cylinder flow. The leading eigenmode of the slowly varying mean-flow field is proven to be able to predict the mean flow distortion. In the post-transient regime, the mean flow $\overline{\boldsymbol{U}}$ is marginally stable ($\sigma = 0$), and the leading eigenmode $\boldsymbol{u}_1$ and the angular frequency ω are closely associated with the limit cycle after saturation. However, the leading mode is not always sufficient to reproduce the nonlinear mean-field flow and frequency. Bengana & Tuckerman (2021) discusses the validity of the self-consistent model and proves the necessity of higher-order modes for a good approximation of the mean flow.

1.5 Deep mean-field modelling

In physics and statistics, mean-field theory, also known as self-consistent field theory, is a simplification of studying high-dimensional stochastic system by statistical averaging. The study of the mean flow can reveal some characteristics of the original system with a lower computational cost and has been applied to a wide range of scientific disciplines. In fluid mechanics, the mean flow usually refers to a time-averaged solution of the periodic flow, with the fluctuating components vanishing. Unlike the steady solution, the mean flow can be obtained both numerically and experimentally. The analysis of this solution has a broader range of application and may contribute to a deeper understanding of nonlinear interaction during the transition.

As mentioned in the previous section § 1.4.2, the mean flow stability analysis can successfully predict the nonlinear frequency of the saturated state. It turns out that the RZIF property is a direct result of the nonlinear saturation of oscillatory instability. Mean-field considerations also lead to the Reynolds decomposition of the flow field (1.34) and Reynolds-averaged Navier–Stokes equations (1.27). The self-consistent model proposed in Mantič-Lugo *et al.* (2014) uses a sole oscillatory mode to predict the nonlinear saturation dynamics of a supercritical Hopf bifurcation.

In this section, we will briefly discuss the flow under the mean-field consideration § 1.5.1, which leads to two kinds of mean-field modelling (MFM) strategies in the next two subsections. In § 1.5.2, a cluster-based MFM is introduced by incorporating the Reynolds decomposition into the cluster-based modelling. In § 1.5.3, we introduce the Galerkin framework under the mean-field consideration, which leads to least-order mean-field models based on the instabilities undergone by the system when increasing the Reynolds number. Based on the least-order mean-field model, we can reproduce the unsteady force dynamics in § 1.5.4 with a sparse formula of the drag and lift dynamics, simplified under the symmetry consideration for mean-field modelling, revealing the modes involved in the production of drag and lift.

1.5.1 Mean-field ansatz

Under the mean-field consideration, the triple decomposition of the flow field is similar to Reynolds & Hussain (1972)

$$\boldsymbol{u}(\boldsymbol{x},t) = \underbrace{\langle \boldsymbol{u}(\boldsymbol{x},t)\rangle_T}_{\omega \ll \omega_c} + \underbrace{\tilde{\boldsymbol{u}}(\boldsymbol{x},t)}_{\omega \sim \omega_c} + \underbrace{\boldsymbol{u}'(\boldsymbol{x},t)}_{\omega \gg \omega_c}, \tag{1.34}$$

where the dominant angular frequency ω_c is defined as the dominant peak in the Fourier spectrum of the velocity field. Here, the velocity field is decomposed into a slowly-varying mean-flow field $\langle \boldsymbol{u}\rangle_T$, a coherent component on time-scales of order $2\pi/\omega_c$, involving coherent structures $\tilde{\boldsymbol{u}}$, and a remaining (supposedly) non-coherent small scale fluctuations $\boldsymbol{u}'$. This kind of decomposition can also be found in the low-order Galerkin models of Tadmor *et al.* (2011) and in the weakly nonlinear modelling of Rigas *et al.* (2017*a*).

The slowly-varying mean-flow field $\langle \boldsymbol{u}\rangle_T$ can be defined as the average of the velocity field $\boldsymbol{u}$ over one local period $T \approx 2\pi/\omega_c$ of the coherent structures,

$$\langle \boldsymbol{u}(\boldsymbol{x},t)\rangle_T := \frac{1}{T}\int_{t-T/2}^{t+T/2} d\tau\, \boldsymbol{u}(\boldsymbol{x},\tau), \tag{1.35}$$

which eliminates both the coherent contribution from $\tilde{\boldsymbol{u}}$ and the non-coherent contribution from $\boldsymbol{u}'$. Unlike the mean-flow field defined by the post-transient limit,

$$\bar{\boldsymbol{u}}(\boldsymbol{x}) = \lim_{T\to\infty}\frac{1}{T}\int_0^T \boldsymbol{u}(\boldsymbol{x},\tau)d\tau, \tag{1.36}$$

the finite-time averaged-flow field considered in this study owns a slowly varying dynamics. From the mean-field theory of Stuart (1958), the slowly-varying mean-flow field evolves out of the steady solution under the action of the Reynolds stress associated with the most unstable eigenmode(s). The mean-flow field deformation $\boldsymbol{u}_\Delta$ is used to describe the difference between the slowly-varying mean-flow field and the invariant steady solution $\boldsymbol{u}_s(\boldsymbol{x})$, which reads

$$\langle \boldsymbol{u}(\boldsymbol{x},t)\rangle_T = \boldsymbol{u}_s(\boldsymbol{x}) + \boldsymbol{u}_\Delta(\boldsymbol{x},t). \tag{1.37}$$

The mean-flow deformation comes from the effect of the Reynolds stresses of the coherent component $\overline{(\tilde{\boldsymbol{u}}\cdot\nabla)\tilde{\boldsymbol{u}}}$ and of the non-coherent small scale fluctuations $\overline{(\boldsymbol{u}'\cdot\nabla)\boldsymbol{u}'}$.

The most simple mean-field model was proposed by Landau (1944) and Stuart (1958) for a nonlinear oscillating dynamics in the cylinder flow, near the threshold of the supercritical Hopf bifurcation. The weakly nonlinear evolution of the underlying instability can be modelled with the Stuart-Landau equation, which reads:

$$\frac{dA}{dt} = \sigma A + \alpha |A|^2 A, \tag{1.38}$$

where A is a complex quantity describing the fluctuation, for instance, the oscillation amplitude. The complex growth rate σ can be determined from the linear instability analysis, and the Landau constant α can be obtained from the asymptotic regimes with the nonlinear saturation characterised by a stable limit cycle. The Stuart-Landau equation reveals deep insights into the coupling between the fluctuations and the mean flow, e.g. the damping mechanism of unstable modes by Reynolds stress. The Landau constant can also indicate the type of Hopf bifurcation, with negative values for supercritical criticality and positive values for subcritical criticality. The Stuart-Landau equation has been widely used for flows undergoing supercritical Hopf bifurcations, and even for more complicated flows with multiple instabilities. Barkley *et al.* (2000) proposed Landau-like amplitude equations to describe the complex transition process for the three-dimensional vortex shedding in the wake of a circular cylinder. Fabre *et al.* (2008); Meliga *et al.* (2009); Rigas *et al.* (2017*a*) derived similar equations based on symmetry arguments for the wake of axisymmetric bodies, for instance, a sphere and a disk.

1.5.2 Hierarchical clutser-based framework

Automated reduced-order modelling is one of the most challenging and exciting directions for complex nonlinear dynamics. The traditional modelling techniques, like Galerkin modelling, always project the original system onto a low-dimentional subspace, and model an approximate dynamical system with an optimal basis. The projection basis decides the closeness of the approximate dynamics to the full dynamics. Data-driven modelling can liberate us from the issue of choosing the projection basis, and provide us with novel and promising modelling strategies from mathematics, data science, and statistical physics. Nowadays, machine learning can be seen as the most popular multi-disciplinary scientific area. It has been applied to data analysis and automated modelling in industry, transport, socio-economics, finance, biology and many other fields.

In this subsection, we briefly overview the machine learning techniques applied in our work: clustering, Markov chain and network science. The critical point is the hierarchical modelling strategy, which presents great consistency with the Reynolds decomposition and shows its great potential for multi-scale and multi-frequency modelling.

Cluster-based reduced-order modelling

Clustering is one of the most popular unsupervised machine learning techniques, which can automatically explore the natural groups (clusters) of the data in the feature space. The most common k-means algorithm partitions the data into k clusters, and features each state with the nearest centroid $\boldsymbol{c}_k$.

For fluid flows, the data are usually sampled into time-resolved snapshots of the flow field $\boldsymbol{u}^m(\boldsymbol{x}) = \boldsymbol{u}(\boldsymbol{x}, t^m)$, continuously sampled at times $t^m = m\Delta t$, $m = 1, \ldots, M$, with the time step Δt. The clustering algorithm will find the optimal locations of the K centroids, determined as solutions of a minimisation problem of the within-cluster distances between the snapshots and the related centroids, as described in section 5.3.1. The M snapshots will be partitioned into K clusters according to the Euclidean distance to the nearest centroid in the state space. Snapshots in the same cluster share the identical attributes featured by the cluster centroid. Thereby M snapshots are kinematically compressed to K centroids with clustering. The clustering process is unsupervised with only a sole control paremeter K.

The cluster-based reduced-order modelling (CROM) has been proposed by Kaiser *et al.* (2014) and exemplified on the mixing layer and the Ahmed body wake. The flow data is partitioned into 10 clusters by the k-means++ algorithm, and the flow dynamics is described from a statistical point of view with a cluster-based Markov model (CMM). CROM provides us with a novel modelling strategy, liberating us from the issue of choosing low dimensional spaces for the traditional projection methods.

Clutser-based Markov model and network model

The main difference between cluster-based Markov models (CMM) and cluster-based network models (CNM) is the cluster-based modelling process, while the kinematic compression with clustering is identical. CMM concentrates on stochastic modelling, providing a probability description of the state and the probability evolution over time. In contrast, CNM focuses on dynamical modelling, which can achieve the flow dynamics reconstruction by the state propagation with a sequence of visited clusters over time.

Clutser-based Markov model (CMM) provides a stochastic model to describe the transient dynamics between clusters using a Markov chain. In probability theory and statistics, the Markov property assumes the memoryless property of a stochastic process, which means that the current state depends on the previous state only.

As the M snapshots have been divided into K clusters, they are labeled with a cluster index $k^m = k$, $k = 1, \ldots, K$. The total number of snapshots in $\mathcal{C}_j$ is n_j, $i, j = 1, \ldots, K$. There exist $M - 1$ one-step transitions for M snapshots. By noting n_{ij} the number of snapshots departing from $\mathcal{C}_j$ to $\mathcal{C}_i$, the probability of transition from $\mathcal{C}_j$ to $\mathcal{C}_i$ is defined as $P_{ij} = n_{ij}/n_j$, and all the possible transitions are identified with the cluster transition matrix $\mathbf{P}$.

The state at time t^m is described by a cluster probability vector $\boldsymbol{p}^m = [p_1^m, \ldots, p_K^m]$, with probabilities p_k^m for the state in $\mathcal{C}_k$. The time evolution of the cluster probability vector is a Markov process, which can be described as

$$\boldsymbol{p}^{m+1} = \mathbf{P}\,\boldsymbol{p}^m. \tag{1.39}$$

Let $\boldsymbol{p}^1$ be the initial probability distribution at time t^1, the cluster probability vector at time t^m is computed by the time-step iteration,

$$\boldsymbol{p}^m = \mathbf{P}^{m-1}\,\boldsymbol{p}^1. \tag{1.40}$$

The CMM is the first CROM in the cluster-based framework, which results into a stochastic model of probability distribution. Nair *et al.* (2019) applied an extended CMM to the nonlinear feedback flow control of turbulent post-stall separated flows and introduced the directed network (Newman, 2018) for the dynamical modelling. The Markov chain is considered as a random walk on a directed network with the clusters being the nodes and the transitions between clusters being the edges. The within-cluster transitions are ignored under this consideration. For the non-trivial transitions between two different clusters, we re-define n_j as the total number of departing

snapshots from $\mathcal{C}_j$, with $n_j = \sum_{i=1}^{K} (1 - \delta_{ij}) n_{ij}$. The transition probability can be represented by:

$$P_{ij} = \frac{(1 - \delta_{ij}) n_{ij}}{n_j}, \quad i, j = 1, \cdots, K. \tag{1.41}$$

This extended CMM eliminates the self-loops residing in clusters and focuses on the non-trivial transitions, which significantly contributes to the transition dynamics.

Cluster-based network models (CNM), as proposed by Fernex *et al.* (2021) and Li *et al.* (2021), further develop the extended CMM by using a direct transition matrix for the non-trivial transitions, and a transition time matrix for the within-cluster transitions. The transition matrix $\mathbf{P}$ still acts as a propagator in terms of probability, which predicts the evolution of the cluster probability distributions over time. The residing time matrix $\mathbf{T}$ records the transition times for all the possible transitions in $\mathbf{P}$. In order to reproduce the transient dynamics for the original system, a sequence of visited clusters over time $[k_0, k_1, \ldots, k_N]$ are needed to trigger the cluster transitions. The transition times can be synthesized as

$$\hat{t}_n = t_0 + \sum_{i=0}^{n} T_{k_{i+1} k_i}, \quad n = 0, \ldots, N - 1. \tag{1.42}$$

Linear interpolation between two clusters can be introduced for a uniform state propagation between two clusters, as in Fernex *et al.* (2021) and Li *et al.* (2021). The flow state is represented by the nearest centroid without any additional assumption.

Hierarchical network model

Clustering is optimal for the spatial distributions of the centroids but not for the dynamics. The classical CROM may face difficulties with multi-scale problems in transient and post-transient dynamics because the clustering results are highly dependent on the spatial distances in the state space. The critical enabler for applying CNM to the transient and post-transient dynamics is the triple decomposition under the mean-field consideration, as in Eq. (1.34). By decomposing the flow field into a hierarchy of components, we can systematically model the dynamics at different scales.

A self-supervised hierarchical clustering is performed from top to bottom, which prioritizes the modelling of global trends for the mean flow distortion, then refines the local dynamics with sub-clusters, as described in § 5.3.2. Both the global trends and the local structure during the transition can be well preserved by a fewer number of clusters in the hierarchical structure, which leads to a better understanding of the physical mechanisms involved in the flow dynamics. In summary, the hierarchical cluster-based reduced-order model (HiCROM) inherits the excellent recognition performance of CROM and provides a generalized modelling strategy for complex dynamics with multiple scales and frequencies.

The cluster-based hierarchical network model (HiCNM) presented in Chapter 5 is based on the hierarchical Markov model of Fine *et al.* (1998), which introduces the hierarchical structure into the standard Markov model for stochastic modelling. In a HiCNM, each state of the parent layer will be considered separately and a new network model of its sub-states will be built in the child layer.

1.5.3 Galerkin framework

The POD-Galerkin method is based on the Galerkin projection of the governing equations onto an optimal low-dimensional basis, consisting of the spatial modes from the proper orthogonal decomposition (POD). This method has a long history as a data-driven reduced-order modelling technique and has been widely used for model reduction in fluid dynamics. However, the truncation errors,

stability, transient behavior and robustness for varying operating conditions are still challenging for this method. The choice of the low-dimensional subspace is critically important and will decide the representation of the flow dynamics. In this work, we use a least-order mean-field model to represent the flow dynamics with the fewest number of degrees of freedom. The basic modes of the low-dimensional space are chosen with the knowledge of the instabilities that the system has undergone while increasing the Reynolds number. Based on an optimal low-dimensional basis, a general Galerkin framework can be applied as following:

The Galerkin method is based on an inner product in the space of the square-integrable vector fields $\mathcal{L}^2(\Omega)$ in the observation domain Ω. The standard inner product between two velocity fields $\boldsymbol{u}(\boldsymbol{x})$ and $\boldsymbol{v}(\boldsymbol{x})$ reads:

$$(\boldsymbol{u}, \boldsymbol{v})_\Omega := \int_\Omega d\boldsymbol{x}\, \boldsymbol{u}(\boldsymbol{x}) \cdot \boldsymbol{v}(\boldsymbol{x}). \tag{1.43}$$

A traditional Galerkin approximation decompose the velocity field into a constant mode $\boldsymbol{u}_0$ and a fluctuating contribution of N orthonormal expansion modes $\boldsymbol{u}_i(\boldsymbol{x})$, $i = 1, \ldots, N$ with time-dependent amplitudes $a_i(t)$.

$$\boldsymbol{u}(\boldsymbol{x}, t) = \boldsymbol{u}_0(\boldsymbol{x}) + \sum_{i=1}^{N} a_i(t)\boldsymbol{u}_i(\boldsymbol{x}). \tag{1.44}$$

Generally, the time-averaged flow $\overline{\boldsymbol{u}}$ is chosen as the basic mode for a data set of the saturated state near an attractor. However, when considering the transient and post-transient dynamics, it is better to choose an exact solution, for instance, the steady solution $\boldsymbol{U}_s$ of the Navier-Stokes equations, corresponding to a fixed point of the Galerkin system in the following derivation. The basic mode $\boldsymbol{u}_0$ is constant with $a_0 \equiv 1$ (Rempfer & Fasel, 1994*b*). All the modes in the Galerkin expansion (1.44) satisfy the incompressibility condition. For the boundary conditions, the basic mode inherits the same boundary conditions for the original flow fields (1.4), the fluctuating modes satisfy the homogeneous boundary condition. The orthonormality condition for the fluctuating modes reads $(\boldsymbol{u}_i, \boldsymbol{u}_j)_\Omega = \delta_{ij}, \quad i, j \in \{1, \ldots, N\}$.

A Galerkin projection of the Navier-Stokes equation (1.1) onto the modes(1.44) leads to the following linear-quadratic ordinary differential equations (ODE), known as the Galerkin system (Fletcher, 1984),

$$\frac{d}{dt}a_i = \nu \sum_{j=0}^{N} l_{ij}^{\nu} a_j + \sum_{j,k=0}^{N} q_{ijk}^{c} a_j a_k + \sum_{j,k=0}^{N} q_{ijk}^{p} a_j a_k, \tag{1.45}$$

with coefficients

$$l_{ij}^{\nu} = (\boldsymbol{u}_i, \triangle \boldsymbol{u}_j)_\Omega\,, \tag{1.46a}$$
$$q_{ijk}^{c} = (\boldsymbol{u}_i, \nabla \cdot \boldsymbol{u}_j \otimes \boldsymbol{u}_k)_\Omega\,, \tag{1.46b}$$
$$q_{ijk}^{p} = (\boldsymbol{u}_i, -\nabla p_{jk})_\Omega\,, \tag{1.46c}$$

for the viscous, convective and pressure terms in the Navier-Stokes equations (1.1), respectively. The pressure term vanishes for sufficiently large domains (Noack *et al.*, 2005) and is neglected in the following.

When the steady solution is taken as the basic mode $\boldsymbol{u}_0 = \boldsymbol{U}_s$ with $a_0 \equiv 1$, the projection of the steady Navier-Stokes equation (1.5) onto the modes (1.44) leads to the constant relations $\nu l_{i0}^{\nu} + q_{i00}^{\nu} + q_{i00}^{p} = 0$. Exclude $a_0 \equiv 1$ from the state vector, and $\boldsymbol{a} = 0$ is a fixed point of the following Galerkin system:

$$\frac{d}{dt}a_i = \sum_{j=1}^{N} l_{ij} a_j + \sum_{j,k=1}^{N} q_{ijk} a_j a_k, \tag{1.47}$$

where $l_{ij} = \nu l_{ij}^{\nu} + q_{ij0}^{c} + q_{i0j}^{c}$ and $q_{ijk} = q_{ijk}^{c}$ for $i, j, k \in \{1, \ldots, N\}$.

Consider for illustration an oscillatory instability under mean-field considerations. The flow can be decomposed into the coherent fluctuation $\tilde{\boldsymbol{u}}$ and the distorting mean-flow $\boldsymbol{u}^D(\boldsymbol{x},t) = \langle \boldsymbol{u}(\boldsymbol{x},t)\rangle_T$. The mean-field deformation $\boldsymbol{u}_\Delta$ is well approximate by the difference between the distorted mean-flow and the steady solution as $\boldsymbol{u}^D(\boldsymbol{x},t) = \boldsymbol{U}_s(\boldsymbol{x}) + \boldsymbol{u}_\Delta(\boldsymbol{x},t)$.

Consider the symmetry of those flow components. The steady solution, distorted mean-flow and mean-field deformation can be expected to be symmetric while the dominant fluctuation is anti-symmetric. Therefore, the anti-symmetric part of the Navier-Stokes equations drives the dynamics of the fluctuation field:

$$\partial_t \tilde{\boldsymbol{u}} + \nabla \cdot \left[\boldsymbol{u}^D \otimes \tilde{\boldsymbol{u}} + \tilde{\boldsymbol{u}} \otimes \boldsymbol{u}^D\right] = \nu \triangle \tilde{\boldsymbol{u}} - \nabla \tilde{p}. \tag{1.48}$$

Analogously, the symmetric part describes the dynamics of the distorted mean-flow:

$$\partial_t \boldsymbol{u}_\Delta + \nabla \cdot [\boldsymbol{U}_s \otimes \boldsymbol{u}_\Delta + \boldsymbol{u}_\Delta \otimes \boldsymbol{U}_s + \boldsymbol{u}_\Delta \otimes \boldsymbol{u}_\Delta + \tilde{\boldsymbol{u}} \otimes \tilde{\boldsymbol{u}}] = \nu \triangle \boldsymbol{u}_\Delta - \nabla p_\Delta \tag{1.49}$$

Next, consider $\tilde{\boldsymbol{u}}$ and $\boldsymbol{u}_\Delta$ as small perturbations around the fixed point $\boldsymbol{U}_s$. Let $\tilde{\boldsymbol{u}} \in O(\varepsilon)$ and $\boldsymbol{u}_\Delta \in O(\delta)$ where ε and δ are smallness parameters. Hence, $\boldsymbol{u}_\Delta \otimes \boldsymbol{u}_\Delta \in O(\delta^2)$ can be neglected in comparison to the $O(\delta)$ terms $\boldsymbol{U}_s \otimes \boldsymbol{u}_\Delta$, $\boldsymbol{u}_\Delta \otimes \boldsymbol{U}_s$. We follow Stuart's original idea to separate the fluctuation $\tilde{\boldsymbol{u}}$ driven by the instability and the resulting mean-field deformation $\boldsymbol{u}_\Delta$ and arrive at the unsteady linearized Reynolds equation,

$$\partial_t \boldsymbol{u}_\Delta + \nabla \cdot [\boldsymbol{U}_s \otimes \boldsymbol{u}_\Delta + \boldsymbol{u}_\Delta \otimes \boldsymbol{U}_s + \langle \tilde{\boldsymbol{u}} \otimes \tilde{\boldsymbol{u}} \rangle] = \nu \triangle \boldsymbol{u}_\Delta - \nabla p_\Delta. \tag{1.50}$$

The mean-field deformation $\boldsymbol{u}_\Delta$, characterized by the scale δ, is seen to respond linearly to the Reynolds stress force $-\nabla \cdot \langle \tilde{\boldsymbol{u}} \otimes \tilde{\boldsymbol{u}} \rangle$ scaling with ϵ^2. Hence, $\delta \propto \epsilon^2$.

In a least-order mean-field model, the Galerkin approximation consists of a basic mode $\boldsymbol{u}_0(\boldsymbol{x}) = \boldsymbol{U}_s(\boldsymbol{x})$, a pair of conjugated modes $\boldsymbol{u}_1(\boldsymbol{x})$, $\boldsymbol{u}_2(\boldsymbol{x})$ for the oscillatory fluctuation, and a shift mode $\boldsymbol{u}_3(\boldsymbol{x}) \propto \boldsymbol{u}_\Delta(\boldsymbol{x})$ for the Reynolds-stress effect of the fluctuation on the mean-field deformation. After projection onto the Navier-Stokes equations, it yields the mean-field Galerkin system with harmonic balance on the oscillatory coefficients $a_1(t)$, $a_2(t)$ and the slowly varying coefficient $a_3(t)$. Moreover, a slaving relation $a_3 \propto (a_1^2 + a_2^2)$ leads to the mean-field manifold.

We notice that the Galerkin system has a sparse form under symmetry considerations, as defined in subsection 1.3.2. For each $d\,a_i/dt$, the linear-quadratic ordinary differential equation has 5 linear terms and 15 quadratic terms. However, only half of them are preserved under symmetry assumptions, as shown in figure 1.19.

The Galerkin system, directly derived from the Galerkin projection, cannot capture the transient and post-transient dynamics correctly. It introduces extremely long transients due to the underestimated growth rate. Our strategy is to correct some of the system coefficients by the knowledge provided by the linear stability analysis and the asymptotic dynamics. The remaining coefficients of the nonlinear interaction terms are identified by a sparse regression algorithm. The details can be found in Chapter 2.

1.5.4 Galerkin force model

In this work, we also provide systematic investigations and interpretations of the aerodynamic forces in the Galerkin framework. Here is a brief overview of the motivation and the critical derivation.

Considering a reduced-order model (ROM) that can achieve an accurate reconstruction of the flow fields, a natural thought arises: *can this ROM reproduce the instantaneous forces on the body?* Intuitively, the answer should be yes. One can calculate the body forces directly from the reconstructed flow fields. Since the aerodynamic forces depend on the viscous and pressure forces on the body, the prediction of the forces needs the ROM to reconstruct the velocity field and the pressure field simultaneously. In general, a POD-Galerkin system only provides a reconstruction of the

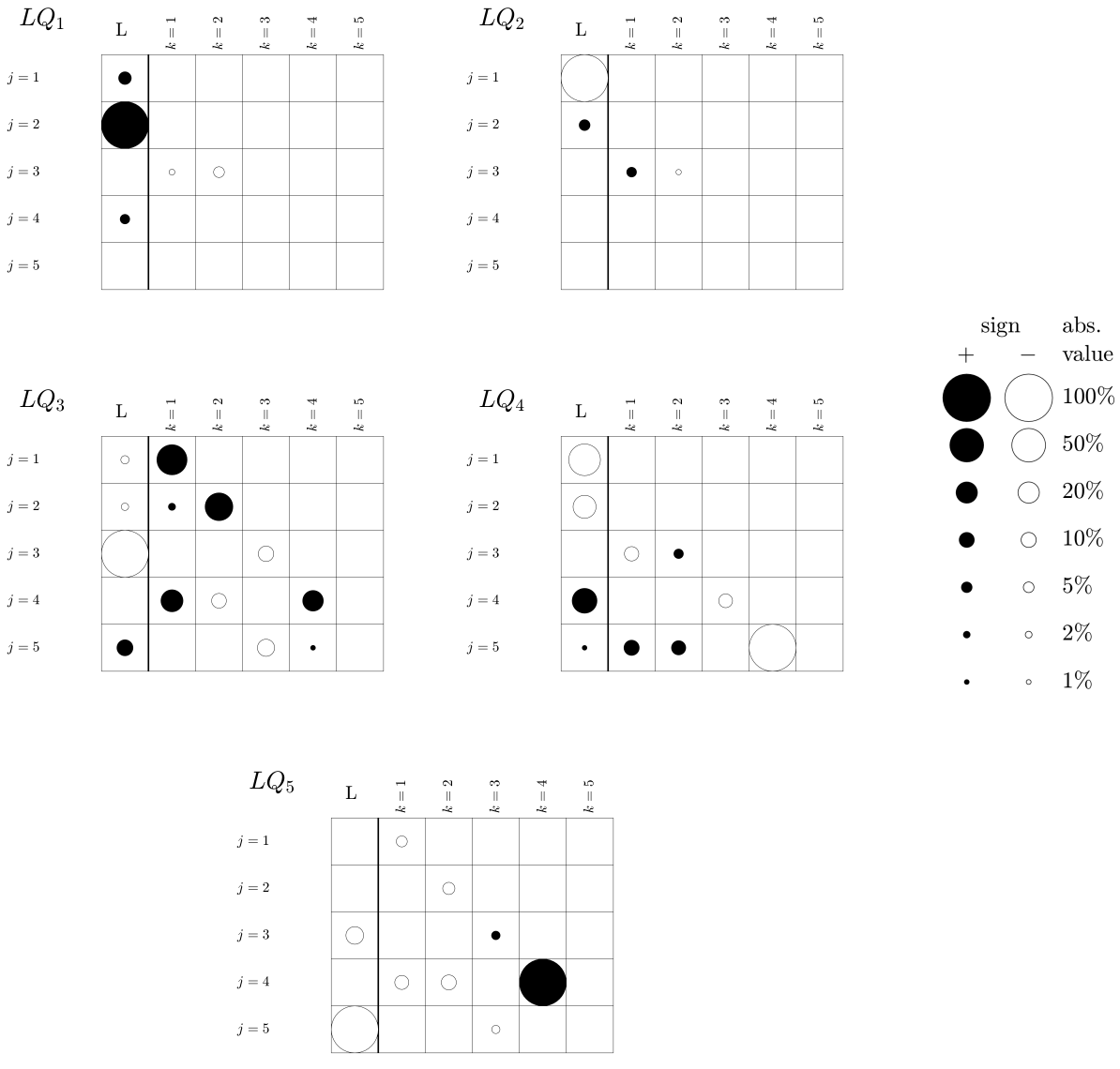


Figure 1.19: The active terms in the linear-quadratic ODE from the Galerkin projection. The considered Galerkin system consists of five degrees of freedom at $Re = 80$, and the details can be found in section 2.3.5. The positive coefficients are presented with filled circles and open circles for the negative coefficients. Their size has been normalized to the largest coefficient for each $\frac{d}{dt}a_i$, with $i = 1, \ldots, 5$ in the subscript of LQ.

velocity field. The pressure field reconstruction is still challenging for both computation cost and boundary issues. There are two different approaches to achieve the pressure reconstruction from the velocity-based ROM. One is to decompose the pressure field in the same Galerkin expansion as the velocity field, then assumes that the velocity expansion and pressure expansion share the same temporal coefficients. Another one is to solve the pressure Poisson equation with suitable boundary conditions; but this approach needs considerable effort to solve all the possible modal interactions between every two modes.

The critical idea of both two approaches is to present the pressure force with the velocity-based ROM. Obviously, the second approach has a theoretical basis, but is troublesome numerically. From the same idea, Noca *et al.* (1999) proposed force equations in an incompressible flow without the need to evaluate explicitly the pressure field and is only based on the knowledge of the velocity field and its time derivative. Liang & Dong (2014) applied it to the velocity-based POD modes, and derived a force expression in terms of the force associated with each individual POD mode and the force associated with the interaction between the POD modes. It showed that the leading six POD modes are enough to predict the drag force with 5% error.

POD modes are orthonormalized spatial structures and therefore their force contribution is constant. The same is true for the forces from the interactions between every two POD modes. Therefore, the body force can be derived as a constant-linear-quadratic function of the mode amplitudes.

The lift and drag forces can be derived from the general Galerkin expansion 1.44 of the velocity field $\boldsymbol{u}(\boldsymbol{x},t)$ as a constant-linear-quadratic function of the mode amplitudes $a_i(t)$. For an body immersed in the flow domain Ω, the α-component F_α^ν ($\alpha = x, y$) of the viscous force vector $\boldsymbol{F}^\nu$ on the boundary Γ is expressed by

$$F_\alpha^\nu = \boldsymbol{F}^\nu \cdot \boldsymbol{e}_\alpha = 2\nu \oint_\Gamma \sum_{\beta=x,y} S_{\alpha,\beta}\, n_\beta \, dS \,, \tag{1.51}$$

where $\boldsymbol{n}$ is the unit normal pointing outward the surface element dS, $\boldsymbol{e}_\alpha$ is the unit vector in the α-direction and $S_{\alpha,\beta} = (\partial_\alpha u_\beta + \partial_\beta u_\alpha)/2$ is the strain rate tensor with indices $\alpha, \beta = x, y$.

The pressure force in the α-direction is expressed as

$$F_\alpha^p = \boldsymbol{F}^p \cdot \boldsymbol{e}_\alpha = -\oint_\Gamma dS\, n_\alpha p. \tag{1.52}$$

The drag and lift forces are defined as the projection on $\boldsymbol{e}_x$ and $\boldsymbol{e}_y$ of the pressure and viscous forces exerted on the body

$$F_D(t) = F_x^p(t) + F_x^\nu(t), \tag{1.53a}$$
$$F_L(t) = F_y^p(t) + F_y^\nu(t). \tag{1.53b}$$

The velocity field $\boldsymbol{u}(\boldsymbol{x},t)$ is replaced by the Galerkin approximation (1.44). The viscous force (1.51) can be re-written as

$$F_\alpha^\nu = \sum_{j=0}^{N} q_{\alpha;j}^\nu a_j, \tag{1.54}$$

where $q_{\alpha;j}^\nu$ can easily be derived from (1.51) with the corresponding $S_{\alpha,\beta}$ of the velocity mode $\boldsymbol{u}_j$, with the form

$$q_{\alpha;j}^\nu = 2\nu \oint_\Gamma \sum_{\beta=x,y,z} S_{\alpha,\beta}(\boldsymbol{u}_j)\, n_\beta \, dS. \tag{1.55}$$

Similarly, from the pressure Poisson equation

$$\nabla^2 p = \nabla \cdot (-\nabla \cdot \boldsymbol{u} \otimes \boldsymbol{u}) = -\sum_{\alpha=x,y} \sum_{\beta=x,y} \partial_\alpha u_\beta \partial_\beta u_\alpha, \tag{1.56}$$

the expression of the pressure field is derived as

$$p(\boldsymbol{x},t) = \sum_{j,k=0}^{N} p_{jk}(\boldsymbol{x})\, a_j(t)\, a_k(t), \tag{1.57}$$

with

$$\nabla^2 p_{jk} = \nabla\cdot(-\nabla\cdot\boldsymbol{u}_j\otimes\boldsymbol{u}_k) = -\sum_{\alpha=x,y}\sum_{\beta=x,y}\partial_\alpha u_\beta(\boldsymbol{u}_j)\,\partial_\beta u_\alpha(\boldsymbol{u}_k). \tag{1.58}$$

The homogeneous Neumann boundary conditions for partial pressure field p_{jk} are needed to solve the above Poisson equation (Noack *et al.*, 2005). Integrating (1.52) with (1.57) shows that the pressure force is a quadratic polynomial of the a_j's

$$F_\alpha^p = \sum_{j,k=0}^{N} q_{\alpha;jk}^p a_j a_k, \quad \text{where} \quad q_{\alpha;jk}^p = -\oint_\Gamma dS\, n_\alpha p_{jk}. \tag{1.59}$$

The total force can be expressed as a constant-linear-quadratic expression in terms of the mode coefficients

$$F_\alpha = F_\alpha^\nu + F_\alpha^p = c_\alpha + \sum_{j=1}^{N} l_{\alpha;j} a_j + \sum_{j,k=1}^{N} q_{\alpha;jk} a_j a_k, \tag{1.60}$$

where

$$c_\alpha = q_{\alpha;0}^\nu + q_{\alpha;00}^p, \qquad l_{\alpha;j} = q_{\alpha;j}^\nu + q_{\alpha;j0}^p + q_{\alpha;0j}^p, \qquad q_{\alpha;jk} = q_{\alpha;jk}^p. \tag{1.61}$$

The drag and lift formula can be simplified under symmetry considerations and reveal the modes that contribute to the draf and the lift Liang & Dong (2015). A second simplification is performed with a sparse calibration of the remaining coefficients. The sparsity parameter λ penalizes any non-vanishing term and yields sparse human-interpretable expressions for the force model. The detailed work can be found in Chapter 4.

1.6 Structure of the thesis

This thesis aims to facilitate the automated reduced-order modelling for a sparse human-interpretable model using first principles and machine learning techniques. More specifically, we focus on the complex wake dynamics generally found in fluid mechanics and develop different strategies for mean-field modelling. A key enabler for constructing a mean-field model is the mean-field assumption, where slowly-varying mean-field deformations are due to the fluctuating field through the Reynolds stress, resulting in a Reynolds-like decomposition. The mean-field models proposed in this study provide a challenging benchmark example for automatable reduced-order modelling.

In this introductory chapter, we addressed the motivation of reduced-order modelling and provided an overview of the principles and methodologies for the cluster-based and projection-based mean-field modelling strategies:

(a) **Cluster-based mean-field modelling :** The clustering and modelling are performed based on mean-field considerations. A hierarchical structure is incorporated from theoretical considerations into the data-driven modelling. The clustering achieves a kinematic compression of the input data, featured by statistical averages of the grouping data (centroids) in the original data space. The data is labelled with a cluster index, which automatically groups the states in the state space without approximation. The identified dynamics can preserve the structure of the original system depending on the cluster distribution. The hierarchical model can systematically identify the multi-scale dynamics, including the transitions between different solutions, the bifurcating dynamics into different attractors, and the local structures for the distorted mean flow in time.

(b) **Projection-based mean-field modelling :** The original system is projected onto a linearly independent low-dimensional subspace, and the nonlinear dynamics will be discussed in the chosen subspace. The basis of the low-dimensional representation is essential for the approximate dynamics, reflecting the stability, transient behaviour and robustness of the identified model. The mean-field assumption leads to a constrained Galerkin system, where the degrees of freedom are optimal with respect to the underlying instabilities in the system. The knowledge from both the stability analysis and the saturated states makes it possible to determine some coefficients of the mean-field model and increases the robustness of the sparse Galerkin model by only identifying the remaining coefficients by a sparse regression.

(c) **Projection-based force modelling :** The force modelling is based on the projection of the mathematical force expression onto a linearly independent low-dimensional subspace. It can be generalized for any other Galerkin model without mean-field assumption. However, the modal symmetry and slaving relations in a mean-field model can introduce additional sparsity and improve the interpretability of the model, identified by sparse Galerkin regression.

All the above-mentioned modelling strategies are exemplified for the transient and post-transient dynamics of the unforced fluidic pinball at different Reynolds numbers. The flow undergoes two successive supercritical Hopf and pitchfork bifurcations, respectively associated with the von Kármán instability of vortex shedding for $Re > 18$ and the symmetry-breaking instability of the base-bleeding jet for $Re > 68$. The two projection-based modelling strategies are applied to the dynamics between six exact Navier-Stokes solutions for the two successive bifurcations, at $68 < Re < 104$. The cluster-based modelling can be much easier to apply as a self-supervised data-driven method. We do not need to find an optimal basis of low-dimensional space for the underlying instabilities. For $104 < Re < 115$, a lower newly introduced frequency modulates the oscillations of the base-bleeding jet around its deflected position in a quasi-periodic regime. At $Re > 115$, the base-bleeding jet randomly switches up and down in a chaotic regime with more complex nonlinear interaction. The cluster-based modelling strategy is applied to the transient and post-transient dynamics of the fluidic pinball in three distinct dynamical regimes : a periodic regime characterized by four unstable and two stable exact solutions of the Navier-Stokes equations at $Re = 80$, a quasi-periodic regime at $Re = 105$ and a chaotic regime at $Re = 130$.

The remaining part of this thesis has been organised as follows:

In Chapter 2, the numerical plant of the fluidic pinball is introduced. The flow behaviours at different Reynolds numbers are numerically investigated. The linear stability analysis of the steady solutions and the Floquet stability analysis of the periodic solutions are performed for the verification of the first two bifurcations. A mean-field Galerkin model can be derived for the primary Hopf bifurcation and the secondary pitchfork bifurcation. The Galerkin system identification is based on the sparse quadratic regression with physics-based constraints. The least-order model is only five-dimensional but can reproduce the key features of the transient and post-transient behaviours of the full dynamics.

In Chapter 3, we discuss the coinciding local bifurcations of the steady and periodic Navier-Stokes solutions found in the fluidic pinball at $Re > Re_2$. The coincidence of two local bifurcations is intriguing since it is non-generic. Such a non-generic coincidence is modelled and explained.

In Chapter 4, we derive an aerodynamic force model associated with a Galerkin model for the first two bifurcations of the fluidic pinball. The identification of the drag and lift formula is simplified by exploiting the modal symmetry under mean-field considerations and sparse calibration. Based on the least-order mean-field model, the basic Galerkin force models are derived for the supercritical Hopf bifurcation and for the supercritical pitchfork bifurcation. Next, the force models include the coupling between the elementary modes associated with both bifurcations. The advantages of this methodology are discussed in comparison to the purely projection-based approach and the POD-based regression model.

In Chapter 5, we propose a self-supervised hierarchical cluster-based reduced-order modelling methodology for bifurcations and topology identification of complex dynamics. We start with the

standard process of a cluster-based network model and introduce the hierarchical structure to enable the modelling of more complex transient dynamics. This methodology is a universal modelling strategy to identify the transient and post-transient dynamics in a self-supervised manner. The HiCNMs are successfully applied to the transient and post-transient dynamics of multiple trajectories at three Reynolds numbers: for the case involving six exact solutions at $Re = 80$, for the quasi-periodic case at $Re = 105$, and for the chaotic case at $Re = 130$, respectively. Moreover, it promises to automate the identification and analysis of complex dynamics with multiple attractors and multiple scales.

Chapter 6 concludes the results of this thesis and provides some perspectives for future work.

Chapter 2

Low-order model for successive bifurcations of the fluidic pinball

Nan Deng[1,2] , Bernd R. Noack[2,3,4], Marek Morzyński[5] and Luc R. Pastur[1]

[1] Institute of Mechanical Sciences and Industrial Applications, ENSTA-Paris, Institut Polytechnique de Paris, 828 Bd des Maréchaux, F-91120 Palaiseau, France

[2] LIMSI, CNRS, Université Paris-Saclay, Bât 507, rue du Belvédère, Campus Universitaire, F-91403 Orsay, France

[3] Institute for Turbulence-Noise-Vibration Interaction and Control, Harbin Institute of Technology, Shenzhen Graduate School, University Town, Xili, Shenzhen 518058, People's Republic of China

[4] Institut für Strömungsmechanik und Technische Akustik (ISTA), Technische Universität Berlin, Müller-Breslau-Straße 8, D-10623 Berlin, Germany

[5] Chair of Virtual Engineering, Poznań University of Technology, Jana Pawla II 24, PL 60-965 Poznań, Poland

Published in *Journal of Fluid Mechanics* (Deng *et al.* (2020)).

We propose the first least-order Galerkin model of an incompressible flow undergoing two successive supercritical bifurcations of Hopf and pitchfork type. A key enabler is a mean-field consideration exploiting the symmetry of the mean flow and the asymmetry of the fluctuation. These symmetries generalize mean-field theory, e.g. no assumption of slow growth-rate is needed. The resulting 5-dimensional Galerkin model successfully describes the phenomenogram of the fluidic pinball, a two-dimensional wake flow around a cluster of three equidistantly spaced cylinders. The corresponding transition scenario is shown to undergo two successive supercritical bifurcations, namely a Hopf and a pitchfork bifurcations on the way to chaos. The generalized mean-field Galerkin methodology may be employed to describe other transition scenarios.

Contents

2.1 Introduction

This study advances mean-field modelling for successive symmetry breaking due to Hopf and pitchfork bifurcations. The theoretical framework is applied to the transition of the flow around a cluster of circular cylinders, termed the *fluidic pinball* for the possibility to control fluid particle by cylinder rotation (Noack & Morzyński, 2017; Ishar *et al.*, 2019).

Mean-field theory was pioneered by Landau (1944) and Stuart (1958) and is a singular triumph of nonlinear reduced-order modelling in fluid mechanics. Already the most simple mean-field model for the supercritical Hopf bifurcation reveals deep insights into the coupling between the fluctuations and the mean flow, e.g. the damping mechanism of unstable modes by Reynolds stress. In addition, Malkus (1956) principle of marginal stability for time-averaged flows, the square root growth law of fluctuation level with increasing Reynolds number, the cubic damping term from a linear-quadratic dynamics, the energetic explanation of this amplitude dynamics, and the slaving principle leading to manifolds driven by ensemble-averaged Reynolds stress are easily derived. Also, the idea of center manifold theory and the surprising success of linear parameter-varying models are analytically illustrated. Historically, Landau was the first to derive the normal form of the dynamics with Krylov-Bogoliubov approximation (an averaging method for spiral phase paths, see *e.g.* Jordan & Smith, 1999) while Stuart could explain how the cubic damping term arises from the distorted mean flow.

Mean-field models for a supercritical Hopf bifurcation with an unstable oscillatory eigenmode have been applied and validated for numerous configurations. The onset of vortex shedding behind a cylinder wake has been thoroughly investigated (Strykowski & Sreenivasan, 1990; Schumm *et al.*, 1994; Noack *et al.*, 2003). Even high-Reynolds number turbulent wake flow can display a distinct mean-field manifold and modeled by a noise-driven mean-field model (Bourgeois *et al.*, 2013).

A supercritical pitchfork bifurcation similarly arises by an unstable eigenmode with a real eigenvalue. The onset of convection rolls in the Rayleigh-Bénard problem is a famous example (Zaitsev & Shliomis, 1971; Swift & Hohenberg, 1977; Cross & Hohenberg, 1993). The features

of a pitchfork bifurcation are observed for the sidewise symmetry breaking of the time-averaged Ahmed body wake (Grandemange *et al.*, 2012, 2013; Cadot *et al.*, 2015; Bonnavion & Cadot, 2018) and more generally in three-dimensional wake flows (Mittal, 1999; Gumowski *et al.*, 2008; Szaltys *et al.*, 2012; Grandemange *et al.*, 2014; Rigas *et al.*, 2014). In contrast, the drag crisis of circular cylinder is associated with a subcritical bifurcation into two asymmetric sheddings with opposite mean lift values (Schewe, 1983).

Not surprisingly, numerous generalizations of mean-field models have been proposed. Landau (1944) and Hopf (1948) have conjectured that high-dimensional fully developed turbulence may be explained by an increasingly rapid succession of Hopf bifurcations. This idea has been discarded as unlikely (see, for instance, Landau & Lifshitz, 1987). The seminal paper by Ruelle & Takens (1971) showed that turbulence does not arise as a successive superposition of oscillators, but irregular chaotic behavior can already appear after few bifurcations. A second direction is the explanation of nonlinear coupling between two incommensurable shedding frequencies (Luchtenburg *et al.*, 2009), also referred to as *frequency crosstalk* in the following. This amplitude coupling over the mean flow has been termed *quasi-laminar* in Reynolds & Hussain (1972) pioneering theoretical foundation of the triple decomposition. The advancements also include subcritical bifurcations (Watson, 1960). More specifically, the case of a codimension two bifurcation, involving both a pitchfork and a Hopf bifurcation, was addressed in Meliga *et al.* (2009), who derived the amplitude equation based on the weakly nonlinear analysis of the wake of a disk. Fabre *et al.* (2008) derived the same equation solely based on symmetry arguments for the wake of axisymmetric bodies. A resolvent analysis follows mean-field considerations in decomposing the flow in a time-resolved linear dynamics and a feedback-term with the quadratic nonlinearity (Gomez *et al.*, 2016; Rigas *et al.*, 2017*b*).

Our study develops a generalized mean-field Galerkin model for the first two bifurcations of the fluidic pinball with increasing Reynolds number. The primary supercritical bifurcation leads to the periodic vortex shedding which is statistically symmetric. At higher Reynolds numbers, the resulting limit cycle undergoes a pitchfork bifurcation into a stable, asymmetric, mirror-symmetric pair of periodic solutions. This local bifurcation has a transverse effect resulting from the decoupling of these two bifurcations (see appendix 2.E), which simultaneously leads to an identical local pitchfork bifurcation of the steady solution, into an unstable, asymmetric, mirror-symmetric pair of steady solutions. The underlying dynamics is modeled with a small number of assumptions. The key simplification results from exploiting the symmetry of the mean flow and the antisymmetry of the fluctuation. The generalized mean-field Galerkin methodology can be expected to be useful for describing other transition scenarios.

The manuscript is organized as follows. In § 2.2, the numerical plant is introduced and the Reynolds-number dependent flow behavior described. This phenomenology drives the mean-field modelling of the first two bifurcations in § 2.3. The resulting models for the Hopf and subsequent pitchfork bifurcation are present in § 2.4 and § 2.5, respectively. § 2.6 summarizes the results and outline future directions of research.

2.2 Flow configuration

In this section we describe the numerical toolkit and the flow features as the Reynolds number is increased. The direct Navier-Stokes solver with MATLAB interfaces, used for the simulation, is described in § 2.2.1. The fluidic pinball configuration and the flow features and route to chaos are described in § 2.2.2 and § 2.2.3, respectively.

2.2.1 Direct Navier-Stokes solver

The unsteady Navier-Stokes solver is based on fully implicit time integration and Finite-Element Method discretization (Noack & Morzyński, 2017; Noack *et al.*, 2003, 2016). The time integra-

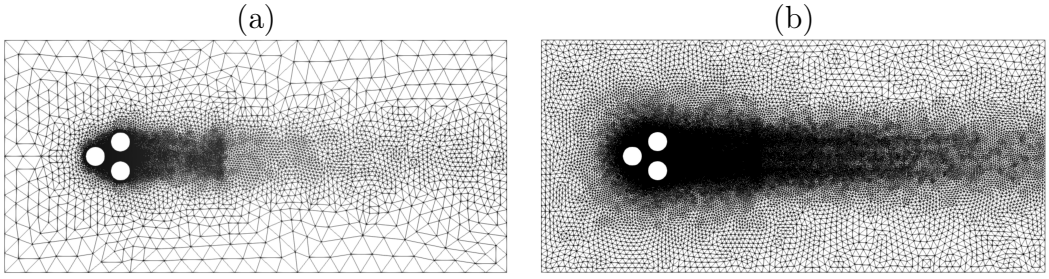


Figure 2.1: Computational grid for the fluidic pinball, with 8633 (a) and 54195 vertices (b).

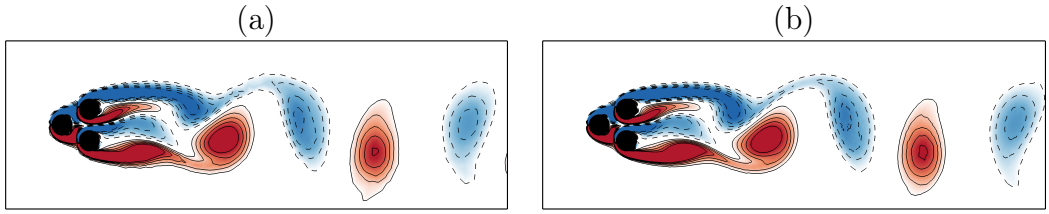


Figure 2.2: DNS computations on a grid with 8 633 nodes (a) and 54 195 nodes (b). Vorticity depicted with color is $[-1.5, 1.5]$, $t = 200$. The initial kick is provided by a rotation of all three cylinders at $t = 0.2$, see text for details.

tion is third-order accurate while FEM discretization employs a second–order Taylor-Hood finite elements (Taylor & Hood, 1973). The solution is obtained iteratively, with the Newton-Raphson type approach. The tangent matrix is updated on each iteration and computations are carried out until the residual is under a prescribed tolerance. The steady solution is obtained in a similar Newton-Raphson iteration for the steady Navier-Stokes equations. The convergence to one of the three steady solutions with different states of the base-bleeding jet is triggered by appropriate "initial" conditions in the iteration, see appendix 2.A. The solver quickly converges to one of the steady states and a final, near-zero residual confirms that this is indeed the steady flow solution sought. The computational domain is discretized on an unstructured grid. Pinball configuration uses a grid with 4225 triangles and 8633 vertices (see figure 2.1(a)). To test the grid dependency of the solution we increased the number of triangles by nearly a factor 4 (26849 elements and 54195 nodes, see figure 2.1(b)). The flow patterns shown in figure 2.2 develop from a steady solution at $Re = 100$ subjected to an instantaneous rotation of cylinders at $T = 0.2$. The upper cylinder rotates counterclockwise, the lower one clockwise and the center cylinder also in a clockwise direction — all with unit circumferential velocity, i.e. the velocity of oncoming flow U_∞. This configuration and boundary conditions result in a vortex shedding shown in figure 2.2 for the time instance $t = 200$. Both simulations prove grid independence and yield dynamically consistent results (see figure 2.2).

2.2.2 Pinball configuration

We refer to the configuration shown in figure 2.1 as the *fluidic pinball* as the rotation speeds allow one to change the paths of the incoming fluid just as flippers manipulate the ball of a conventional pinball machine. The fluidic pinball is a set of three equal circular cylinders with radius R placed parallel to each other in a viscous incompressible uniform flow at speed U_∞. The flow over a cluster of three parallel cylinders has been experimentally studied involving heat transfer, fluid-structure interactions and multiple frequencies interactions over the past few decades (Price & Paidoussis, 1984; Sayers, 1987; Lam & Cheung, 1988; Tatsuno *et al.*, 1998; Bansal & Yarusevych, 2017). For the fluidic pinball, the cylinders can rotate at different speeds creating a kaleidoscope of vortical structures or variety of steady flow solutions. The configuration is used for evaluation of flow controllers (Cornejo Maceda, 2017) as this problem is a challenging task for control methods

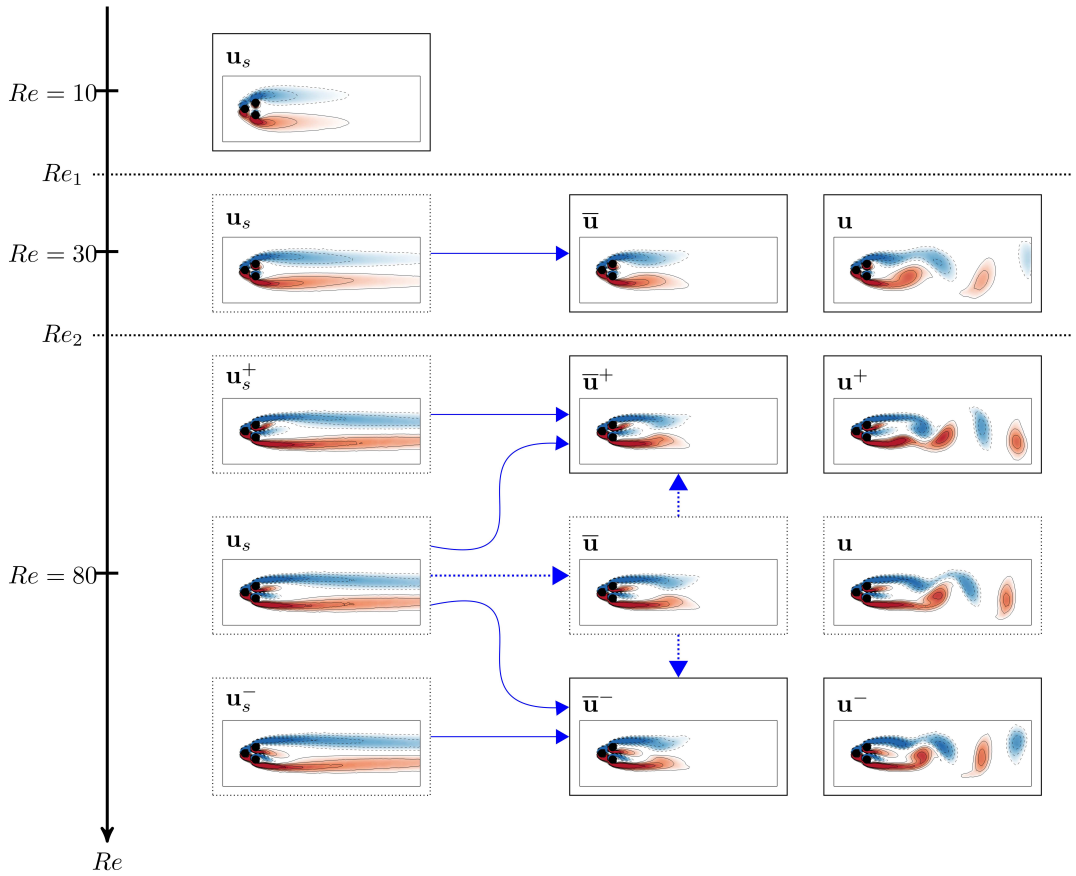


Figure 2.3: Flow states at different values of the Reynolds number: the stable states are labeled with a solid-line box and the unstable states with a dashed-line box. Steady solutions $\boldsymbol{u}_s(\boldsymbol{x})$ (on the left side): three steady solutions $\boldsymbol{u}_s(\boldsymbol{x})$ (symmetric), $\boldsymbol{u}_s^{\pm}(\boldsymbol{x})$ (asymmetric), exist at $Re = 80$. For $Re = 30$ and $Re = 80$, the steady solutions are unstable: the permanent regime $\boldsymbol{u}(\boldsymbol{x},t) = \overline{\boldsymbol{u}}(\boldsymbol{x}) + \boldsymbol{u}'(\boldsymbol{x},t)$ is unsteady with mean flow field $\overline{\boldsymbol{u}}(\boldsymbol{x})$. At $Re = 80$, the instantaneous flow field $\boldsymbol{u}(\boldsymbol{x},t) = \overline{\boldsymbol{u}}(\boldsymbol{x}) + \boldsymbol{u}'(\boldsymbol{x},t)$, transiently explored when starting close to $\boldsymbol{u}_s(\boldsymbol{x})$, is unstable with respect to either $\boldsymbol{u}^{\pm}(\boldsymbol{x},t) = \overline{\boldsymbol{u}}^{\pm}(\boldsymbol{x}) + \boldsymbol{u}'(\boldsymbol{x},t)$.

comprising several frequency crosstalk mechanisms (Noack & Morzyński, 2017). The centers of the cylinders form an equilateral triangle with side length $3R$, symmetrically positioned with respect to the flow. The leftmost triangle vertex points upstream, while the rightmost side is orthogonal to the oncoming flow. The origin of the Cartesian coordinate system is placed in the middle of the top and bottom cylinder. The fluidic pinball computational domain, shown in figure 2.1 is bounded by the rectangle $[-6, 20] \times [-6, 6]$.

Without forcing, the boundary conditions comprise a no-slip condition on the cylinders and a unit velocity in the far field:

$$U_r = 0 \text{ on the cylinders and } U_\infty = \boldsymbol{e}_x \text{ at infinity.} \tag{2.1}$$

The far field boundary conditions are exerted on the inflow, upper and lower boundaries while the outflow boundary is assumed to be a stress–free one, transparent for the outgoing fluid structures. A typical initial condition is the unstable steady Navier-Stokes solution $\boldsymbol{u}_s(\boldsymbol{x})$.

In this study, all three cylinders remain static as we are interested in the natural dynamics of the flow as the Reynolds number is increased.

2.2.3 Flow features

The steady solution $\boldsymbol{u}_s$, shown in figure 2.3 for different values of the Reynolds number Re, is stable up to the critical value $Re_1 \approx 18$. This value corresponds to $5/2 \times Re_1 \approx 45$ with respect to the actual body height $5R$, which is consistent with the critical value of the Reynolds number found for a single cylinder (Ding & Kawahara, 1999; Barkley, 2006). Beyond Re_1, the steady solution becomes unstable with respect to vortices periodically and alternately shed at the top and bottom of the two right-most cylinders, following a Hopf bifurcation (instability of the fixed point *via* a pair of complex-conjugated eigenvalues, see e.g. Strogatz *et al.* (1994)). In addition to the resulting von Kármán street of vortices, the gap between the cylinders makes possible the formation of a jet at the base of the two outer cylinders. The steady solution $\boldsymbol{u}_s$, the mean flow $\overline{\boldsymbol{u}}$ and the instantaneous flow field $\boldsymbol{u}$, are shown in figure 2.3, for $Re = 30$.

The flow passing between the two rearward cylinders, the base-bleeding flow, has a critical impact on the successive bifurcations undergone by the system on the route to chaos. Indeed, beyond a secondary critical value $Re_2 \approx 68$ of the Reynolds number, the system undergoes a pitchfork bifurcation, which affects both the fixed point or the limit cycle, *via* a real eigenvalue, see e.g. Strogatz *et al.* (1994). As a result, the symmetry of both the steady solution and the mean flow is broken with respect to the symmetry plane defined by $y = 0$. This is illustrated by the two mirror-conjugated steady solutions $\boldsymbol{u}_s^{\pm}$ and the two associated mean flows $\overline{\boldsymbol{u}}^{\pm}$, shown in figure 2.3 for $Re = 80$, where the base-bleeding jet appears deflected to either upward or downward with respect to the symmetry plane. Note, however, that a symmetry-preserving mean flow ($\overline{\boldsymbol{u}}$ in figure 2.3 for $Re = 80$) still exists beyond the secondary bifurcation, so that three mean flows exist beyond Re_2: two of them, $\overline{\boldsymbol{u}}^{\pm}$, are mirror-conjugated and break the symmetry, while the last one, $\overline{\boldsymbol{u}}$ preserves the symmetry. This bifurcation of the limit cycle is coincident with the bifurcation of the fixed point, as three steady solutions can be found beyond Re_2: $\boldsymbol{u}_s^{\pm}$ are mirror-conjugated and break the symmetry, while the last one, $\boldsymbol{u}_s$ preserves the symmetry. Yet, all three of them are unstable with respect to the cyclic shedding of von Kármán vortices, in which symmetry properties of the steady solution are succeeded in the resulting mean flow. When the initial condition is close to the symmetric steady solution $\boldsymbol{u}_s$, the flow regime arrives after a long transient on a limit cycle whose mean flow $\overline{\boldsymbol{u}}$ is symmetric, as illustrated in figure 2.3 for $Re = 80$. However, the dynamics of this limit cycle is only transient, indicating that it is not a stable state. After a new transient, depending on the details of the initial condition, the flow regime eventually reaches one of the two mirror-conjugated limit cycles (centered on either $\overline{\boldsymbol{u}}^{\pm}$). When the initial condition already breaks the symmetry of the flow configuration, the unstable "symmetry-centered" limit cycle is not explored and the system reaches directly one of the two stable limit cycles. The transient dynamics between these six typical states beyond Re_2 illustrate a transverse action on the original state space resulting from the new active symmetric breaking mode decoupling with the primary Hopf bifurcation, as detailed in appendix 2.D. The new active degree of freedom introduced by the pitchfork instability is responsible for the two simultaneous local pitchfork bifurcations, of both the steady solution and the periodic solution, as shown in the appendices 2.B and 2.C. The simultaneous bifurcation of the steady and periodic solution has also been observed for the cylinder wake transition from stability analyses (Noack & Eckelmann, 1994*a*,*b*) and from 3D Navier-Stokes simulations (Zhang *et al.*, 1994). A further discussion about this non-generic situation is recorded in appendix 2.E. Besides, the linear stability analysis (see appendix 2.B) and the Floquet stability analysis (see appendix 2.C) around Re_2 have been performed to prove these two simultaneous bifurcations.

As a result, when the symmetry of vortex shedding is broken, the mean value $\overline{C}_L$ (solid line) of the pressure lift coefficient $C_L = 2F_L/\rho U^2$, where F_L is the total lift force from pressure, no longer vanishes, as shown in figure 2.4. At the precision of our investigation, both the Hopf and pitchfork bifurcations were found to be supercritical.

The fluctuation amplitude of the lift coefficient is minimum for $Re \approx 80 > Re_2$, as shown in figure 2.4 (dashed curve). It starts to decrease around $Re = 30$, when the jet starts to grow

at the base of the two outer cylinders. Henceforth, the growth of the base-bleeding jet, as the Reynolds number is increased, seems to be fed with the energy of the fluctuations. Transfers of energy between the dynamically dominant degrees of freedom will be made clear in § 2.3.

When the Reynolds number is further increased up to a critical value $Re_3 \approx 104$, a new frequency rises in the power spectrum of the lift coefficient $C_L(t)$. This frequency is about one order of magnitude smaller than the natural frequency of the vortex shedding, as illustrated in figure 2.5(a) for $Re = 105$ and in the movie QP.MP4 of the additional materials (Deng *et al.*, 2020). The low and natural frequencies couple to generate combs of sharp peaks in the power spectrum, while the background level depends on the length of the time series. The new frequency is associated with modulations of the base-bleeding jet around its deflected position. A visual inspection of both the time series and the phase portrait at $Re = 105$ indicates that the new frequency also modulates the amplitude (figure 2.5(b)) of the main oscillator and thickens the limit cycle associated with the main oscillator (figure 2.5(c)). All these features are typical of a quasi-periodic dynamics, indicating that the system has most likely undergone a Neimark-Säcker bifurcation, e.g. a secondary Hopf bifurcation, at $Re = Re_3$, after which two mirror-conjugated 2-tori exist in the state space of the system.

At even larger values of the Reynolds number, $Re \geq Re_4 \approx 115$, the main peak in the power spectral density of the lift coefficient widens significantly, as shown in figure 2.6(a) for $Re = 130$. In this new regime, the instantaneous flow field is characterized by random switches between an upward or downward base-bleeding jet, see the movie CHAOS.MP4 in the additional materials (Deng *et al.*, 2020), and the mean flow $\overline{\boldsymbol{u}}$ is symmetric, as shown in figure 2.4 for $Re > 115$. The time series exhibits neither periodic nor quasi-periodic features anymore (see figure 2.6(b)) and the phase portrait exhibits a much more complex dynamics (see figure 2.6(c)). The dynamical regime henceforth exhibits many features of a chaotic regime, indicating that the system has most likely followed the Ruelle-Takens-Newhouse route to chaos (Newhouse *et al.*, 1978).

2.3 Low-dimensional modelling

We derive a mean-field Galerkin model for the primary and secondary bifurcations of the fluidic pinball. First (§ 2.3.1), the Galerkin method is recapitulated as a very general approach to reduced-order models. In § 2.3.2, the constitutive equations of the mean-field model are derived from a minimal set of assumptions. Then, mean-field Galerkin models are derived for the Hopf bifurcation (§ 2.3.3), the pitchfork bifurcation (§ 2.3.4) and the succession of both bifurcations (§ 2.3.5).

2.3.1 Galerkin method

The starting point is the non-dimensionalized incompressible Navier-Stokes equations:

$$\partial_t \boldsymbol{u} + \nabla \cdot \boldsymbol{u} \otimes \boldsymbol{u} = \nu \triangle \boldsymbol{u} - \nabla p, \tag{2.2}$$

where $\nu = 1/Re$. The velocity field satisfies the no-slip condition $\boldsymbol{u} = 0$ on the cylinders, the free-stream condition $\boldsymbol{u} = (1, 0)$ at the inflow, a no-slip condition at the top and bottom boundary and the no-stress condition at the outflow. The steady solution $\boldsymbol{u}_s$ satisfies the steady Navier-Stokes equations:

$$\nabla \cdot \boldsymbol{u}_s \otimes \boldsymbol{u}_s = \nu \triangle \boldsymbol{u}_s - \nabla p_s. \tag{2.3}$$

The Galerkin method is based on an inner product in the space of square-integrable vector fields $\mathcal{L}^2(\Omega)$ in the observation domain Ω. The standard inner product between $\boldsymbol{u}(\boldsymbol{x})$ and $\boldsymbol{v}(\boldsymbol{x})$ reads:

$$(\boldsymbol{u}, \boldsymbol{v})_\Omega := \int_\Omega d\boldsymbol{x}\, \boldsymbol{u}(\boldsymbol{x}) \cdot \boldsymbol{v}(\boldsymbol{x}). \tag{2.4}$$

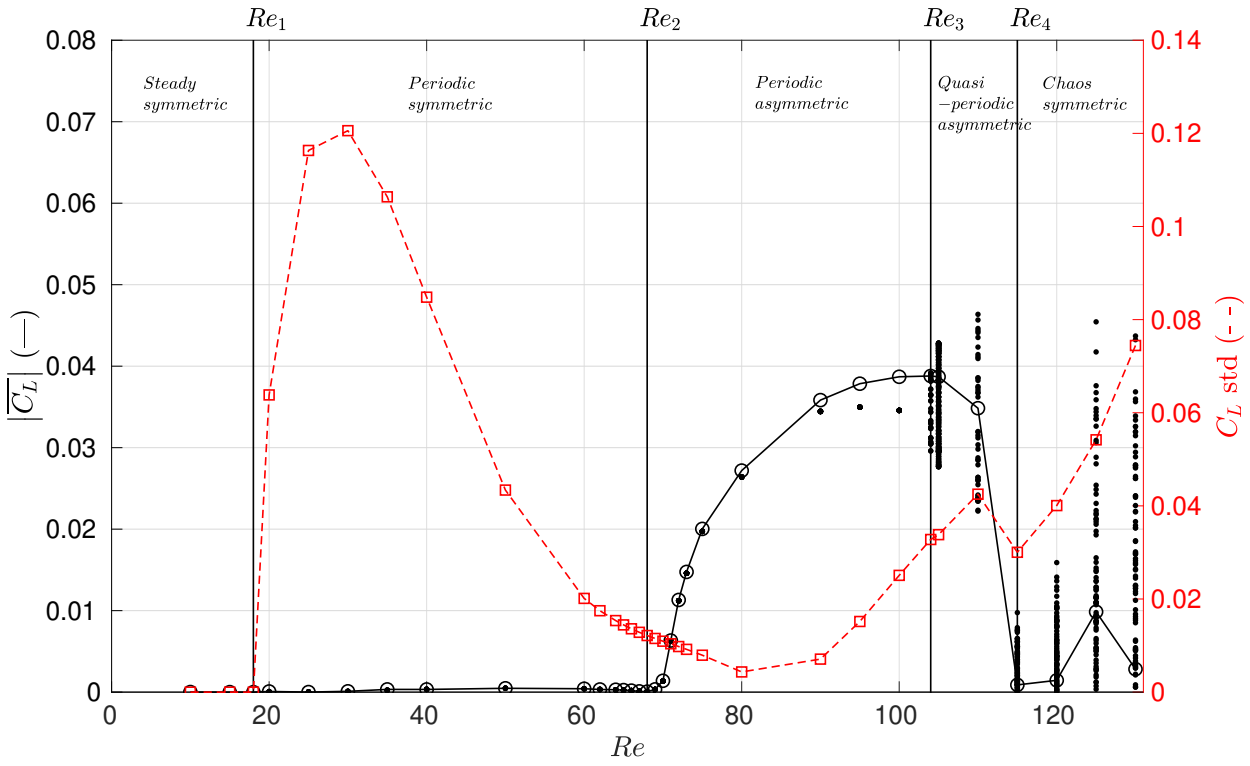


Figure 2.4: Bifurcation diagram based on the absolute value of the mean pressure lift coefficient $|\overline{C}_L|$ (circles + solid line) and its standard deviation (squares + dashed line). By symmetry, each non-vanishing mean lift value is associated with a positive and negative sign for the two attractors. The vertically distributed black dots that are visible for $Re > Re_3$, are median values $\bar{c}_n = (C_n + C_{n+1})/2$ between successive local optima $C_n = C_L(t_n)$ of $C_L(t)$, at a given Reynolds number, where the t_n are times at which $\dot{C}_L(t_n) = 0$. Transition to unsteadiness occurs at $Re_1 \approx 18$ (Hopf bifurcation), the average symmetry is broken beyond $Re_2 \approx 68$ (pitchfork bifurcation), a secondary (incommensurable) frequency rises in the power spectrum at $Re_3 \approx 104$ (Neimark-Säcker bifurcation), and transition to chaos occurs at $Re_4 \approx 115$. Note that the symmetry is statistically recovered in the chaotic regime ($\overline{C}_L \approx 0$).

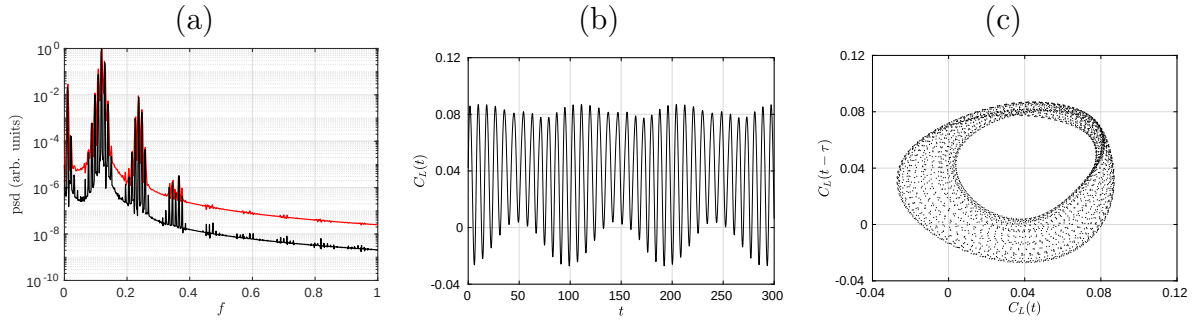


Figure 2.5: Quasi-periodic dynamics at $Re = 105$ displayed by (a) the power spectral density on time series of length $T_{\text{data}} = 400$ (red curve), $T_{\text{data}} = 900$ (black curve), (b) the time series and (c) the phase portrait of the pressure lift coefficient C_L.

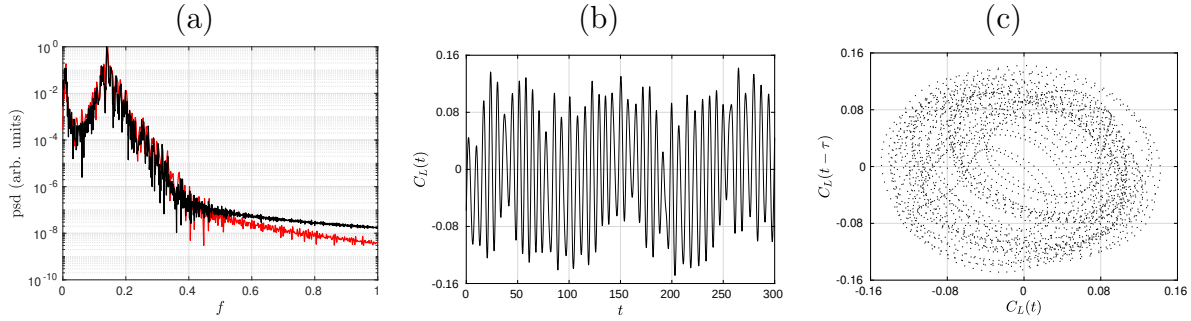


Figure 2.6: Chaotic dynamics at $Re = 130$ displayed by (a) the power spectral density on time series of length $T_{\text{data}} = 400$ (red curve), $T_{\text{data}} = 900$ (black curve), (b) the time series and (c) the phase portrait of the pressure lift coefficient C_L.

A traditional Galerkin approximation with a basic mode $\boldsymbol{u}_0$, for instance, the steady solution and N orthonormal expansion modes $\boldsymbol{u}_i(\boldsymbol{x})$, $i = 1, \ldots, N$ with time-dependent amplitudes $a_i(t)$, reads:

$$\boldsymbol{u}(\boldsymbol{x}, t) = \boldsymbol{u}_0(\boldsymbol{x}) + \sum_{i=1}^{N} a_i(t)\boldsymbol{u}_i(\boldsymbol{x}). \tag{2.5}$$

Orthonormality implies:

$$(\boldsymbol{u}_i, \boldsymbol{u}_j)_\Omega = \delta_{ij}, \quad i, j \in \{1, \ldots, N\}. \tag{2.6}$$

The projection of (2.5) on (2.2) leads to the linear-quadratic Galerkin system (Fletcher, 1984),

$$\frac{d}{dt}a_i = \nu \sum_{j=0}^{N} l_{ij}^{\nu} a_j + \sum_{j,k=0}^{N} q_{ijk}^{c} a_j a_k. \tag{2.7}$$

Following Rempfer & Fasel (1994*b*), $a_0 = 1$ is introduced. The coefficients $l_{ij}^{\nu} = (\boldsymbol{u}_i, \triangle \boldsymbol{u}_j)_\Omega$ and $q_{ijk}^{c} = (\boldsymbol{u}_i, \nabla \cdot \boldsymbol{u}_j \otimes \boldsymbol{u}_k)_\Omega$ parametrize the viscous and convective Navier-Stokes terms. The pressure term vanishes for sufficiently large domains and is neglected in the following.

In the following, the steady solution is taken as the basic mode $\boldsymbol{u}_0 = \boldsymbol{u}_s$. This implies that $\boldsymbol{a} = 0$ is a fixed point of (2.7) and the constant term $\nu l_{i0}^{\nu} + q_{i00}^{\nu} = 0$ vanishes as the projection of (2.3) onto the ith mode $\boldsymbol{u}_i$. In this case, (2.7) can be re-written as a linear-quadratic system of ordinary differential equations :

$$\frac{d}{dt}a_i = \sum_{j=1}^{N} l_{ij} a_j + \sum_{j,k=1}^{N} q_{ijk} a_j a_k, \tag{2.8}$$

where $l_{ij} = \nu l_{ij}^{\nu} + q_{ij0}^{c} + q_{i0j}^{c}$ and $q_{ijk} = q_{ijk}^{c}$ for $i, j, k \in \{1, \ldots, N\}$.

2.3.2 Mean-field modelling

Mean-field modelling allows a dramatic simplification of a general Galerkin system (2.8) close to bifurcations. In this section, we derive constitutive equations with a small number of more general assumptions.

In the spirit of the Reynolds decomposition, the velocity field is decomposed into a slowly varying distorted mean flow $\boldsymbol{u}^D$ and fluctuation $\boldsymbol{u}'$ with first-order (relaxational) or second-order (oscillatory) dynamics:

$$\boldsymbol{u}(\boldsymbol{x}, t) = \boldsymbol{u}^D(\boldsymbol{x}, t) + \boldsymbol{u}'(\boldsymbol{x}, t), \qquad \boldsymbol{u}^D(\boldsymbol{x}, t) = \boldsymbol{u}_s(\boldsymbol{x}) + \boldsymbol{u}_\Delta(\boldsymbol{x}, t) \tag{2.9}$$

Here, the *mean-field deformation* $\boldsymbol{u}_\Delta$ is the difference between the distorted mean flow and the steady solution. For the oscillatory dynamics considered, the distorted mean flow can be defined as an average over one local fluctuation period T denoted by $\langle\cdot\rangle$. Thus,

$$\boldsymbol{u}^D(\boldsymbol{x},t) = \langle \boldsymbol{u}(\boldsymbol{x},t)\rangle := \frac{1}{T}\int\limits_{t-T/2}^{t+T/2} d\tau\, \boldsymbol{u}(\boldsymbol{x},\tau). \tag{2.10}$$

After the pitchfork bifurcation into two mirror-conjugated flows $\boldsymbol{u}^+$, $\boldsymbol{u}^-$, a symmetric distorted mean flow is enforced via

$$\boldsymbol{u}^D(\boldsymbol{x},t) = \frac{1}{2}(\langle \boldsymbol{u}^+(\boldsymbol{x},t)\rangle + \langle \boldsymbol{u}^-(\boldsymbol{x},t)\rangle). \tag{2.11}$$

We note that $\boldsymbol{u}^D(\boldsymbol{x},t)$ is not the mean flow, which is defined by the post-transient limit

$$\bar{\boldsymbol{u}}(\boldsymbol{x}) = \lim_{T\to\infty}\frac{1}{T}\int\limits_0^T \boldsymbol{u}(\boldsymbol{x},\tau)d\tau. \tag{2.12}$$

The distorted mean flow coincides with mean flow for the post-transient phase before the pitchfork bifurcation. Technically, a harmonic fluctuation is assumed and this one period average is computed as an average of all phases in $[0, 2\pi]$. The somewhat loaded term "distorted mean flow" is directly adopted from the original publications of mean-field theory (Stuart, 1958). J.T. Stuart considers this flow as "distorted" from the steady solution by the Reynolds stress associated with the instability mode(s).

For a nominally symmetric cylindrical obstacle, the distorted mean flow can be expected to be symmetric while the dominant fluctuation is antisymmetric. This leads to a symmetry-based decomposition of the flow into a symmetric contribution $\boldsymbol{u}^s = (u^s, v^s) \in \mathcal{U}^s$ with

$$u^s(x,-y) = u^s(x,y), \quad v^s(x,-y) = -v^s(x,y) \tag{2.13}$$

and an antisymmetric component $\boldsymbol{u}^a = (u^a, v^a) \in \mathcal{U}^a$ satisfying

$$u^a(x,-y) = -u^a(x,y), \quad v^a(x,-y) = v^a(x,y). \tag{2.14}$$

Here, $\mathcal{U}^s$ and $\mathcal{U}^a$ denote the set of symmetric and antisymmetric vector fields, respectively. The resulting decomposition reads

$$\boldsymbol{u}(\boldsymbol{x},t) = \boldsymbol{u}^s(\boldsymbol{x},t) + \boldsymbol{u}^a(\boldsymbol{x},t). \tag{2.15}$$

In the sequel, we will identify the distorted mean flow with the symmetric component and the fluctuation with the antisymmetric one:

$$\mathbf{u}^D(\mathbf{x},t) = \mathbf{u}^s(\mathbf{x},t), \quad \boldsymbol{u}'(\boldsymbol{x},t) = \boldsymbol{u}^a(\boldsymbol{x},t). \tag{2.16}$$

This identification is justified for symmetry-breaking bifurcations with first- or second-order dynamics with neglected higher harmonics. For brevity, $\mathcal{U}^s$ and $\mathcal{U}^a$ are introduced as symmetric and antisymmetric subsets of $\mathcal{L}^2(\Omega)$.

The convective term is easily shown to have the following symmetry properties:

$$\nabla\cdot\boldsymbol{u}^s\otimes\boldsymbol{u}^s \in \mathcal{U}^s, \tag{2.17a}$$
$$\nabla\cdot\boldsymbol{u}^a\otimes\boldsymbol{u}^a \in \mathcal{U}^s, \tag{2.17b}$$
$$\nabla\cdot\boldsymbol{u}^s\otimes\boldsymbol{u}^a \in \mathcal{U}^a, \tag{2.17c}$$
$$\nabla\cdot\boldsymbol{u}^a\otimes\boldsymbol{u}^s \in \mathcal{U}^a. \tag{2.17d}$$

The antisymmetric component is derived starting with (2.2), subtracting the steady version of (2.3) and exploiting the symmetry of $\boldsymbol{u}^D$, the antisymmetry of $\boldsymbol{u}'$ as well as the symmetry relations (2.17). The fluctuation dynamics reads:

$$\partial_t \boldsymbol{u}' + \nabla \cdot \left[\boldsymbol{u}^D \otimes \boldsymbol{u}' + \boldsymbol{u}' \otimes \boldsymbol{u}^D\right] = \nu \triangle \boldsymbol{u}' - \nabla p'. \tag{2.18}$$

Analogously, the symmetric part describes the distorted mean flow dynamics:

$$\partial_t \boldsymbol{u}_\Delta + \nabla \cdot [\boldsymbol{u}_s \otimes \boldsymbol{u}_\Delta + \boldsymbol{u}_\Delta \otimes \boldsymbol{u}_s + \boldsymbol{u}_\Delta \otimes \boldsymbol{u}_\Delta + \boldsymbol{u}' \otimes \boldsymbol{u}'] = \nu \triangle \boldsymbol{u}_\Delta - \nabla p_\Delta \tag{2.19}$$

Note that this symmetric component of the Navier-Stokes equations has not yet been averaged and the sum of Eqs. (2.3), (2.19) and (2.18) leads to the Navier-Stokes equations (2.2). To this point, all equations are strict identities for the symmetric and antisymmetric part of the Navier-Stokes dynamics.

Next, we follow mean-field arguments and consider $\boldsymbol{u}'$ and $\boldsymbol{u}_\Delta$ as small perturbations around the fixed point $\boldsymbol{u}_s$. Let $\boldsymbol{u}' \in O(\varepsilon)$ and $\boldsymbol{u}_\Delta \in O(\delta)$ where ε and δ are smallness parameters. Hence, $\boldsymbol{u}_\Delta \otimes \boldsymbol{u}_\Delta \in O(\delta^2)$ can be neglected in comparison to the $O(\delta)$ terms $\boldsymbol{u}_s \otimes \boldsymbol{u}_\Delta$, $\boldsymbol{u}_\Delta \otimes \boldsymbol{u}_s$. We follow Stuart's original idea to separate between the fluctuation $\boldsymbol{u}'$ driven by the instability and the resulting mean-field deformation $\boldsymbol{u}_\Delta$ and arrive at the unsteady linearized Reynolds equation,

$$\partial_t \boldsymbol{u}_\Delta + \nabla \cdot [\boldsymbol{u}_s \otimes \boldsymbol{u}_\Delta + \boldsymbol{u}_\Delta \otimes \boldsymbol{u}_s + \langle \boldsymbol{u}' \otimes \boldsymbol{u}' \rangle] = \nu \triangle \boldsymbol{u}_\Delta - \nabla p_\Delta. \tag{2.20}$$

The mean-field deformation $\boldsymbol{u}_\Delta$ characterized by the scale δ is seen to respond linearly to the Reynolds stress force $-\nabla \cdot \langle \boldsymbol{u}' \otimes \boldsymbol{u}' \rangle$ scaling with ϵ^2. Hence, $\delta \sim \epsilon^2$.

Summarizing, Eqs. (2.18) and (2.20) are the constitutive equations of mean-field theory exploiting only symmetry and smallness of the mean-field deformation.

Close to the critical Reynolds number Re_c, the temporal growth rate can be Taylor expanded to $\sigma = \alpha(Re - Re_c)$ and can be assumed to be small. In this case, $\partial_t \boldsymbol{u}_\Delta \in O(\sigma\delta)$, i.e. the time derivative of (2.20) can be neglected with respect to the other terms $\in O(\delta) = O(\varepsilon^2)$. This leads to the steady linearized Reynolds equation:

$$\nabla \cdot [\boldsymbol{u}_s \otimes \boldsymbol{u}_\Delta + \boldsymbol{u}_\Delta \otimes \boldsymbol{u}_s + \langle \boldsymbol{u}' \otimes \boldsymbol{u}' \rangle] = \nu \triangle \boldsymbol{u}_\Delta - \nabla p_\Delta. \tag{2.21}$$

This equation is also true for the post-transient solution, e.g. the limit cycle of a Hopf bifurcation or the asymmetric state of a pitchfork bifurcation. Often, the distorted mean flow $\boldsymbol{u}^D$ quickly responds to the Reynolds stress even far away from the bifurcation.

2.3.3 Supercritical Hopf bifurcation

At low Reynolds numbers, a symmetric stable steady solution $\boldsymbol{u}_s \in \mathcal{U}^s$ is observed. Periodic vortex shedding sets in with the occurrence of an unstable oscillatory antisymmetric eigenmode at $Re \geq Re_1$. The Reynolds-number dependent initial growth rate and frequency are denoted by σ_1 and ω_1, respectively. The real and imaginary parts of this eigenmode are $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, respectively, both antisymmetric modes. In the following, these modes are assumed to be orthonormalized.

This oscillation generates a Reynolds stress, which changes the mean flow via (2.20). The mean flow deformation is described by the symmetric shift mode $\boldsymbol{u}_3$ with unit norm. By symmetry, the first two modes are orthogonal with respect to the shift mode. Thus, the modes form an orthonormal basis. The resulting Galerkin expansion reads:

$$\boldsymbol{u}(\boldsymbol{x}, t) = \boldsymbol{u}_s(\boldsymbol{x}) + \underbrace{a_1(t)\boldsymbol{u}_1(\boldsymbol{x}) + a_2(t)\boldsymbol{u}_2(\boldsymbol{x})}_{\boldsymbol{u}'} + \underbrace{a_3(t)\boldsymbol{u}_3(\boldsymbol{x})}_{\boldsymbol{u}_\Delta}. \tag{2.22}$$

Moreover, polar coordinates are introduced $a_1(t) = r(t)\cos\theta(t)$, $a_2(t) = r(t)\sin\theta(t)$, $d\theta/dt = \omega(t)$ where r and ω are assumed to be slowly varying functions of time.

Substituting (2.22) in (2.18), projecting on $\boldsymbol{u}_i$, $i = 1, 2$ and applying the Krylov-Bogoliubov (Jordan & Smith, 1999) averaging method yields:

$$da_1/dt = \sigma a_1 - \omega a_2, \qquad \sigma = \sigma_1 - \beta a_3, \tag{2.23a}$$

$$da_2/dt = \sigma a_2 + \omega a_1, \qquad \omega = \omega_1 + \gamma a_3. \tag{2.23b}$$

Here, $\sigma_1, \omega_1, \beta > 0$ for a supercritical Hopf bifurcation. We refer to Noack *et al.* (2003) for details.

Krylov-Bogoliubov averaging implies a harmonic balancing on the slowly varying amplitude and frequency of oscillatory $a_{1,2}$ and the slowly varying a_3 dynamics. The corresponding original theorem includes a convergence proof of this approximation for a second-order ordinary differential equation for oscillations in the limit of small nonlinearity. We cannot perform this limit but justify the operation on the a priori observation that quadratic Galerkin system terms q_{ijk} are typically two orders of magnitude smaller than the dominant linear coefficients, i.e. describe a small nonlinearity. A posteriori the operation is justified by the results, i.e. by obtaining amplitudes and frequencies with up to a few percent error.

Substituting (2.22) in (2.21) replaces (2.25) by the mean-field manifold:

$$a_3 = \kappa \left(a_1^2 + a_2^2\right) \tag{2.24}$$

with derivable proportionality constant κ.

Alternatively, the mean-field manifold may be obtained from the Galerkin system. Substituting (2.22) in (2.20) and projecting on $\boldsymbol{u}_3$ yields:

$$da_3/dt = \sigma_3 a_3 + \beta_3 \left(a_1^2 + a_2^2\right), \tag{2.25}$$

where $\sigma_3 < 0$ and $\beta_3 > 0$ are necessary for a globally stable limit cycle. Note that (2.25) can be rewritten as:

$$da_3/dt = \sigma_3 \left[a_3 - \kappa \left(a_1^2 + a_2^2\right)\right]. \tag{2.26}$$

Now, the slaving process which leads to the mean-field manifold of (2.24) can be appreciated from the mean-field Galerkin system. If $|\sigma_3| \gg \sigma_1$, the timescale of slaving a_3 to the fluctuation level $a_1^2 + a_2^2$ is much smaller than the timescale of the transient and da_3/dt can be set to zero.

Eqs. (2.23) and (2.24) yield the famous Landau equations with the cubic damping term:

$$dr/dt = \sigma_1 r - \beta\kappa r^3, \quad d\theta/dt = \omega_1 + \gamma\kappa r^2. \tag{2.27}$$

The Landau oscillator leads to a stable limit cycle with $r^\circ = \sqrt{\sigma_1/\beta\kappa}$, frequency $\omega^\circ = \omega_1 + \sigma_1\gamma/\beta$ and shift-mode amplitude $a_3^\circ = \sigma_1/\beta$. The three nonlinearity parameters β, γ and κ can be uniquely derived from the limit cycle parameters r°, ω°, and a_3°. The growth rate σ_3 needs to be chosen sufficiently large, e.g. $\sigma_3 = -10\sigma_1$ to ensure slaving on the manifold.

Eqs. (2.23), (2.25) are the mean-field Galerkin system, while Eqs. (2.23), (2.24) characterize the original mean-field model, *i.e.* the slaved Galerkin system. Near the Hopf bifurcation, when $\sigma_1(Re)$ crosses the zero line at Re_1, the growth rate is approximated by $\sigma_1 = \alpha\,(Re - Re_1)$ implying the square-root law $r^\circ = \sqrt{\alpha/\beta\kappa}\sqrt{Re - Re_1}$.

The Landau equation has been proposed by Landau (see, *e.g.* Landau & Lifshitz, 1987), derived from the Navier-Stokes equation by Stuart (1958), generalized for Galerkin systems by Noack *et al.* (2003), and validated in numerous simulations and experiments for cylinder wakes (Schumm *et al.*, 1994) and other soft onsets of oscillatory flows. We note that the proposed derivation from symmetry considerations constrains the model to symmetric obstacles but liberates the mean-field Galerkin model from typical assumptions, likes closeness to the Hopf bifurcation or the need for frequency filtered Navier-Stokes equations.

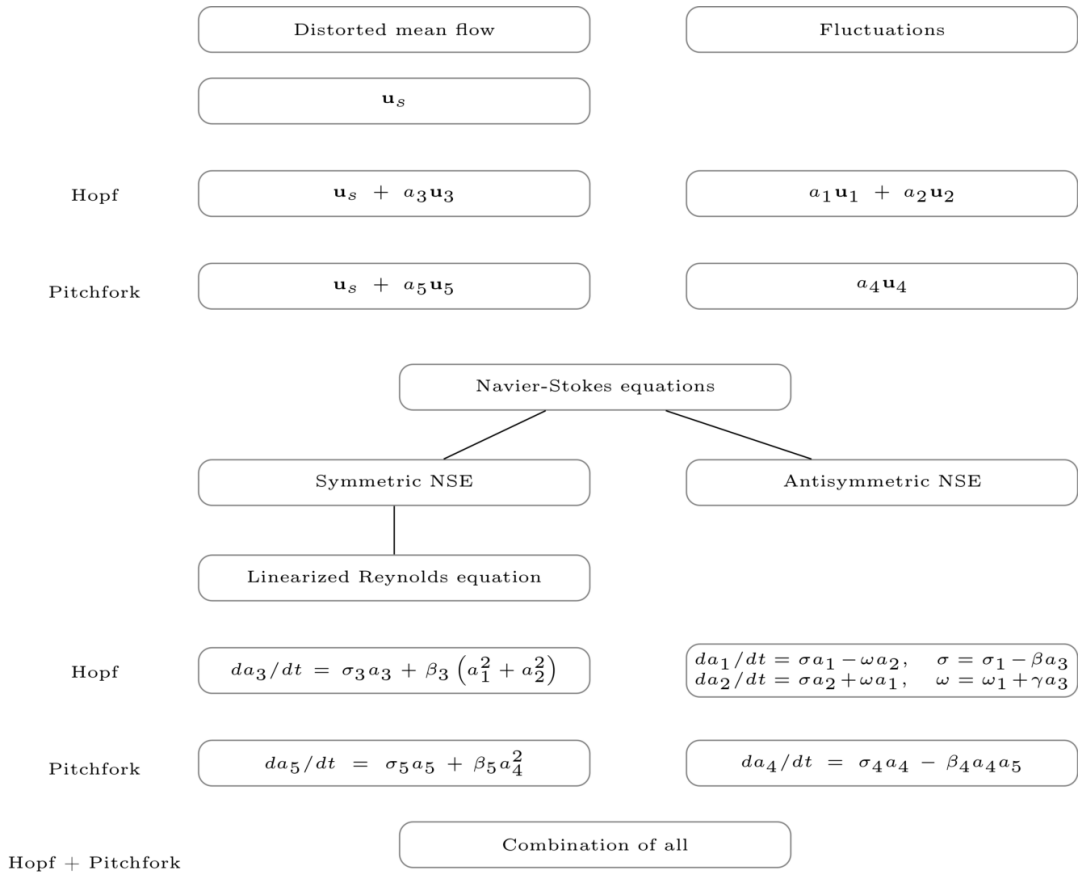


Table 2.1: Symmetries and hierarchy of equations.

2.3.4 Supercritical pitchfork bifurcation

Next, the symmetry-breaking pitchfork bifurcation of a steady symmetric Navier-Stokes equation is considered. Now, mode $\boldsymbol{u}_4$ describes the antisymmetric instability with positive growth rate σ_4. The shift mode $\boldsymbol{u}_5$ prevents unbounded exponential growth. The corresponding Galerkin expansion reads

$$\boldsymbol{u}(\boldsymbol{x},t) = \boldsymbol{u}_s(\boldsymbol{x}) + \underbrace{a_4(t)\boldsymbol{u}_4(\boldsymbol{x})}_{\boldsymbol{u}'} + \underbrace{a_5(t)\boldsymbol{u}_5(\boldsymbol{x})}_{\boldsymbol{u}_\Delta}. \tag{2.28}$$

Substituting (2.28) in Eqs. (2.18) and (2.20), and exploiting the symmetry of the modes, yields:

$$da_4/dt = \sigma_4 a_4 - \beta_4 a_4 a_5, \tag{2.29a}$$
$$da_5/dt = \sigma_5 a_5 + \beta_5 a_4^2. \tag{2.29b}$$

Note that a linear a_5 term and quadratic a_4a_4 and a_5a_5 terms in (2.29a) are ruled out by symmetry. Similarly, a linear a_4 term or a mixed quadratic term a_4a_5 in (2.29b) are prohibited by symmetry. The quadratic a_5a_5 term is not consistent with the linearized Reynolds equation (2.20). The pitchfork bifurcation can be considered as a Hopf bifurcation with $\omega = 0$ and a single mode. Replacing a_1 by a_4, a_3 by a_5 and setting $a_2 = 0$ yields (2.29a) from (2.23) and (2.29b) from (2.25)—modulo names of the coefficients.

Eqs. (2.29a), (2.29b) are the mean-field Galerkin system. Substituting (2.28) in (2.21) yields the manifold:

$$a_5 = \kappa_5 a_4^2, \tag{2.30}$$

with $\kappa_5 = -\beta_5/\sigma_5$. The asymmetric steady solutions read $a_4^{\pm} = \pm\sqrt{\sigma_4/\beta_4\kappa_5}$, $a_5 = \sigma_4/\beta_4$. The two nonlinearity parameters κ_4, β_4 are readily determined from the two asymptotic values a_4 and a_5.

The growth rate can be set in analogy to the previous model to $\sigma_5 = -10\sigma_4$ to ensure slaving on the manifold.

From (2.29a) and (2.30), the famous unstable dynamics with cubic damping term is obtained:

$$da_4/dt = \sigma_4 a_4 - \kappa_5 \beta_4 a_4^3,$$

where $\kappa_5\beta_4 > 0$ for a supercritical bifurcation. Eqs. (2.29a), (2.29b) are the mean-field Galerkin system.

Near the secondary pitchfork bifurcation, $\sigma_4 = \alpha_2 (Re - Re_2)$ and $a_4 \propto \sqrt{Re - Re_2}$. The parameters of the pitchfork Galerkin system can be derived from the eigenmode and the asymptotic state in complete analogy to § 2.3.3. The growth rate $\sigma_5 = -10\sigma_4$ will ensure the slaving of (2.30). We emphasize that this pitchfork model is derived primarily from symmetry considerations and does not require closeness to the critical parameter.

2.3.5 Pitchfork bifurcation of periodic solution

In the final modelling effort, a low-dimensional model from a primary supercritical Hopf bifurcation at $Re = Re_1$ and a secondary supercritical pitchfork bifurcation at $Re = Re_2 > Re_1$ is derived following the numerical observations of the fluidic pinball in § 2.2. For simplicity, closeness to the secondary bifurcation is assumed. For the same reason, the mean-field Galerkin system shall still describe the periodic solution. In this case, the generalized 5-mode mean-field expansion:

$$\boldsymbol{u}(\boldsymbol{x}, t) = \boldsymbol{u}_s(\boldsymbol{x}) + \sum_{i=1}^{5} a_i(t)\boldsymbol{u}_i(\boldsymbol{x}) \tag{2.31}$$

describes the flow where $a_1, a_2, a_3 \in O(1)$ and $a_4 \in O(\varepsilon)$ and $a_5 \in O(\delta)$, ε, δ being smallness parameters associated with the pitchfork bifurcation. We project Eqs. (2.18) and (2.20) on (2.31). The $O(1)$ terms encapsulate the original Hopf model while the low-pass filtered $O(\varepsilon, \delta)$ terms yield the original pitchfork system. This yields the following generalized mean-field system:

$$\begin{aligned}
da_1/dt &= \sigma a_1 - \omega a_2, \qquad \sigma = \sigma_1 - \beta a_3 && \text{(2.32a)}\\
da_2/dt &= \sigma a_2 + \omega a_1, \qquad \omega = \omega_1 + \gamma a_3 && \text{(2.32b)}\\
da_3/dt &= \sigma_3 a_3 + \beta_3 \left(a_1^2 + a_2^2\right) && \text{(2.32c)}\\
da_4/dt &= \sigma_4 a_4 - \beta_4 a_4 a_5 && \text{(2.32d)}\\
da_5/dt &= \sigma_5 a_5 + \beta_5 a_4^2 && \text{(2.32e)}
\end{aligned}$$

The linear instability parameters σ_1, ω_1, σ_4 are obtained from the corresponding global stability analysis. Slaving is ensured with $\sigma_3 = -10\sigma_1$ and $\sigma_5 = -10\sigma_4$. The nonlinearity parameters β, γ, β_3, β_4 and β_5 are determined from the limit cycle parameters r°, ω° and a_3° and pitchfork parameters $a_4^\pm$ and $a_5^\pm$ in the asymptotic regime.

As the amplitude of the pitchfork bifurcation grows, the smallness argument does not hold and we get cross-terms, like $\sigma = \sigma_1 - \beta a_3 - \beta_{15} a_5$. We shall not pause to elaborate on the possible generalizations now, but will return to the topic in the result section.

2.3.6 Sparse Galerkin model from mean-field considerations

The mean-field Galerkin system (2.32) with decoupled Hopf and pitchfork dynamics can, by construction, only be expected to hold near the pitchfork bifurcation $Re \approx Re_2$. At higher Reynolds numbers $Re > Re_2$, cross-terms will appear, e.g., the growth rate σ may also depend on the pitchfork-related shift mode amplitude a_5. The most general Galerkin system (2.8) contains $5 \times 5 = 25$ linear terms and $5 \times 5 \times 6/2 = 75$ quadratic terms.

The assumed symmetry of the modes excludes roughly half of these 100 coefficients. Let $\chi_i = 0$ for symmetric mean flow modes $\boldsymbol{u}_i$, $i = 3, 5$ and $\chi_i = 1$ for the antisymmetric fluctuation modes $\boldsymbol{u}_i$, $i = 1, 2, 4$. The linear coefficients

$$l_{ij} = -\nu \left(\boldsymbol{u}_i, \triangle \boldsymbol{u}_i\right)_\Omega + \left(\boldsymbol{u}_i, \nabla \cdot \boldsymbol{u}_s \otimes \boldsymbol{u}_j\right)_\Omega + \left(\boldsymbol{u}_i, \nabla \cdot \boldsymbol{u}_j \otimes \boldsymbol{u}_s\right)_\Omega$$

can be shown to vanish if $\quad \bmod(\chi_i + \chi_j, 2) = 1$. In other words, the coefficients l_{ij} vanish if the modes $\boldsymbol{u}_i$ and $\boldsymbol{u}_j$ have opposite symmetries. This excludes 12 of the 25 linear coefficients. Analogously, the quadratic coefficient q_{ijk} can be shown to vanish if $\chi_i \neq \quad \bmod(\chi_j + \chi_k, 2)$. In other words, the q_{ijk} vanishes if the symmetry of $\boldsymbol{u}_i$ does not coincide with the symmetry of quadratic term $\boldsymbol{u}_j \otimes \boldsymbol{u}_k$. The quadratic term is symmetric if the modes $\boldsymbol{u}_j$ and $\boldsymbol{u}_k$ are both symmetric or both antisymmetric and is antisymmetric if both modes have opposite symmetries. In summary, q_{ijk} vanishes if one or three modes are antisymmetric.

An additional sparsity of the coefficients arises from the temporal dynamics. Modes $\boldsymbol{u}_i$, $i = 1, 2$ have oscillatory behaviour with angular frequency ω, while the other modes show first-order dynamics, i.e., relaxation to asymptotic values. We apply the Krylov-Bogoliubov approximation with oscillatory a_1, a_2 and slow a_3, a_4, a_5 dynamics. Thus, for instance, $l_{41}a_1$ vanishes on a one-period average and should not contribute to da_4/dt. The linear coefficient l_{41} can hence be set to zero. Taking the quadratic terms for example, a_1a_3 generates a first harmonic. Hence, $q_{413}a_1a_3$ cannot contribute to da_4/dt but $q_{113}a_1a_3$ can contribute to the oscillatory behaviour of da_1/dt. Similarly, a_1a_2 generates a second harmonic, $a_1a_2 = r\cos\omega t \times r\sin\omega t = (1/2)\, r^2 \sin 2\omega t$ does not have a steady contribution, so q_{312} can be set to zero.

From symmetry and Krylov-Bogoliubov considerations, only 9 coefficients contribute to the linear term: l_{11}, l_{12}, l_{21}, l_{22}, l_{33}, l_{35}, l_{44}, l_{53}, l_{55}. Note that the oscillator equations contain the 2×2 block, while the shift-mode equations $i = 3, 5$ have cross-terms and the pitchfork amplitude dynamics $i = 4$ has no cross-terms. Similarly, only 16 quadratic coefficients survive. The first 8 coefficients q_{113}, q_{115}, q_{123}, q_{125}, q_{213}, q_{215}, q_{223}, q_{225} are consistent with the Landau oscillator but with cross-terms to the pitchfork-related shift-mode amplitude, i.e. $\sigma = \sigma_1 - \beta a_3 - \beta_{15}a_5$ and $\omega = \omega_1 + \gamma a_3 + \gamma_{15}a_5$, introducing β_{15} and γ_{15} as new coefficients. The first and second shift-mode equations $i = 3, 5$ may contain six quadratic terms q_{311}, q_{322}, q_{344}, q_{511}, q_{522}, q_{544} from the Reynolds stresses. The amplification of the pitchfork dynamics is affected by the shift modes via q_{443}, q_{445}.

2.4 Primary flow regime

The primary flow regime covers the range of Reynolds numbers $Re_1 < Re < Re_2$. We consider the flow and reduction of the dynamics at $Re = 30$, as a representative case. In § 2.4.1, a linear stability analysis is done on the steady solution and the three degrees of freedom of the flow dynamics are identified. In § 2.4.2, we propose a least-order model of the flow dynamics at $Re = 30$ and compare its performance with respect to the full flow dynamics.

2.4.1 Eigenspectra of the steady solution

The steady solution becomes unstable beyond $Re = Re_1$, as reported in § 2.2. A linear stability analysis indicates that one pair of complex-conjugated eigenmodes have a positive growth rate on the range $Re_1 < Re < Re_2$, as shown in figure 2.7 for $Re = 30$. These two leading eigenmodes are associated with vortical structures shed downstream in the wake, at the angular frequency $1/2$. As the instability grows, the distorted mean flow $\boldsymbol{u}^D = \boldsymbol{u}_s + \boldsymbol{u}_\Delta$ changes, as expected by Eq. (2.3) and (2.19). The shift mode $\boldsymbol{u}_3$, involved in $\boldsymbol{u}_\Delta$ at $Re = 30$, is shown in figure 2.8(c). In the permanent (time-periodic) flow regime, the distorted mean flow $\boldsymbol{u}^D$ eventually matches the asymptotic mean flow field $\overline{\boldsymbol{u}}$, and vortex shedding is well established with frequency 8.7×10^{-2}.

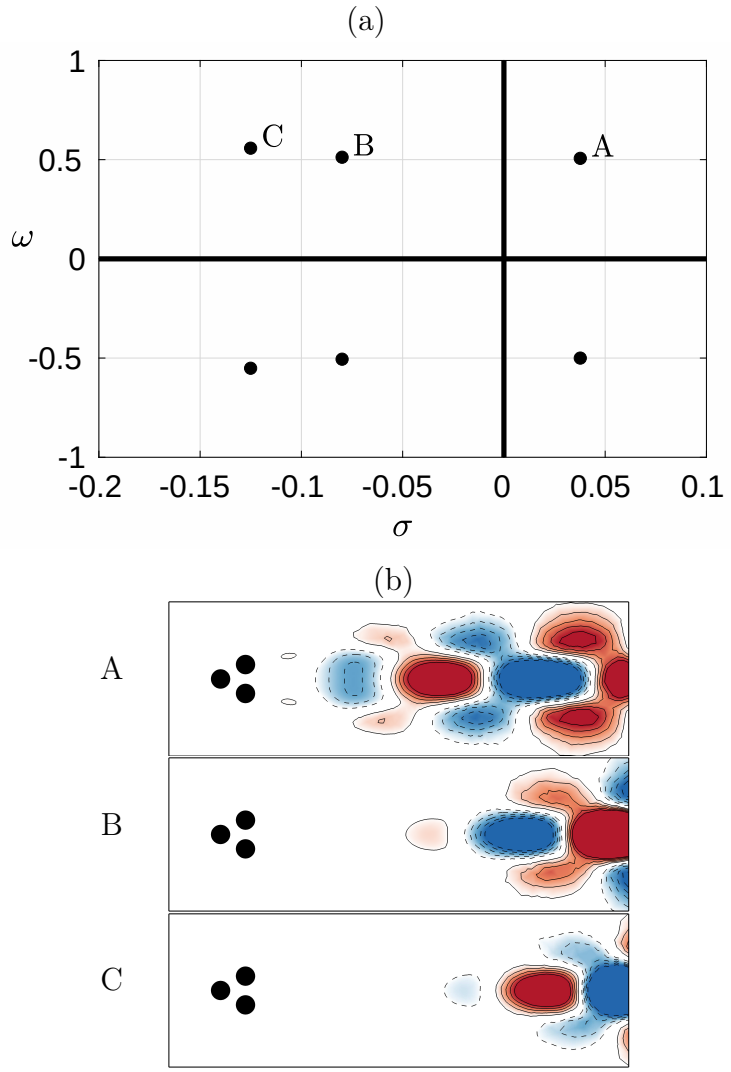


Figure 2.7: (a) Eigenspectrum resulting from the linear stability analysis of the steady solution $\boldsymbol{u}_s$, together with (b) the first three leading eigenmodes, at $Re = 30$. Only the real part of the complex eigenmodes is shown. Red color and solid contours are positive values of the vorticity, blue color and dashed contours are negative values.

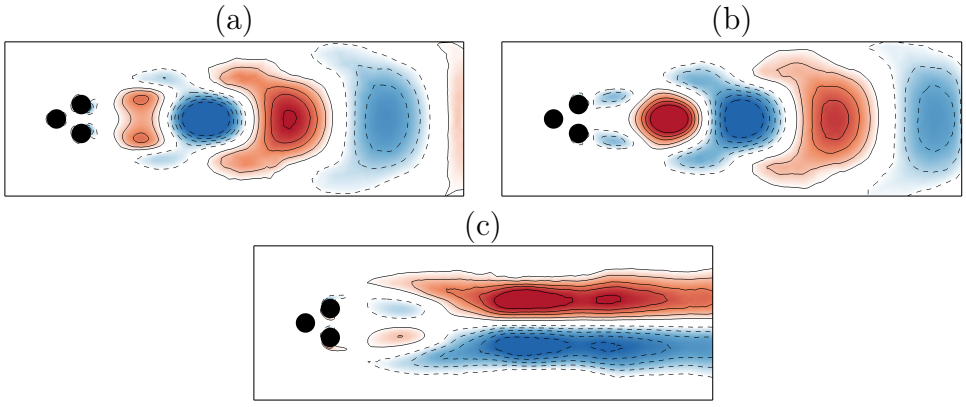


Figure 2.8: First two leading POD modes $\boldsymbol{u}_{1,2}$ at $Re = 30$ (a) & (b) and shift mode $\boldsymbol{u}_3$ (c). Red color and solid contours are positive values of the vorticity, blue color and dashed contours are negative values.

σ_1	ω_1	σ_3	β	γ	κ
3.80×10^{-2}	5.00×10^{-1}	$-10\sigma_1$	1.40×10^{-2}	1.70×10^{-2}	2.10×10^{-1}

Table 2.2: Coefficients of the reduced-order model (ROM) at $Re = 30$. See text for details.

2.4.2 Reduced-order model (ROM) of the primary flow regime

As introduced in § 2.3, the Galerkin ansatz for the Hopf bifurcation reads:

$$\boldsymbol{u}(\boldsymbol{x},t) \approx \boldsymbol{u}_s(\boldsymbol{x}) + a_1(t)\boldsymbol{u}_1(\boldsymbol{x}) + a_2(t)\boldsymbol{u}_2(\boldsymbol{x}) + a_3(t)\boldsymbol{u}_3(\boldsymbol{x}). \tag{2.33}$$

The von Kármán modes $\boldsymbol{u}_{1,2}$ could be chosen as the real and imaginary part of the first eigenmode, respectively. This choice would make sense to describe the transient dynamics close to the steady solution. A better choice for describing the dynamics on the asymptotic limit cycle is to choose $\boldsymbol{u}_{1,2}$ as the first two modes of a Proper Orthogonal Decomposition (POD) of the limit cycle data. The first two POD modes actually contribute to almost 95% of the total fluctuating kinetic energy at $Re = 30$ and are clearly associated with the von Kármán street of shed vortices, as shown in figure 2.8 (a)&(b). For the construction of the ROM, POD modes $\boldsymbol{u}_{1,2}$ are preferred to the two leading eigenmodes, because we demand an accurate representation of the asymptotic periodic dynamics. Following Eq. (2.23)-(2.24), the dynamical system resulting from the Galerkin projection of ansatz (2.33) on the Navier-Stokes equations, after Krylov-Bogoliubov simplifications, reads

$$da_1/dt = \sigma a_1 - \omega a_2 \tag{2.34}$$
$$da_2/dt = \sigma a_2 + \omega a_1 \tag{2.35}$$
$$da_3/dt = \sigma_3 \left(a_3 - \kappa(a_1^2 + a_2^2)\right), \tag{2.36}$$

with $\sigma = \sigma_1 - \beta a_3$ and $\omega = \omega_1 + \gamma a_3$. The value of the coefficients at $Re = 30$ for the resulting ROM are summarized in table 2.2. Note that all coefficients but γ and σ_3 are fixed by either the linear stability analysis or the asymptotic dynamics, see § 2.3. The coefficient σ_3 can be chosen arbitrarily large as a_3 is slaved to a_1, a_2 (here we chose $\sigma_3 = -10\sigma_1$), while γ had to be calibrated in order to better match the asymptotic angular frequency.

The dynamics of both the fluidic pinball (solid blue curve) and the ROM (dashed red curve) are compared in the three-dimensional subspace spanned by a_1, a_2, a_3, see the top of figure 2.9. In figure 2.9 are also shown the individual time series of a_1 to a_3 for both the fluidic pinball and the ROM (same representation). As expected from the POD modes $\boldsymbol{u}_{1,2}$, the dynamics on the asymptotic (permanent) limit cycle is well described in amplitude r and angular frequency ω by the ROM. Moreover, the ROM also captures the transient dynamics on the parabolic manifold $a_3 \equiv \kappa(a_1^2 + a_2^2)$. Henceforth, although all coefficients but one are fixed, the Galerkin system (2.33) is able to reproduce the most salient dynamical features of the flow in both the transient and the permanent regimes.

2.5 Secondary flow regime

The secondary flow regime ranges over $Re_2 < Re < Re_3$. For illustration, we focus on the flow at $Re = 80$, *i.e.* at a finite distance from the secondary bifurcation. In § 2.5.1 a linear stability analysis of the resulting three steady solutions is performed. A least-order model is proposed and discussed in § 2.5.2.

2.5.1 Eigenspectra of the steady solutions

As a result of the pitchfork bifurcation, there exist three steady solutions beyond Re_2: the symmetric steady solution $\boldsymbol{u}_s$, unstable to the periodic vortex shedding beyond Re_1, and two mirror-

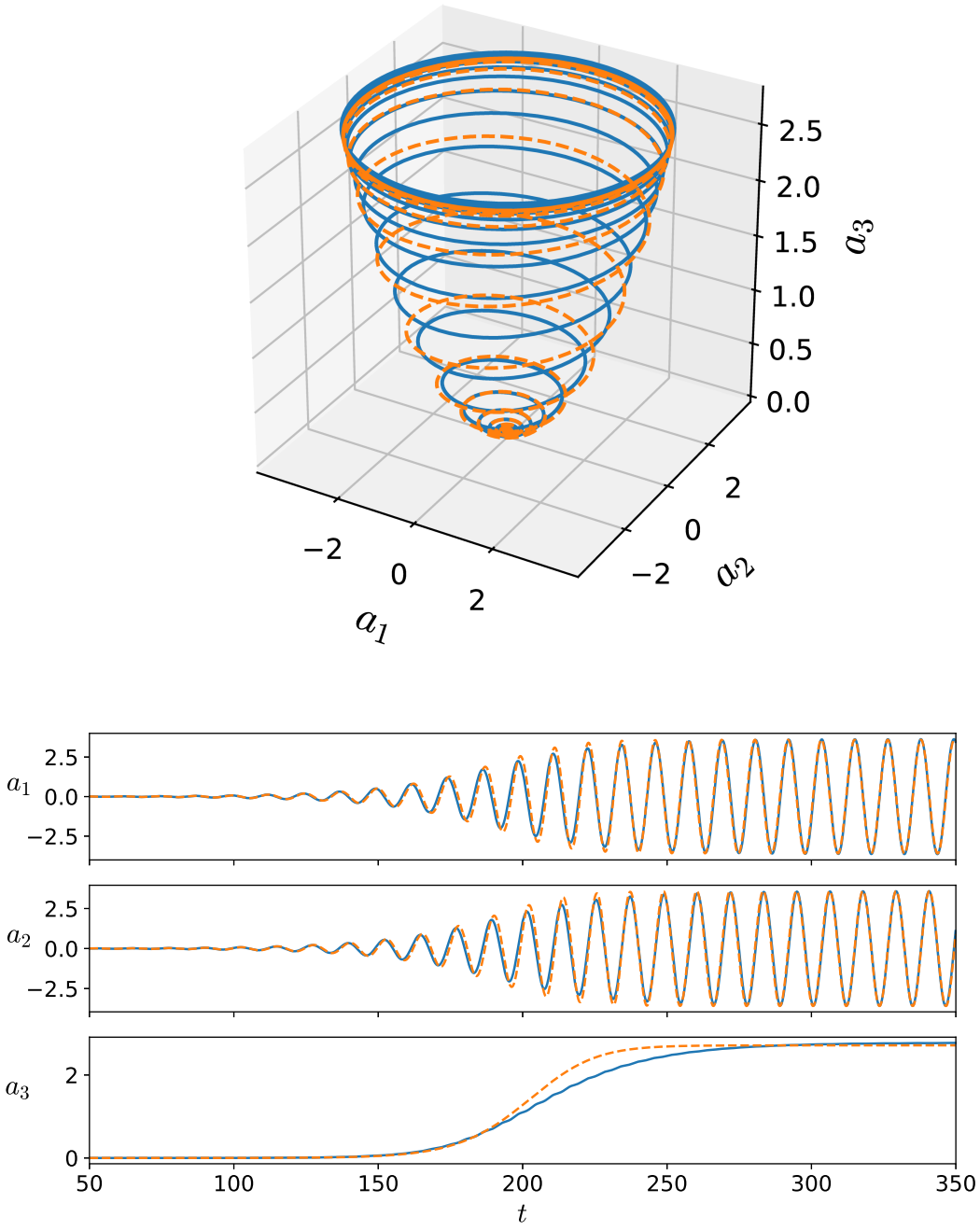


Figure 2.9: Top figure: three-dimensional state space spanned by a_1, a_2, a_3, at $Re = 30$. From direct numerical simulations of the fluidic pinball (solid blue line) and from the mean-field reduced-order model (dashed red line). The initial condition, identical in both systems, starts close to the fixed point (steady solution) before evolving on the parabolic manifold $a_3 = \kappa(a_1^2 + a_2^2)$ toward the asymptotic limit cycle. Bottom figure: corresponding time series for a_1 to a_3.

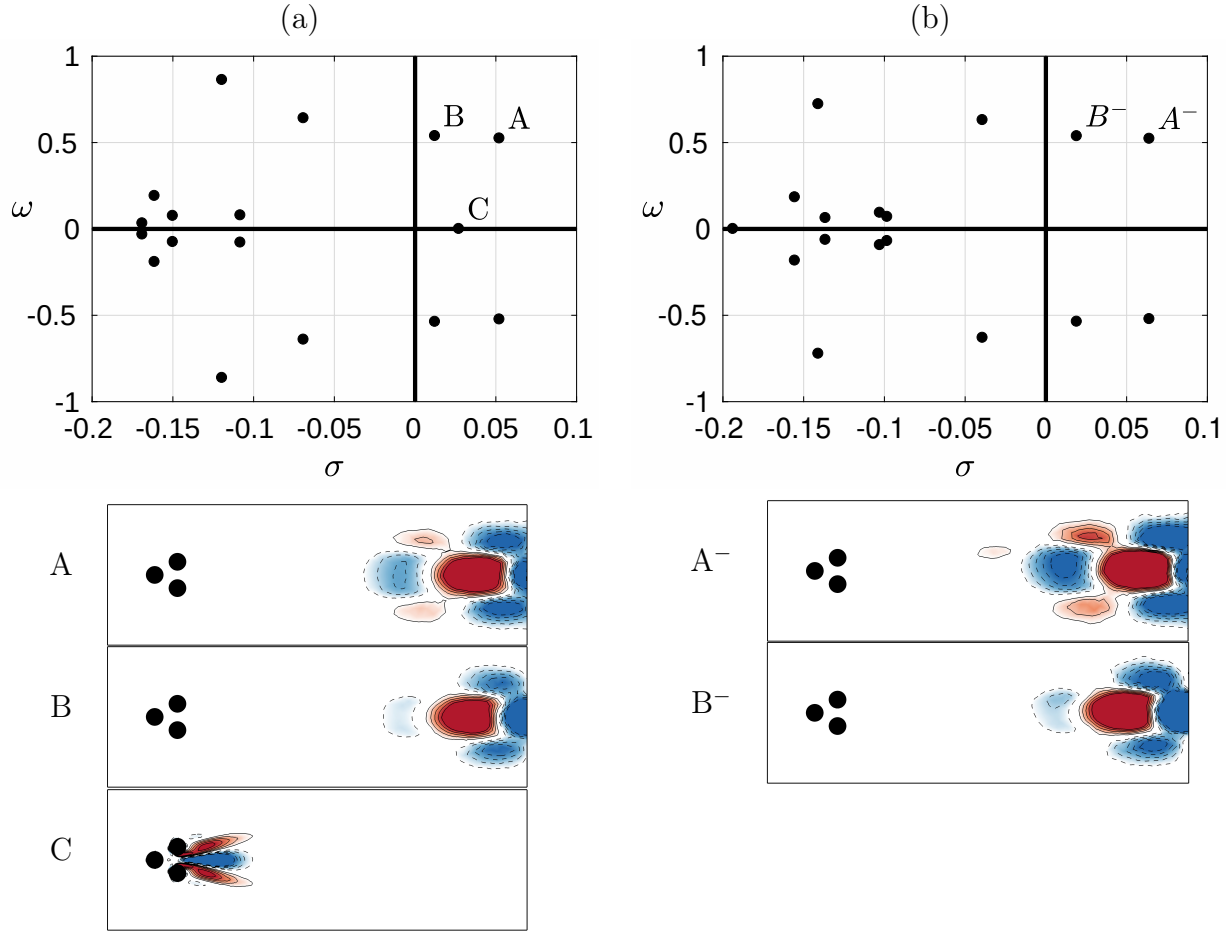


Figure 2.10: Eigenspectrum (top) and real part of the eigenvectors (bottom) of the symmetry-preserving steady solution $\boldsymbol{u}_s$ (left), of the symmetry-breaking steady solution $\boldsymbol{u}_s^-$ (right), both at $Re = 80$. The red color and solid contours in the eigenvectors are positive values of the vorticity, blue color and dashed contours are negative values.

conjugated asymmetric steady solutions $\boldsymbol{u}_s^{\pm}$, also unstable to vortex shedding but only existing beyond Re_2, see figure 2.3.

The linear stability analysis of $\boldsymbol{u}_s$ reveals two pairs of complex-conjugated eigenmodes with positive growth rate, and one eigenmode of zero frequency, see figure 2.10(a). The steady eigenmode is antisymmetric and reflects the symmetry broken by the pitchfork bifurcation. It is clearly associated with the base-bleeding jet, with all its energy concentrated in the near-field. The two pairs of complex-conjugated eigenmodes are each associated with von Kármán streets of shed vortices. Both pairs of complex eigenmodes are antisymmetric and have quite similar angular frequencies. A closer view of the second pair of complex eigenmodes indicates that its growth rate cancels when the real eigenmode crosses the zero axis. This indicates that the new oscillatory mode is intimately connected to the symmetry breaking occurring at Re_2. At $Re > Re_2$, this gives rise to the only stable limit cycle for the flow dynamics, while the limit cycle associated with the leading pair of complex eigenmode has become unstable and can only be visited transiently in time.

The linear stability analysis of $\boldsymbol{u}_s^-$ (resp. $\boldsymbol{u}_s^+$) reveals two pairs of complex-conjugated eigenmodes with positive growth rate, centered on an asymmetric mean flow, see figure 2.10(b). All eigenmodes are asymmetric, a property inherited from the steady solution.

In the permanent regime, the mean flow field will inherit the symmetry of one of the three (unstable) steady solutions, depending on the details of the initial perturbation.

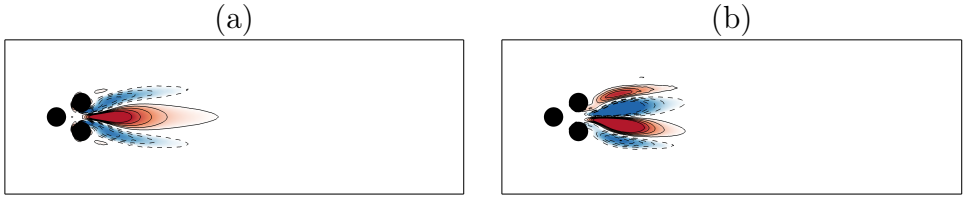


Figure 2.11: Additional modes arising from the pitchfork bifurcation, at $Re = 80$, (a) mode $\boldsymbol{u}_4$, (b) mode $\boldsymbol{u}_5$. Red color and solid contours are positive values of the vorticity, blue color and dashed contours are negative values. See text for details about the computation of these two modes.

σ_1	5.22×10^{-2}	β	1.31×10^{-2}
ω_1	5.24×10^{-1}	γ	2.95×10^{-2}
σ_3	-5.22×10^{-1}	β_3	1.53×10^{-1}
σ_4	2.72×10^{-2}	β_4	2.45×10^{-1}
σ_5	-2.72×10^{-1}	β_5	2.14×10^{-1}

Table 2.3: Coefficients of the least reduced-order model (2.32) at $Re = 80$.

2.5.2 Reduced-order model in the secondary flow regime

As discussed in § 2.3, an ansatz of the flow state can now be written as:

$$\boldsymbol{u}(\boldsymbol{x},t) \approx \underbrace{\boldsymbol{u}_s(\boldsymbol{x})}_{} + \underbrace{a_1(t)\boldsymbol{u}_1(\boldsymbol{x}) + a_2(t)\boldsymbol{u}_2(\boldsymbol{x})}_{\text{leading POD modes at } Re=80} + \underbrace{a_3(t)\boldsymbol{u}_3(\boldsymbol{x})}_{\text{shift mode}} + \underbrace{a_4(t)\boldsymbol{u}_4(\boldsymbol{x}) + a_5(t)\boldsymbol{u}_5(\boldsymbol{x})}_{\text{pitchfork degrees of freedom}} \tag{2.37}$$

It is worthwhile noticing that, in the frame of this ansatz, the two asymmetric steady solutions $\boldsymbol{u}_s^{\pm}$ are related to the symmetric steady solution $\boldsymbol{u}_s$ *via* the additional antisymmetric mode $\boldsymbol{u}_4$:

$$\boldsymbol{u}_s^{\pm} = \boldsymbol{u}_s \pm \overline{a}_4\boldsymbol{u}_4 + \overline{a}_5\boldsymbol{u}_5, \tag{2.38}$$

where $\overline{a}_4$ and $\overline{a}_5$ are the time-averaged coefficients in the permanent regime. Consequently, $\boldsymbol{u}_4$ can be easily computed as:

$$\boldsymbol{u}_4 \propto (\boldsymbol{u}_s^+ - \boldsymbol{u}_s^-), \tag{2.39}$$

and further orthonormalized to u_1, u_2, u_3 by a Gram-Schmidt procedure. The resulting mode $\boldsymbol{u}_4$ is shown in figure 2.11(a). A comparison with the eigenmode associated with the real eigenvalue, in figure 2.10(a), shows that the shift mode $\boldsymbol{u}_4$ is just the real eigenmode against which the symmetric steady solution is unstable at $Re = 80$, as expected by the definition of mode $\boldsymbol{u}_4$.

In a similar way, the additional mode $\boldsymbol{u}_5$ can be constructed as:

$$\boldsymbol{u}_5 \propto (\boldsymbol{u}_s^+ + \boldsymbol{u}_s^-)/2 - \boldsymbol{u}_s. \tag{2.40}$$

Mode $\boldsymbol{u}_5$ is shown in figure 2.11(b) after orthonormalization.

Close to the pitchfork bifurcation, the resulting dynamical system is described by Eqs. (2.32). At the threshold, the degrees of freedom a_4, a_5 associated with the pitchfork bifurcation are expected to be fully uncoupled to the degrees of freedom a_1, a_2, a_3 associated with the Hopf bifurcation, and reciprocally, see § 2.3.3. In this case, an accurate linear and nonlinear dynamics from a Galerkin projection relies on deformable modes from eigenmodes near the fixed point to POD modes near the limit cycle (Loiseau *et al.*, 2018*a*). We avoid this complication by model identification from simulation data. The coefficients of the mean-field system (2.32) reported in table 2.3 were identified from the linear stability analysis of the symmetric steady solution and the asymptotic dynamics on the unstable symmetric preserving limit cycle and the stable symmetric-breaking limit cycle (see § 2.3.3 to 2.3.5). The resulting ROM dynamics is compared to the flow

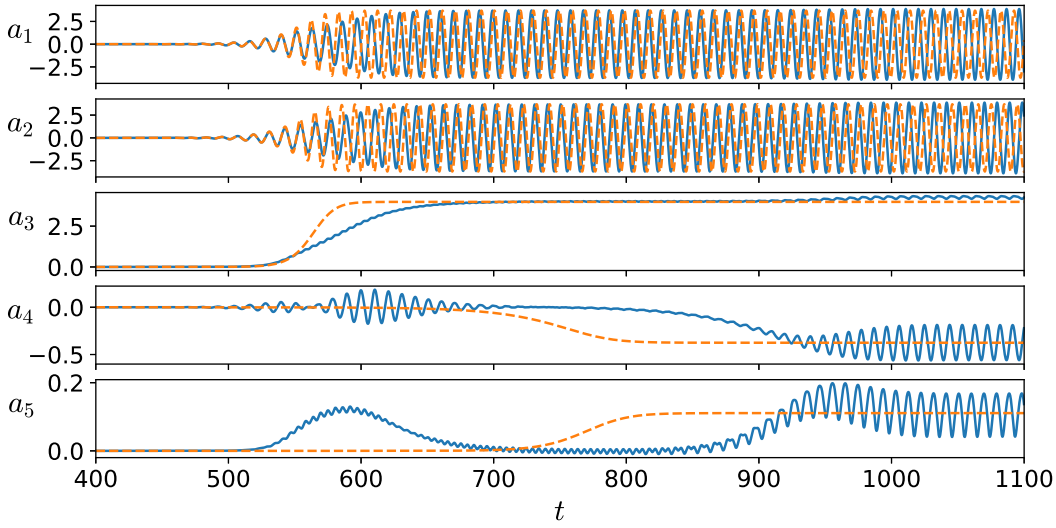


Figure 2.12: Time evolution of coefficients a_1 to a_5 in the full flow dynamics (solid blue line) and the ROM (dashed red line) without cross-terms and coefficients fixed by the linear stability analysis and the asymptotic dynamics. The initial condition for both systems is the same.

dynamics in figure 2.12. Inspection of figure 2.12 shows that such a model, reduced to only five degrees of freedom, is able to reproduce many features of the original dynamics: both the early transient and asymptotic dynamics of a_1 to a_3 are well reproduced, as well as the large timescale evolution of a_4, a_5. However, the growth of coefficients a_1 to a_3 appears to be faster for the ROM, and a_4, a_5 also reach their asymptotic value significantly sooner than their values for the full simulations of the fluidic pinball. In addition, the transient kick in a_4, a_5 is absent from the mean-field system (2.32), as well as the transient and asymptotic oscillations of a_4 and a_5 visible in figure 2.12, which would require coupling to a_1 or a_2, or both. All these features indicate that at Reynolds number $Re = 80$, the Krylov-Bogoliubov assumption of pure harmonic behaviour with slowly varying amplitude and frequency no longer holds.

As a consequence, cross-terms must be included in the ROM. The assumption of non-oscillatory dynamics of the shift mode amplitude a_3 and of the two pitchfork modes a_4, a_5 is relaxed to reproduce the oscillatory behaviour evidenced in figure 2.12. Following § 2.3.6, the model identification process reads:

Step 1: Keep the five-dimensional linear-quadratic form of the dynamical system from the Galerkin projection with 25 linear and 75 quadratic terms.

Step 2: Remove vanishing terms arising from the symmetry of the modes. Thus, only 13 linear and 36 quadratic terms are left to be determined.

Step 3: Enforce the linear dynamics of the unstable Hopf and pitchfork eigenmodes from stability analysis in the Galerkin system. This implies that the growth rate σ_1 and frequency ω_1 characterize the initial growth and angular frequency of a_1, a_2 and σ_4 represents the linear growth rate for a_4.

Step 4: The slaving of the Reynolds-stress-induced modes $\boldsymbol{u}_3$ and $\boldsymbol{u}_5$ to the fluctuation level is imposed by setting the damping rate 10 times larger than the growth rate of the corresponding fluctuation: $\sigma_3 = -10\sigma_1$, $\sigma_5 = -10\sigma_4$. This strong damping rate quickly forces the trajectory onto the mean-field manifold.

Step 5: Enforce phase invariance for a_1, a_2 in the first two equations. This is implied by the mean-field theory and is found to be a good approximation from numerical inspection. Thus, the oscillatory dynamics of a_1, a_2 are governed by σ_1, ω_1, β, β_{15}, γ, γ_{15}, and β_3.

Step 6: Impose the asymptotic dynamics of the unstable symmetric limit cycle by fixing β, γ, and β_3.

Step 7: Apply the SINDy algorithm (Brunton *et al.*, 2016*a*) to the remaining unknown terms, i.e. 6 linear terms and 28 quadratic terms. This step alone typically fails to yield a physics-based globally stable Galerkin system. This is not surprising in view of the necessary and sufficient conditions for global boundedness of Galerkin systems by Schlegel & Noack (2015). A Galerkin system identification with $\ell 1$-norm penalization of the coefficients is found to have a performance similar to SINDy for the chosen constraints.

Step 8: Two additional simplifying physics-based constraints are found to make the dynamic system identified by SINDy robust for a large range of initial conditions. Enforcing $q_{145} = q_{245} = 0$ avoids the initial oscillatory dynamics being influenced by the initial kick of a_5, and setting $q_{344} = q_{434} = 0$ leads to the right asymptotic transition of the pitchfork bifurcation.

The resulting ROM reads:

$$
\begin{aligned}
da_1/dt &= a_1(\sigma_1 - \beta\ a_3 - \beta_{15}\ a_5) - a_2(\omega_1 + \gamma\ a_3 + \gamma_{15}\ a_5) + l_{14}\ a_4 + q_{134}\ a_3 a_4, && \text{(2.41a)}\\
da_2/dt &= a_2(\sigma_1 - \beta\ a_3 - \beta_{15}\ a_5) + a_1(\omega_1 + \gamma\ a_3 + \gamma_{15}\ a_5) + l_{24}\ a_4 + q_{234}\ a_3 a_4, && \text{(2.41b)}\\
da_3/dt &= \sigma_3\ a_3 + \beta_3\ r^2 + l_{35}\ a_5 + q_{314}\ a_1 a_4 + q_{335}\ a_3 a_5 + q_{355}\ a_5^2, && \text{(2.41c)}\\
da_4/dt &= \sigma_4\ a_4 - \beta_4\ a_4 a_5 + a_1(l_{41} + q_{413}\ a_3 + q_{415}\ a_5) + a_2(l_{42} + q_{423}\ a_3 + q_{425}\ a_5), && \text{(2.41d)}\\
da_5/dt &= \sigma_5\ a_5 + \beta_5\ a_4^2 + l_{53}\ a_3 + q_{514}\ a_1 a_4 + q_{533}\ a_3^2 + q_{535}\ a_3 a_5. && \text{(2.41e)}
\end{aligned}
$$

where $r^2 = a_1^2 + a_2^2$. The coefficients are summarized in table 2.4. The dynamics of the system (2.41) (dashed red line) is compared to the full flow dynamics of the fluidic pinball (solid blue line) in figure 2.13. Now the initial stage of the dynamics is much better reproduced, as well as the asymptotic oscillations of a_3, a_4 and a_5. Interestingly, the faster growth in a_1 to a_3, on the time-range around 600, could not be completely corrected. It is worthwhile noticing that this range of time also corresponds to oscillations in a_4, which could not be reproduced by any cross-terms compatible with the symmetries of the system. Noticing that $a_4 \neq 0$ on this range of time may also question our choice for $\sigma_1 = 5.22 \times 10^{-2}$, which is the linear growth rate of the leading eigenmode around the symmetric steady solution $\boldsymbol{u}_s$. Indeed, although the initial condition is close to this point, the large amplitude oscillations of a_4 on the time range around 600 mean that the trajectory transiently escapes the symmetric subspace not only along a_3, but also along a_4 and a_5, before coming back close to the a_3 axis, on the time range from 700 to about 800. Therefore, it would be reasonable here to keep all the coefficients of the model unconstrained, but the number of free parameters is now too large for the identification process to be computationally tractable — for instance, it provides positive σ_3 when it must necessarily be strongly negative.

Model identification can be very challenging for a number of reasons. First, the conditions for global boundedness of the attractor for the linear-quadratic Galerkin system are rather restrictive (Schlegel & Noack, 2015). Second, the Galerkin method assumes fixed expansion modes. Yet, least-order models often have deformable modes changing with the fluctuation level (Tadmor *et al.*, 2011). The von Kármán vortex shedding metamorphosis from stability modes to POD modes may serve as an example. These deformations may also affect the structure of the dynamical system. Third, the Navier-Stokes dynamics may live on a strongly attracting manifold. This restriction of the state space may make certain Galerkin system coefficients numerically unobservable, like σ_3 and σ_5 in our case. Despite these challenges, the model identified in table 2.4, as exemplified by figure 2.13, constitutes a faithful least-order model for the fluidic pinball at $Re = 80$. Although

σ_1	5.22×10^{-2}	β	1.31×10^{-2}	l_{14}	2.93×10^{-1}	l_{24}	-4.87×10^{-1}
ω_1	5.24×10^{-1}	γ	2.95×10^{-2}	q_{134}	-5.87×10^{-2}	q_{234}	1.18×10^{-1}
σ_3	-5.22×10^{-1}	β_3	1.53×10^{-1}	l_{41}	3.14×10^{-2}	l_{42}	-5.14×10^{-2}
σ_4	2.72×10^{-2}	β_4	5.78×10^{-2}	q_{413}	-7.56×10^{-3}	q_{423}	1.28×10^{-2}
σ_5	-2.72×10^{-1}	β_5	1.91×10^{-1}	q_{415}	2.99×10^{-2}	q_{425}	1.71×10^{-1}
		β_{15}	-2.42×10^{-2}	l_{35}	4.28	l_{53}	2.89×10^{-2}
		γ_{15}	1.70×10^{-2}	q_{335}	-1.11	q_{533}	-7.22×10^{-3}
				q_{355}	-5.13×10^{-1}	q_{535}	1.48×10^{-2}
				q_{314}	1.57×10^{-2}	q_{514}	-9.44×10^{-3}

Table 2.4: Coefficients of the reduced-order model at $Re = 80$. See text for details.

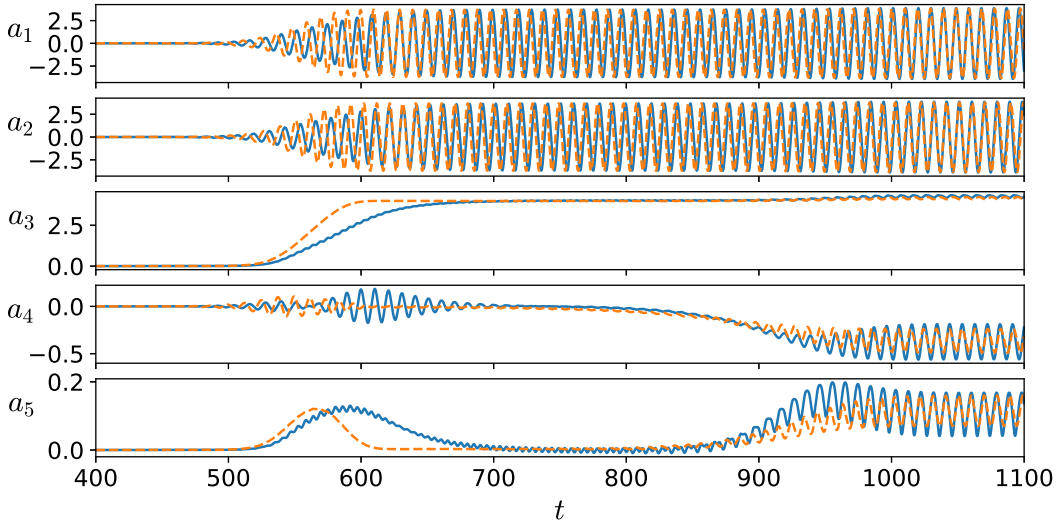


Figure 2.13: Performance of the ROM with cross-terms. Time evolution of coefficients a_1 to a_5 in the full flow dynamics (solid blue line) and for the ROM (red dashed line). The initial condition is the same for the ROM and the full flow dynamics.

the model is only five-dimensional, it can reproduce most of the key features, timescales, transient and asymptotic behaviour of the full dynamics.

2.6 Conclusions and outlooks

Reduced-order models (ROM) serve a number of purposes. For instance, ROMs facilitate a deeper understanding of the physical mechanisms at play in a flow configuration, by extracting the low-dimensional manifold on which evolve the dynamics (Manneville, 2010). In that respect, the linear stability analysis of the steady solution shows the nonlinear amplitude saturation mechanism through the distorted mean flow. The difference between steady solution and the mean flow is caused by the Reynolds stress, captured by the shift mode $\boldsymbol{u}_\Delta$, and affects the stability properties (Noack *et al.*, 2003; Barkley, 2006; Sipp & Lebedev, 2007; Turton *et al.*, 2015). Deeper investigations would certainly deserve to be carried out on the relation between the stability analysis of fixed points, Floquet analysis of limit cycles and Lyapunov exponents of chaotic flow regimes. In addition, because Hopf and pitchfork bifurcations are generic bifurcations in fluid flows, the nonlinear dynamics identified in this study are expected to be extended and generalizable to other flows exhibiting similar bifurcations. Last but not least, a ROM provides fast estimators for predicting

the forward evolution of the system. Such estimators could be used for control purposes (Brunton & Noack, 2015; Rowley & Dawson, 2017). All these considerations motivated the present study, whose main results are summarized in § 2.6.1. Outlooks of this work are listed in § 2.6.2.

2.6.1 Concluding remarks and discussion

Flow configurations undergoing successive Hopf and pitchfork bifurcations are common in fluid mechanics. This is, for instance, the case of three-dimensional wake flows such as spheres (Mittal, 1999; Gumowski *et al.*, 2008; Szaltys *et al.*, 2012; Grandemange *et al.*, 2014) or bluff body wake flows (Grandemange *et al.*, 2012, 2013; Cadot *et al.*, 2015; Bonnavion & Cadot, 2018; Rigas *et al.*, 2014). The drag crisis and stalled flows are also characterized by the pitchfork bifurcation of a primarily Hopf-bifurcated flow, but the secondary transition is subcritical in this case.

In this study, we have considered the fluidic pinball on its way to chaos and have identified least-order models of the flow dynamics in the primary Hopf-bifurcated and secondary pitchfork-bifurcated flow regimes. Reduced-order modelling of Hopf bifurcations was already addressed in Noack *et al.* (2003) for the cylinder wake flow, while Meliga *et al.* (2009) derived the amplitude equation for a codimension two bifurcation (pitchfork and Hopf) based on a weakly nonlinear analysis in the wake of a disk, and Fabre *et al.* (2008) derived the same equation solely based on symmetry arguments, in the wake of axisymmetric bodies. In the present contribution, we could demonstrate that the dynamics resulting from the successive Hopf and pitchfork bifurcation could be well-captured by a five-dimensional model whose degrees of freedom couple through quadratic non-linearities, as imposed by the Navier-Stokes equations.

For the fluidic pinball, the route to chaos is characterized by a primary supercritical Hopf bifurcation at $Re \approx 18$, followed by a secondary supercritical pitchfork bifurcation at $Re \approx 68$. The Hopf bifurcation corresponds to the destabilization of the steady solution with respect to vortex shedding, while the pitchfork bifurcation occurs when the mean flow breaks the symmetry with respect to the mirror-plane. The fluctuation amplitude of the von Kármán street is reduced, over a finite range of the Reynolds number around Re_2, when the base-bleeding jet is rising. This means that energy is withdrawn from the fluctuations to feed the mean flow transformation. Before the next transition occurs, the fluctuation amplitude starts to grow again, up to the largest value of the Reynolds number considered in this work.

There is strong evidence that the next transition is a Neimark-Säcker bifurcation. The resulting flow regime is most likely quasi-periodic over the range $[Re_3, Re_4]$. In this regime, a new oscillatory phenomenon takes place, characterized by slow oscillations of the base-bleeding jet. Three additional degrees of freedom might be necessary to deal with the newly arising oscillator. The flow dynamics eventually bifurcates into a chaotic regime, characterized by the random switching of the base-bleeding jet between two symmetric deflected positions. The overall route to chaos is summarized in the phenomenogram of figure 2.14.

The reduced-order models derived by Galerkin projections of the Navier-Stokes equations, based on the symmetry of the individual degrees of freedom, under Krylov-Bogoliubov simplifications, faithfully extract the manifolds on which the flow dynamics sets in. The ROM for the primary flow regime is only three dimensional: two degrees of freedom are associated with the asymptotic stable limit cycle resulting from the vortex shedding. The third degree of freedom is a mode slaved to the two dominant modes and is mandatory for the description of the transient flow dynamics from the unstable steady solution to the post-transient mean flow, as already demonstrated in Noack *et al.* (2003). The least-order model in the secondary flow regime has only five degrees of freedom, three of which are associated with the Hopf bifurcation, the two remaining degrees of freedom being associated with the pitchfork bifurcation. In the phenomenogram of figure 2.14 are reported the structure of both reduced-order models close to the Hopf and the pitchfork bifurcations. When the two sets of degrees of freedom are fully uncoupled, some features of the flow dynamics are well-reproduced (asymptotic mean behaviour, parabolic manifolds of the Hopf and pitchfork bi-

furcations), but many details are missing. To reproduce most of the transient and asymptotic flow features far from the bifurcation point, additional cross-terms have been included in the model, which relaxes the steadiness constraint usually assumed for the shift modes.

2.6.2 Outlook

The current generalized mean-field model captures the Hopf bifurcation and subsequent pitchfork bifurcation of the steady solution and limit cycles. The following onset of a quasi-periodic regime with slow oscillations of the deflected base-bleeding jet might presumably be incorporated by another Hopf bifurcation, leading to a 8-dimensional mean-field Galerkin model. The transition to chaos is accompanied by a return to a statistically symmetric flow, i.e. the base-bleeding jet oscillates around one asymmetric state before it stochastically switches to the other mirror-symmetric one. This behaviour is reminiscent of the transition to chaos of a harmonically forced Duffing oscillator. In the case of the fluidic pinball, the vortex shedding would constitute a forcing. Hence, one may speculate that the transition to chaos may already be resolved by the 8-dimensional Galerkin model in which the effect of vortex shedding on the jet oscillation becomes stronger with an increasing Reynolds number.

An alternative direction is to increase the accuracy of the mean-field Galerkin model. While the structure of the Galerkin system prevails for a large range of Reynolds numbers, the modes and all Galerkin system coefficients change in a non-trivial manner, e.g., the growth-rate formula should read $\sigma = \sigma_1(Re) - \beta(Re)a_3$. The transients can be expected to be much more accurately resolved by the mean flow dependent modes, e.g. $\boldsymbol{u}_i(Re, a_3, a_5, \vec{x})$, $i = 1, 2$ for the resolution of vortex shedding (Loiseau *et al.*, 2018*a*). More generally, the flow lives on a low-dimensional manifold which includes mode deformations (Noack, 2016). Locally linear embedding (LLE) is a powerful technique for identifying the dimension and a parameterization in an automatic manner (Roweis & Saul, 2000). The normal form of the bifurcations can be expected to coincide with the dynamics on the LLE feature coordinates.

A third direction follows an observation of Rempfer (1994) that Galerkin systems of many fluid flows can be considered as nonlinearly coupled oscillators. For two incommensurable shedding frequencies, this observation has been formalized in a generalized mean-field model by Noack *et al.* (2008); Luchtenburg *et al.* (2009). Such multi-frequency models may be extended to resolve broadband frequency dynamics taking, for instance, the most dominant DMD modes (Rowley *et al.*, 2009; Schmid, 2010). Strengths and weaknesses of techniques currently used for model reduction are discussed in Taira *et al.* (2017). While mean-field consideration expressly ignores non-trivial triadic interactions, their quantitative effect on the frequency crosstalk may still be well approximated by the mean flow interaction terms. Such multi-frequency mean-field models may eventually describe the effect of open-loop forcing on turbulence, see for instance the recent thorough review by Jiménez (2018) on turbulent flow modelling.

A fourth direction aligned with the large success of machine learning / artificial intelligence is the automated learning of state spaces, modes, and dynamical systems. For the latter, SINDy provides an established elegant framework (Brunton *et al.*, 2016*a*). The choice of the state spaces might be facilitated by manifold learning from many solution snapshots (Gorban & Karlin, 2005). The authors actively pursue all the mentioned directions.

Acknowledgements

This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissement d'Avenir" program, through the "iCODE Institute project" funded by the IDEX Paris-Saclay, ANR-11-IDEX-0003-02, by the ANR grants 'ACTIV_ROAD' and 'Flow-Con' (ANR-17-ASTR-0022), and by Polish Ministry of Science and Higher Education (MNiSW) under the Grant No.: 05/54/DSPB/6492.

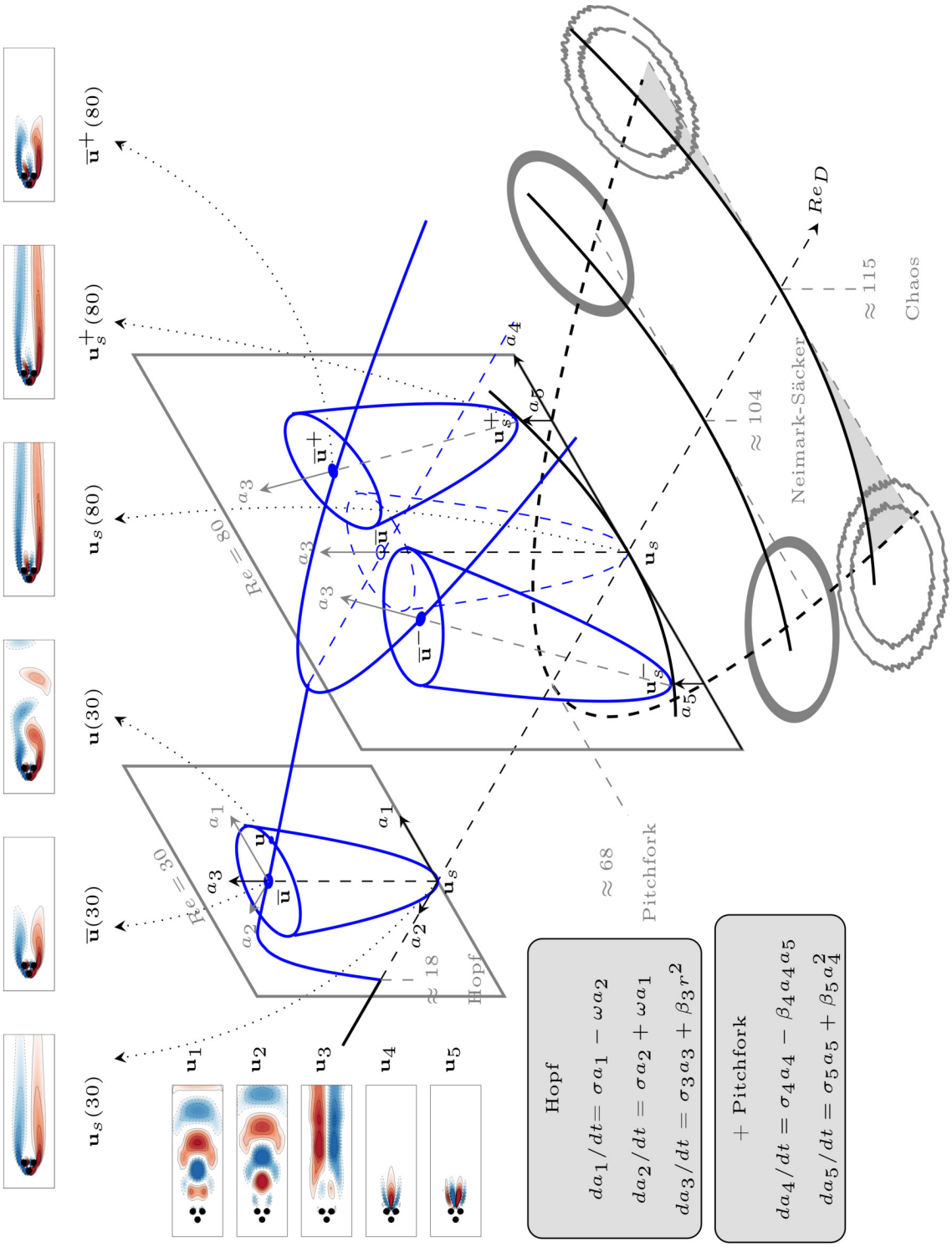


Figure 2.14: Phenomenogram. The route to chaos develops along the Re_D-axis. On this route are highlighted the least-order models identified at $Re = 30$ and $Re = 80$, where the manifolds on which the dynamics take place are schematically drawn. The branches of the steady solution and periodic solution are presented by the black and blue curves, with the solid/dashed curves for the stable/unstable states. The degrees of freedom a_1,a_2 span the limit cycle subspace, a_3 is the axis of the parabolically-shaped manifold, while a_4 is transversally associated with the pitchfork bifurcation, together with a_5 which slightly bend the surface to which the steady solutions $\boldsymbol{u}_s^{\pm}$ and $\boldsymbol{u}_s$ belong. The gray-shaded shadows of the quasi-periodic and chaotic regimes are represented for the sake of illustration. Gray-shaded inserts provide the least-reduced order models at the threshold for both the Hopf and pitchfork bifurcations under the constraint of the Navier-Stokes equations. Also shown at the top are the figure snapshots of the steady solutions, mean flow fields, and instantaneous flow field, at either $Re = 30$ or $Re = 80$. The dotted arrows connect these snapshots to the corresponding points in the phenomenogram. In addition, individual degrees of freedom $\boldsymbol{u}_1$ to $\boldsymbol{u}_5$ are shown at the left.

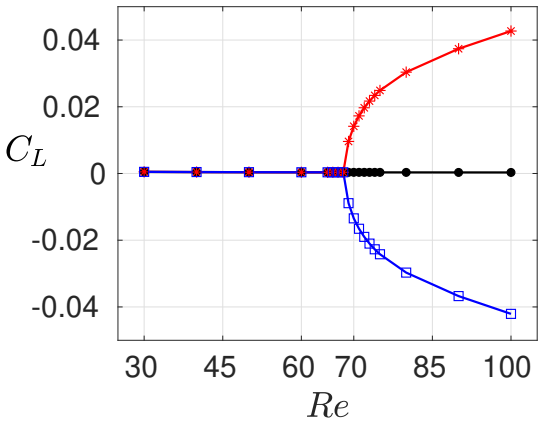


Figure 2.15: Pressure lift coefficient of steady solutions at different values of the Reynolds number resulting from the steady Navier-Stokes solver starting with the steady solutions: $\boldsymbol{u}_s(\boldsymbol{x})$ (point + black curve), $\boldsymbol{u}_s^+(\boldsymbol{x})$ (star + red curve), and $\boldsymbol{u}_s^-(\boldsymbol{x})$ (square + blue line), at $Re = 80$. Three curves overlap on the $C_L = 0$ level as $Re \leq 68$.

We appreciate valuable stimulating discussions with Steven Brunton, Alessandro Bucci, Nathan Kutz, Onofrio Semeraro, Yohann Duguet, Laurette Tuckerman and the French-German-Canadian-American pinball team: François Lusseyran, Guy Cornejo-Maceda, Jean-Christophe Loiseau, Robert Martinuzzi, Cedric Raibaudo, Richard Semaan, and Arthur Ehlert.

2.A Asymmetric steady solutions

For a Reynolds number larger than the critical value of the pitchfork bifurcation Re_2, we can obtain two additional asymmetric steady solutions, one associated with the base-bleeding jet deflected upward, the other downward. These two asymmetric steady solutions are obtained by the steady Navier-Stokes solver initialized with a flow field (snapshot) with the same state of the base-bleeding jet. From the two asymmetric vortex shedding, their corresponding asymmetric steady solutions can be obtained by the following steps:

Step 1: Run the unsteady Navier-Stokes solver with a time scale larger than the vortex shedding period, initialized from a snapshot of the asymmetric vortex shedding. The vortex shedding will quickly vanish for this artificially large time scale and approach the corresponding approximation of the steady state.

Step 2: Run the steady Navier-Stokes solver, restarted from this vortex shedding vanished solution to further refine the steady state.

Initialized with the three steady solutions with different states of the base-bleeding jet at $Re = 80$, the steady solutions for other Reynolds numbers can be obtained by the steady Navier-Stokes solver. At $Re = 68$, all three steady solutions converge to a unique solution, which indicates that the critical value of the pitchfork bifurcation of the steady solution is between 68 to 69, as shown in figure 2.15 based on the pressure lift coefficient of the steady solutions.

2.B Linear stability analysis

The linear stability problem for a base flow $(\boldsymbol{U}(\boldsymbol{x}), P(\boldsymbol{x}))$ with small perturbations $(\boldsymbol{u'}(\boldsymbol{x},t), p'(\boldsymbol{x},t))$ is governed by the linearized Navier-Stokes equations, which read:

$$\partial_t \boldsymbol{u'} + (\boldsymbol{U}\cdot\nabla)\,\boldsymbol{u'} + (\boldsymbol{u'}\cdot\nabla)\,\boldsymbol{U} = \nu\triangle\boldsymbol{u'} - \nabla p', \quad \nabla\cdot\boldsymbol{u'} = 0 \tag{2.42}$$

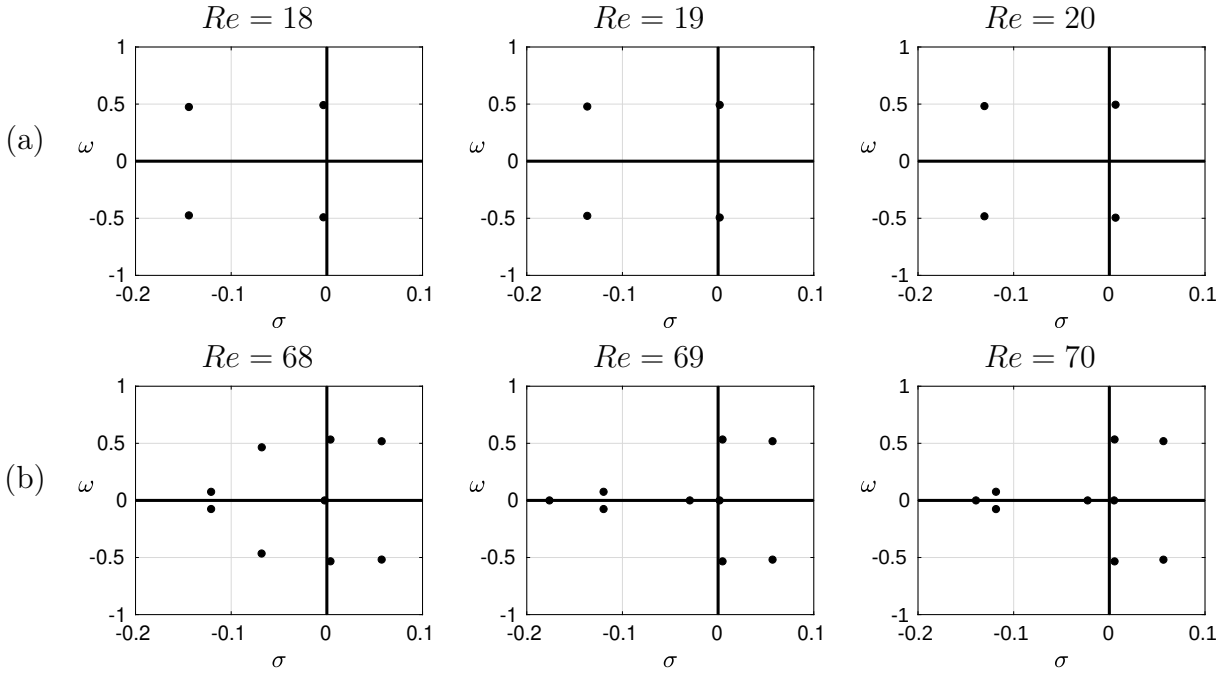


Figure 2.16: Eigenspectrum resulting from the linear stability analysis of the symmetric steady solution $\boldsymbol{u}_s$. With increasing Re, (a) a complex-conjugated eigenvalue pair crosses the imaginary axis at Re changing from 18 to 19, the critical value of the Hopf bifurcation $Re_1 \approx 18$, (b) a real eigenvalue crosses the imaginary axis at Re changing from 68 to 69, the critical value of the pitchfork bifurcation $Re_2 \approx 68$.

The assumption of small perturbation allows to linearize the equations, and we can separate the time and space dependence as:

$$\boldsymbol{u'}(\boldsymbol{x},t) = \hat{\boldsymbol{u}}(\boldsymbol{x})e^{(\sigma+i\omega)t}, \quad p'(\boldsymbol{x},t) = \hat{p}(\boldsymbol{x})e^{(\sigma+i\omega)t} \tag{2.43}$$

By introducing the linear operator $\boldsymbol{L}(\boldsymbol{U})$, all the terms except the time derivative term and the continuity equation can be cast as $\boldsymbol{L}(\boldsymbol{U})\boldsymbol{u'}$, we can rewrite (2.42) as

$$\partial_t \boldsymbol{u'} = \boldsymbol{L}(\boldsymbol{U})\boldsymbol{u'} \tag{2.44}$$

Introducing (2.43) into (2.44), the equations can be written as:

$$(\sigma + i\omega)\hat{\boldsymbol{u}} = \boldsymbol{L}(\boldsymbol{U})\hat{\boldsymbol{u}} \tag{2.45}$$

We use subspace iteration to solve this eigenvalue problem. A detailed review can be found in Morzyński *et al.* (1999).

The global stability analysis of the steady solutions at different Reynolds numbers has been performed on a Krylov subspace of dimension 9-20. This converges after 50-100 iterations. The linear stability analysis of the symmetric steady solution $\boldsymbol{u}_s$ reveals a pair of conjugated eigenvalues with positive real part first appearing as the Reynolds number is changing from 18 to 19, see figure 2.23(a). A real eigenvalue becomes positive between $Re = 68$ and 69, see figure 2.23(b). This confirms that a Hopf bifurcation occurs on the symmetric steady solution at $Re_1 \approx 18$, and a pitchfork bifurcation at $Re_2 \approx 68$.

2.C Floquet stability analysis

Similar to the linear stability framework, the Floquet stability problem works with a T-periodic base flow $(\boldsymbol{U}(\boldsymbol{x},t), P(\boldsymbol{x},t))$. The linear operator now reads:

$$\partial_t \boldsymbol{u'} = \boldsymbol{L}(\boldsymbol{U}(\boldsymbol{x},t))\boldsymbol{u'} \tag{2.46}$$

The linear operator $\boldsymbol{L}(\boldsymbol{U}(\boldsymbol{x},t))$ is T-periodic because of the base flow $\boldsymbol{U}(\boldsymbol{x},t)$. The solutions to (2.46) are seeked as:

$$\boldsymbol{u'}(\boldsymbol{x},t) = \hat{\boldsymbol{u}}(\boldsymbol{x},t)e^{(\sigma+i\omega)t}, \tag{2.47}$$

with the T-periodic Floquet modes $\hat{\boldsymbol{u}}(\boldsymbol{x},t)$ and the corresponding Floquet exponents $\sigma + i\omega$. We define the Floquet operator as the time-integrated $\boldsymbol{L}(\boldsymbol{U}(\boldsymbol{x},t))$ with the pre-stored periodic solutions over one period (Barkley & Henderson, 1996*a*; Schatz *et al.*, 1995), which reads:

$$A_F = \exp\left(\int_0^T \boldsymbol{L}(\boldsymbol{U}(\boldsymbol{x},t))dt\right). \tag{2.48}$$

The method used to solve the eigenproblem is the same as the linear stability analysis. The Floquet multipliers of A_F can be written as $\lambda_F = e^{(\sigma+i\omega)T}$. The considered periodic base flow is unstable if there exists a Floquet multiplier with an absolute value greater than 1.

We performed the block-Arnoldi method to construct a high-dimensional Krylov subspace, with itergrating the linear operator on a symmetry-constrained T-periodic base flow (Shaabani-Ardali *et al.*, 2019). The algorithm computes 20 vectors in parallel at each iteration, improving the construction efficiency. Below the critical Reynolds number, the base-bleeding jet is approximately steady and symmetric. This symmetry is also enforced at higher Reynolds numbers to compute the unstable periodic solution. The constraint is imposed on the central line as:

$$v(0 \leq x \leq 1, |y| \leq 5 \times 10^{-4}) = 0 \tag{2.49}$$

As Eq. (2.49) restricts the vertical velocity of the nodes, the symmetry-constrained periodic solution is very close to the symmetry-preserving periodic solution. Normally, 6–15 iterations are enough to get a converged leading eigenvalue, which means a Krylov subspace with 120–300 vectors, initialized with a group of 20 random orthogonal vectors or a group of eigenvectors computed at nearby Reynolds number with the periodic solution of the same symmetry. We do not attempt to calculate a complete, converged spectrum of all the eigenvalues, as the leading eigenvalue is associated with the instability of interest. The multipliers from the Floquet analysis around the critical value of the pitchfork bifurcation Re_2 are shown in figure 2.17. There always exists a stable

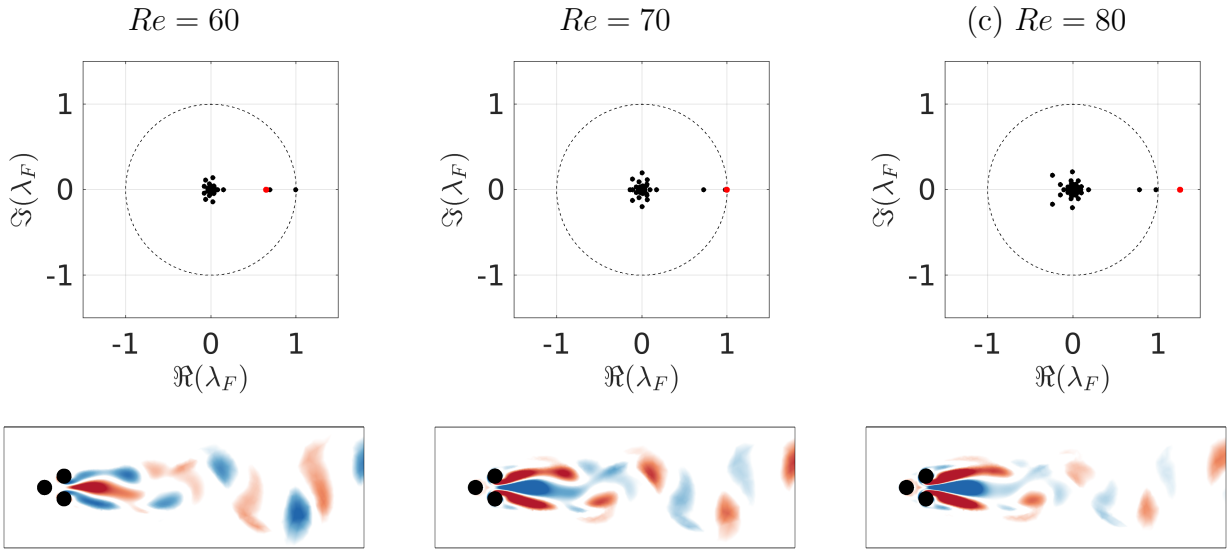


Figure 2.17: Floquet multipliers for the symmetry-preserving periodic solution at $Re = 60$, 70, and 80, together with the Floquet modes for the multipliers marked by red dots.

multiplier close to 1, which is related to the phase shift at each time step. The structure of this phase-shifting eigenmode, for instance, the mode B (B^+ for the asymmetry periodic solution), can

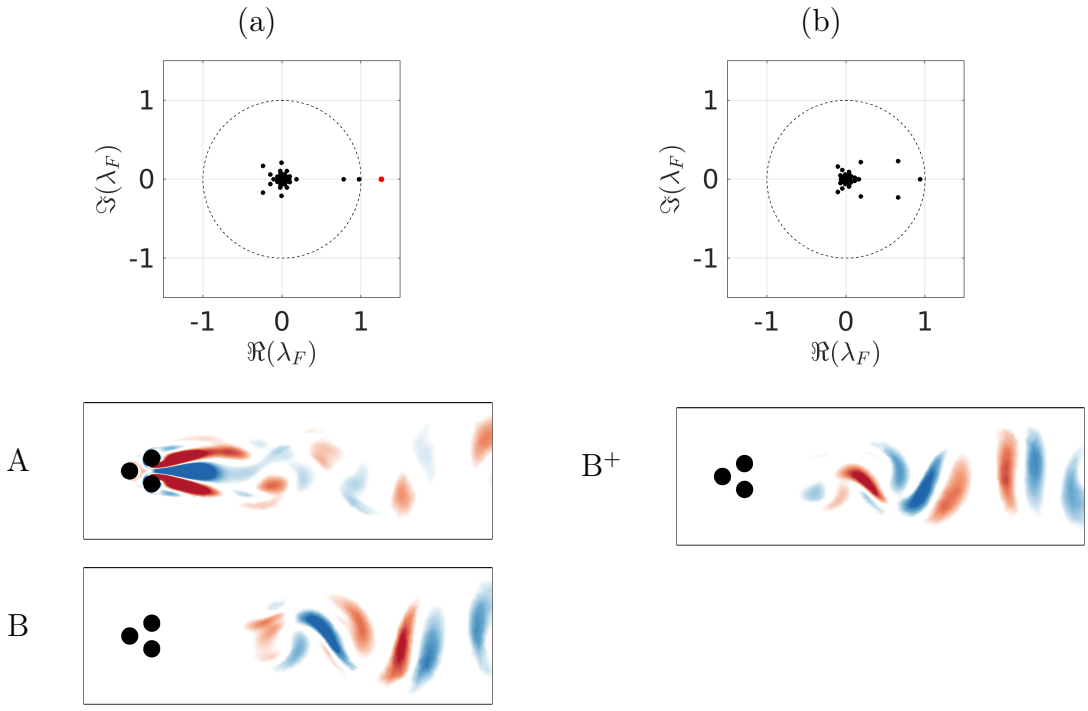


Figure 2.18: Multipliers resulting from the Floquet analysis (top) and the leading modes (bottom) of (a) the symmetry-preserving periodic solution, (b) the asymmetric periodic solution, both at $Re = 80$. Only the real part of the complex eigenmodes is shown. Red color and solid contours are positive values of the vorticity, blue color and dashed contours are negative values. Mode A (unstable multiplier marked with red dot) is related to the symmetry breaking of the periodic solution, and the modal energy is concentrated in the jet structure. Mode B or mode B^+ (stable multiplier close to the unit cycle) is marginally stable and comes from the phase shift at each time step of integrating the linear operator on the periodic base flow.

be found in figure 2.18. When increasing Re, the leading real eigenvalue crosses the unit cycle at $(+1, 0)$ as Re changes from 69 to 70. The critical value of the pitchfork bifurcation is, therefore, $Re_2 \approx 69$, identical to the critical value for the steady solution at the precision of the numerics. Both of them have the same eigenmode. At $Re = 80$, the leading Floquet modes of both the unstable symmetric periodic solution and the stable asymmetric periodic solutions are shown in figure 2.18.

Overall, combined with the result of the linear stability analysis of the steady solutions, the bifurcation scenario at low Reynolds numbers can be shown in figure 2.19. The linear stability analysis of the steady solution and the periodic solution show a highly consistent result: the same kind of bifurcation with nearly the same critical Reynolds number, and the same eigenmodes. Besides, at $Re = 80$, the growth rates of the real eigenmode are very close: 0.0272 from the symmetric steady solution, 0.0232 from the symmetric periodic solution. This similarity is understood as the result of a transverse effect of the symmetric subspace.

2.D Transient dynamics from different steady solutions

In this section, we show in figure 2.20 some typical transient dynamics starting with the unstable symmetric/asymmetric steady solutions at different Reynolds numbers, based on the pressure lift coefficient $C_L(t)$ from the resulting force on the three cylinders. Combining $C_L(t)$ with the pressure drag coefficient $C_D(t)$ and the time-delayed lift coefficient $C_L(t-\tau)$ in which τ is a quarter period, provides the phase portraits of figure 2.21. Three comparative numerical simulations are shown, starting with the symmetric and the two mirror-conjugated asymmetric steady solutions at the

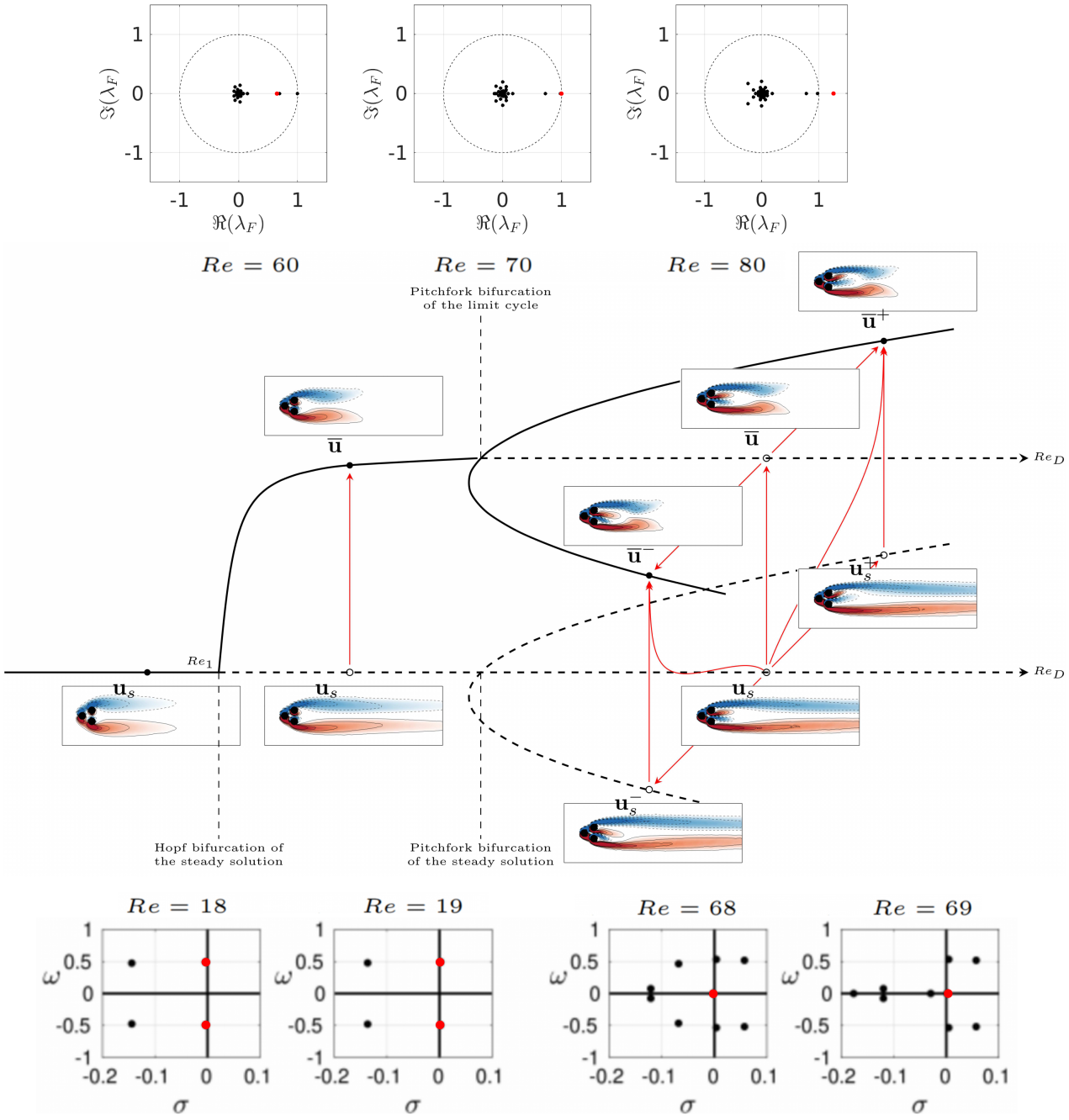


Figure 2.19: Bifurcation scenario along the Re_D-axis for the first two successive instabilities. The black curve indicates the stable branch, and dashed black curve for the unstable branch, combined with the linear/Floquet stability analysis results around the critical Reynolds number. The relevant eigenvalues at the critical value of the bifurcation are marked with red dots. The stability of the steady solutions $\boldsymbol{u}_s$ and $\boldsymbol{u}_s^{\pm}$, and the periodic solutions presented by the mean flow field $\overline{\boldsymbol{u}}$ and $\overline{\boldsymbol{u}}^{\pm}$ are illustrated by a dot for the stable state, or a circle for the unstable state. The red arrows show the possible transitions between them. The flow states in three stages are represented by the flow fields at $Re = 10, 30, 80$.

same Re respectively. The mirror-conjugated initial conditions provide mirror-conjugated transient dynamics.

Figure 2.20(a) shows the transient dynamics at the critical Reynolds number Re_2, initialized with the three steady solutions at $Re = 75$. The lift coefficient starts oscillating quickly, and eventually reaches a unique oscillating state with zero mean value. This is consistent with the Floquet analysis at $Re = 68$, where only one stable symmetry-centered limit cycle exists and any other state will eventually converge to this stable state.

Figure 2.20(b) shows three different scenarios at $Re = 75$ depending on the initial condition: from the symmetric steady solution $\boldsymbol{u}_s$ to the symmetry-centered limit cycle, from the symmetry-centered limit cycle to the asymmetry-centered limit cycles, and from the asymmetric steady solutions $\boldsymbol{u}_s^{\pm}$ to the asymmetry-centered limit cycles. Starting with the symmetric steady solution, it will first reach the unstable symmetry-centered limit cycle, before asymptotically approaching one of two stable asymmetry-centered limit cycles. However, starting with the asymmetric steady solutions, it will directly reach the corresponding stable asymmetry-centered limit cycle. If the initial perturbation introduced to the symmetric steady solution has a certain bias of symmetry, a transition from the symmetric steady solution $\boldsymbol{u}_s$ to one of the two asymmetry-centered limit cycles will occur.

As we keep increasing the Reynolds number up to $Re = 100$, there still exist six states, but the transient scenario from the symmetric steady solution $\boldsymbol{u}_s$ is different. It will first reach one of two unstable asymmetric steady solutions $\boldsymbol{u}_s^{\pm}$, before asymptotically approaching the corresponding asymmetry-centered limit cycle, as shown in figure 2.20(c).

All these transient dynamics mentioned above are highlighted with the red arrows in figure 2.19.

The phase portraits starting with different steady solutions, as shown in figure 2.21, reveal the above-mentioned transient dynamics which is affected by the initial condition and the Reynolds number. At the same time, it also reflects the global effect of the pitchfork bifurcation at Re_2, splitting the state space (see figure 2.21(a)) into a symmetric sub-space and two mirror-conjugated asymmetric sub-spaces (see figure 2.21(b),(c)).

2.E On the simultaneous instability of the fixed point and the limit cycle

In this section, we exemplify the transverse effect of the pitchfork bifurcation on a three-dimensional dynamical system equivalent to the system of Eq. (2.32).

Dynamical system

The dynamical system reads:

$$\begin{cases} \dot{x} &= (\mu - \mu_1 - (x^2+y^2))x + (\omega_0 + (x^2+y^2))y \\ \dot{y} &= (\mu - \mu_1 - (x^2+y^2))y - (\omega_0 + (x^2+y^2))x \\ \dot{z} &= (\mu - \mu_2)z - z^3 \end{cases} \tag{2.50}$$

with $\mu_1 = 1, \mu_2 = 2, \omega_0 = 1$. This system undergoes a supercritical Hopf bifurcation in the (x, y)-plane at $\mu = \mu_1$ and a supercritical pitchfork bifurcation along the z-axis at $\mu = \mu_2$. For $\mu > \mu_1$, the stable fixed point at $(0, 0, 0)$ becomes unstable, and the limit cycle around $(0, 0, 0)$ with radius $r = \sqrt{\mu - \mu_1}$ and angular frequency $\omega = \omega_0 + \mu - \mu_1$ is stable in the (x, y)-plane. Increasing μ until $\mu > \mu_2$, the fixed point undergoes a secondary instability, as well as the limit cycle. Three unstable fixed points $(0, 0, 0)$, $(0, 0, \pm\sqrt{\mu - \mu_2})$, and three limit cycles around these fixed points with radius $r = \sqrt{\mu - \mu_1}$ and angular frequency $\omega = \omega_0 + \mu - \mu_1$ in the (x, y)-plane are found, as shown in figure 2.22.

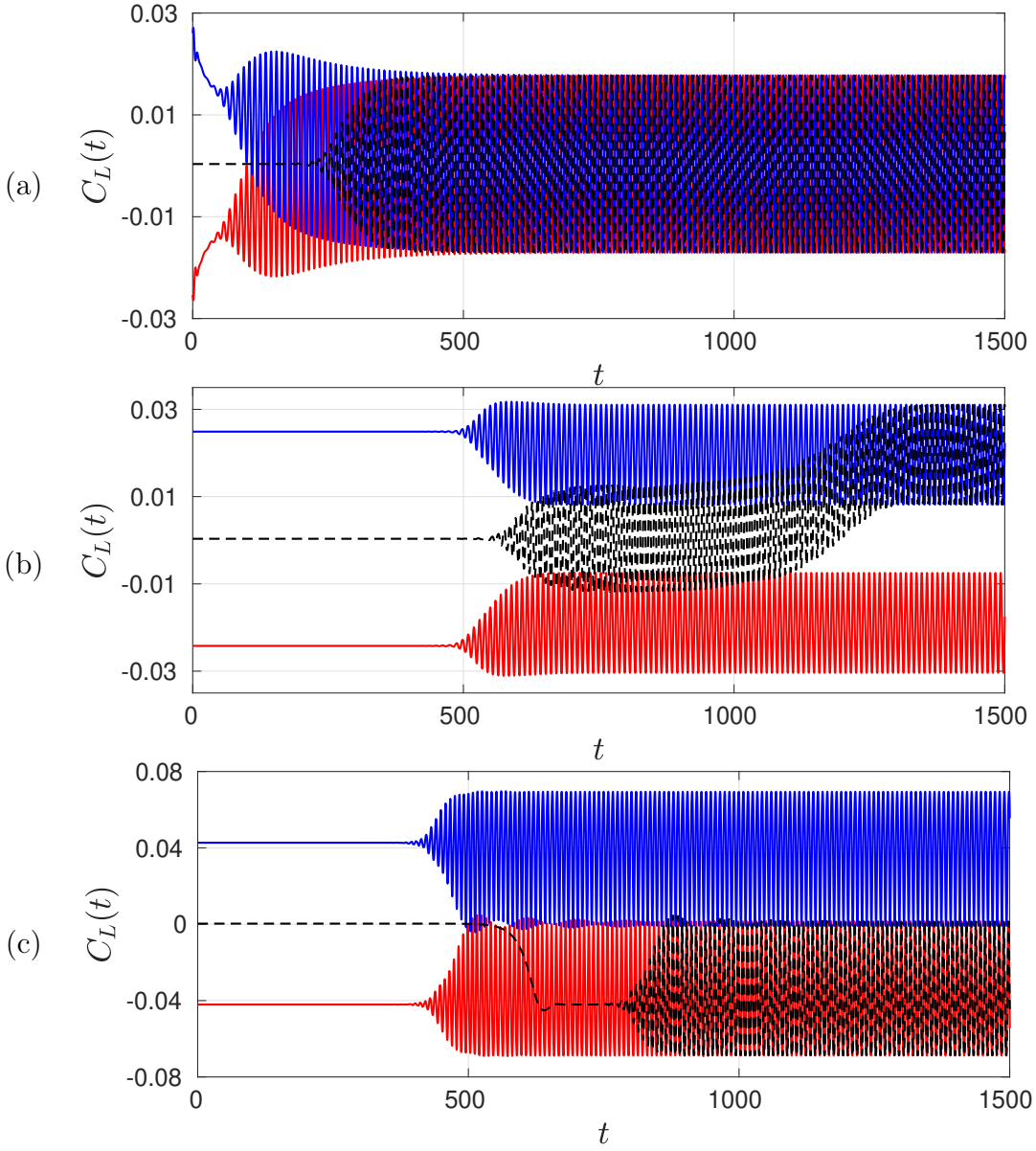


Figure 2.20: Transient dynamics based on the pressure lift coefficient C_L, resulting from the DNS starting with three steady solutions at different Reynolds numbers: (a) $Re = 68$ starting with three steady solutions at $Re = 75$, (b) $Re = 75$ starting with three steady solutions at $Re = 75$, and (c) $Re = 100$ starting with three steady solutions at $Re = 100$. The black dashed curve starts with the symmetric steady solution $\boldsymbol{u}_s$, the red curve starts with $\boldsymbol{u}_s^-$, and the blue curve starts with $\boldsymbol{u}_s^+$.

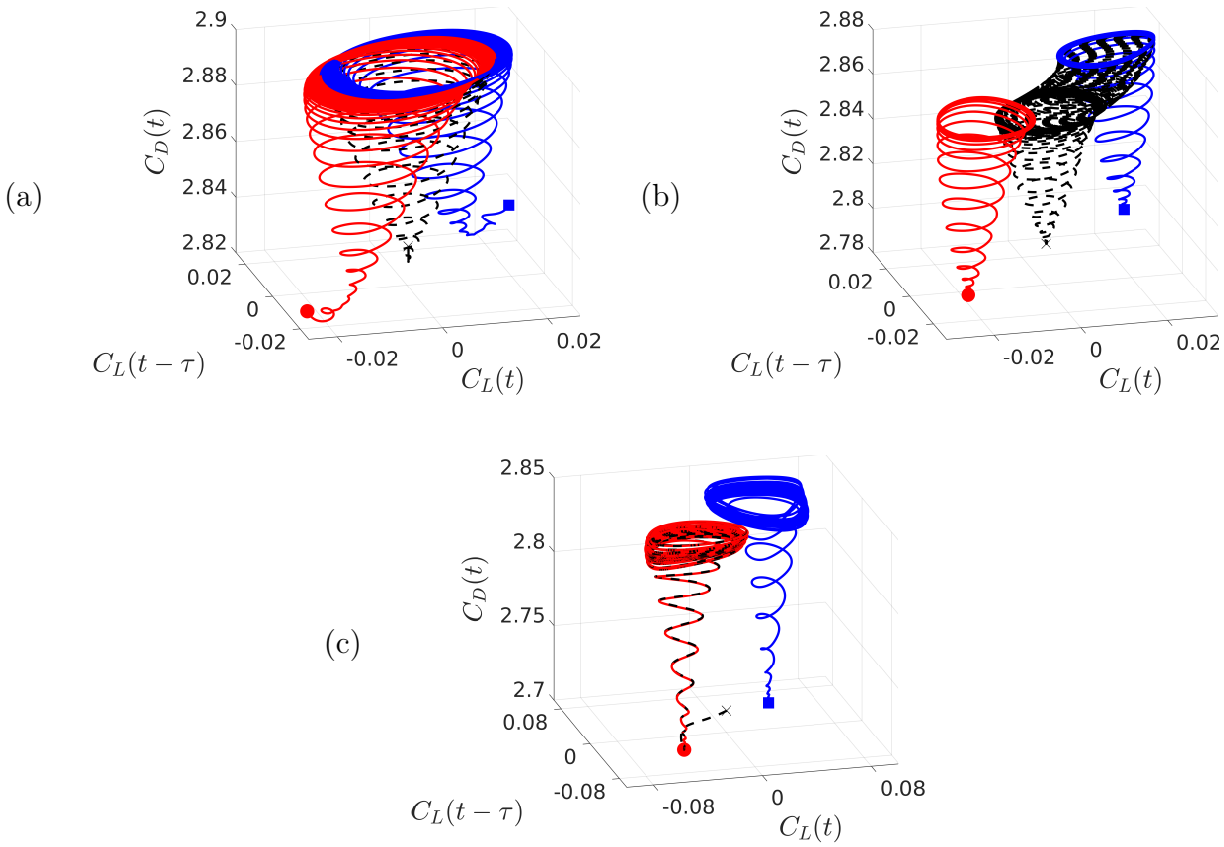


Figure 2.21: Scenarios of transient dynamics, based on the pressure drag coefficient $C_D(t)$, pressure lift coefficient $C_L(t)$ and time-delayed pressure lift coefficient $C_L(t-\tau)$, from the three unstable symmetric/asymmetric steady solutions of (a)$Re = 75$, (b)$Re = 75$ and (c)$Re = 100$, to their asymptotic stable limit cycles at (a)$Re = 68$, (b)$Re = 75$ and (c)$Re = 100$: the black dashed curve from the symmetric steady solution $\boldsymbol{u}_s(\times)$, the red curve from the asymmetric steady solution $\boldsymbol{u}_s^-(\bullet)$, and the blue curve from the asymmetric steady solution $\boldsymbol{u}_s^+(\blacksquare)$.

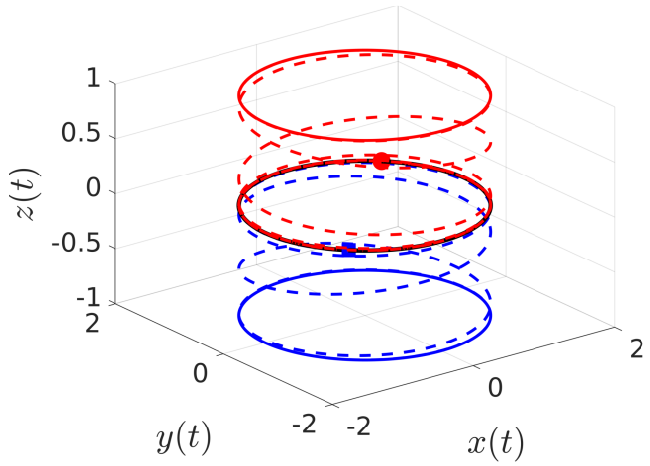


Figure 2.22: Three-dimensional schematic diagram of $\mu = 3$, with the transition from the perturbed initial conditions on the unstable limit cycle $z = 0$ (the black cycle) to the corresponding stable limit cycles $z = \pm\sqrt{\mu - \mu_2}$ (the blue and red cycles): the blue dashed curve from $(-1, -1, -0.0001)(\blacksquare)$, the red dashed curve from $(1, 1, 0.0001)$ $(\bullet)$.

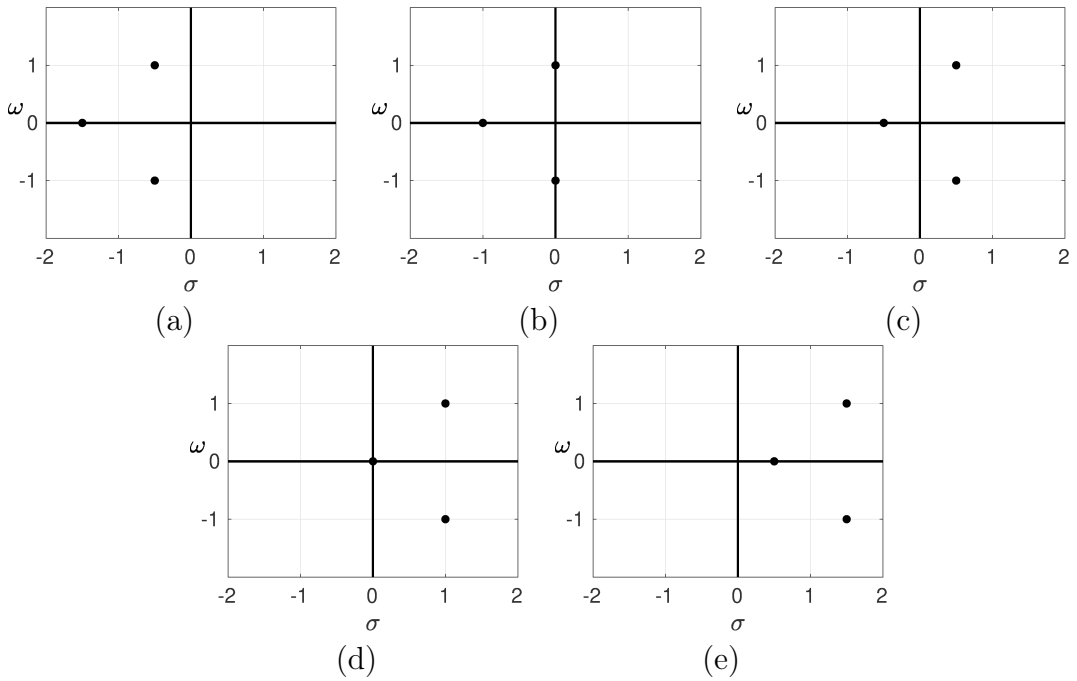


Figure 2.23: Eigenspectrum resulting from the linear stability analysis of the steady solution $\boldsymbol{q}_s^0$ at different μ: (a)$\mu = 0.5$, (b)1.0, (c)1.5, (d)2.0, (b)2.5.

Linear stability analysis

The ODE (2.50) can be writen as:

$$\dot{\boldsymbol{q}} = \boldsymbol{F}(\boldsymbol{q}), \quad \boldsymbol{q} = (x, y, z) \tag{2.51}$$

and we note $\boldsymbol{q}_s$ is the steady state,that is, $\boldsymbol{F}(\boldsymbol{q}_s) = 0$. Consider a small perturbation $\boldsymbol{q'}$ around the steady state $\boldsymbol{q}_s$ by

$$\boldsymbol{q} = \boldsymbol{q}_s + \boldsymbol{q'} \tag{2.52}$$

We derived the linearized evolution equation:

$$\dot{\boldsymbol{q}'} = \boldsymbol{DF}(\boldsymbol{q}_s)\boldsymbol{q'} \tag{2.53}$$

where $\boldsymbol{DF(q_s)}$ is the Jacobian matrix of the considered steady state $\boldsymbol{q}_s$.

The stability of this steady state is determined by the eigenvalues $\sigma + i\omega$ of the Jacobian matrix. The eigenspectrum of the fixed point $\boldsymbol{q}_s^0 = (0, 0, 0)$ is shown in figure 2.23. The growth rate and angular frequency of the pair of conjugated eigenvalues are $\mu - \mu_1$ and ω_0 respectively. The growth rate of the real eigenvalue is $\mu - \mu_2$.

The eigenspectrum of the steady solution $\boldsymbol{q}_s^{\pm} = (0, 0, \sqrt{\mu - \mu_2})$ for $\mu > \mu_2$ is shown in figure 2.24(right). The growth rate is $\mu - \mu_2$ at the steady solution $\boldsymbol{q}_s^0$ and $-2(\mu - \mu_2)$ at the steady solution $\boldsymbol{q}_s^{\pm}$.

Floquet stability analysis

Now, we consider the periodic solution $\boldsymbol{q}_p(t)$ of the system of (2.50), which can be writen as:

$$\boldsymbol{q}_p(t + T) = \boldsymbol{q}_p(t), \quad \text{with} \quad \dot{\boldsymbol{q}}_p(t) = F(\boldsymbol{q}_p(t)) \tag{2.54}$$

Consider a small perturbation $\boldsymbol{q'}$ around the periodic solution by

$$\boldsymbol{q}(t) = \boldsymbol{q}_p(t) + \boldsymbol{q'}(t) \tag{2.55}$$

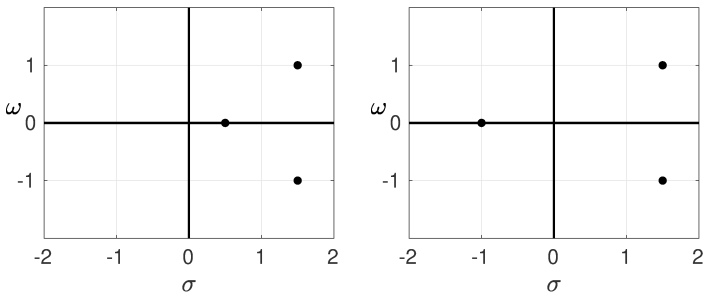


Figure 2.24: Eigenspectrum resulting from the linear stability analysis of the steady solutions $\boldsymbol{q}_s^0$(left) and $\boldsymbol{q}_s^{\pm}$(right), at $\mu = 2.5$

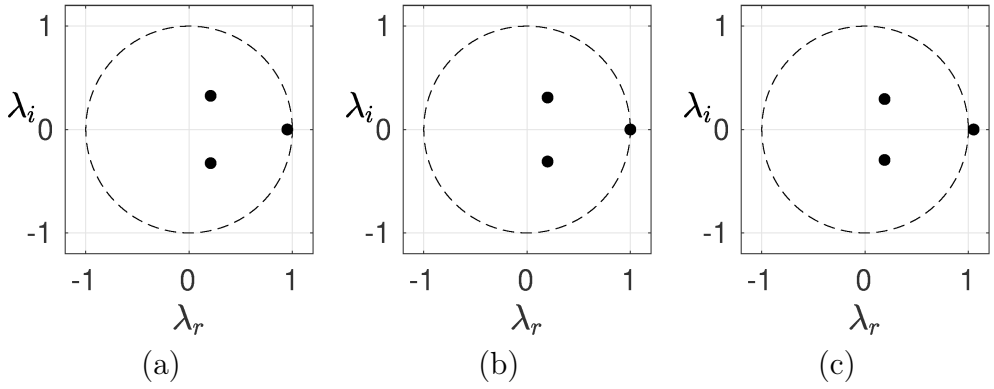


Figure 2.25: Floquet stability analysis of the $z = 0$ periodic solutions at different $\mu =$ (a)1.95, (b)2.0, and (c)2.05

The first variational form reads:

$$\dot{\boldsymbol{q}'}(t) = \boldsymbol{DF}(\boldsymbol{q}_p(t))\boldsymbol{q}'(t) \tag{2.56}$$

where $\boldsymbol{DF}(\boldsymbol{q}_p(t))$ is the Jacobian matrix of the considered periodic solution $\boldsymbol{q}_p(t)$, but now, the linear equation has periodic coefficients.

The monodromy matrix can be written as:

$$M_{\text{mono}} = \exp\left(\int_0^T \boldsymbol{DF}(\boldsymbol{q}_p(t))dt\right) \tag{2.57}$$

or equivalently:

$$M_{\text{mono}} = \exp\left(\frac{T}{2\pi}\int_0^{2\pi} \boldsymbol{DF}(\boldsymbol{q}_p(\theta))d\theta\right). \tag{2.58}$$

The stability of this periodic solution is determined by the multipliers $\lambda = \exp((\sigma_F + i\omega_F)/T)$, being the eigenvalues of the monodromy matrix. As shown in figure 2.25, the periodic solution $\boldsymbol{q}_p^0$: $(\sqrt{\mu - \mu_1}\cos(\theta), \sqrt{\mu - \mu_1}\sin(\theta), 0)$ for $\mu \geq \mu_2$ becomes unstable. A real multiplier λ crosses the unit cycle at $+1$. The Floquet exponent of this real multiplier is $\mu - \mu_2$, equal to the growth rate at the fixed point $\boldsymbol{q}_s^0$.

The multipliers of the other two periodic solutions $\boldsymbol{q}_p^{\pm}$: $(\sqrt{\mu - \mu_1}\cos(\theta), \sqrt{\mu - \mu_1}\sin(\theta), \pm\sqrt{\mu - \mu_2})$ is shown in figure 2.24(right). The Floquet exponent of the leading multiplier is $-2(\mu - \mu_2)$, equal to the growth rate at the fixed point $\boldsymbol{q}_s^{\pm}$.

As a consequence, if the additional degree of freedom z, introduced by the pitchfork bifurcation, do not couple, at the onset of the bifurcation, to the primary degrees of freedom x, y associated with the Hopf bifurcation (see § 2.3.5), then it is easy to understand that both the symmetric steady solution $(x_s, y_s, z_s) \equiv (0, 0, 0)$ and the statistically symmetric periodic solution $(x_c(t + T), y_c(t + T), 0) = (x_c(t), y_c(t), 0)$, T being the period of the limit cycle, will both undergo an instability with

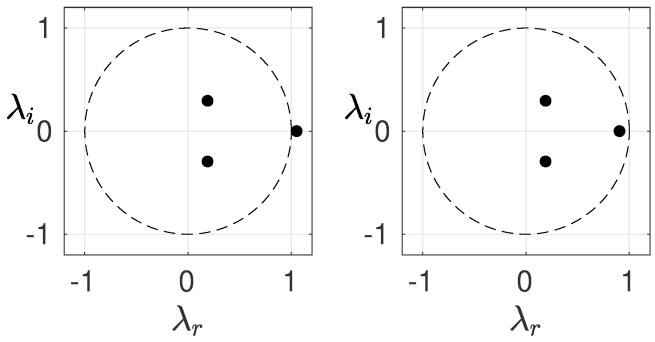


Figure 2.26: Floquet stability analysis of the periodic solutions $\boldsymbol{q}_p^0$(left) and $\boldsymbol{q}_p^{\pm}$(right), at $\mu = 2.05$

respect to the symmetry breaking provoked by the pitchfork bifurcation ($z \neq 0$). This simultaneous instability of both the symmetric fixed point and the statistically symmetric limit cycle looks like an instability of the subspace (x, y) with respect to the transverse direction z, associated with the active degree of freedom introduced by the pitchfork bifurcation.

Chapter 3

Coinciding local bifurcations in the Navier-Stokes equations

Nan Deng[1], Luc R. Pastur[1], Laurette S. Tuckerman[2], Bernd R. Noack[3]

[1] Institute of Mechanical Sciences and Industrial Applications, ENSTA-Paris, Institut Polytechnique de Paris, 828 Bd des Maréchaux, F-91120 Palaiseau, France.

[2] Laboratoire de Physique de Mécanique des Milieux Hétéerogènes (PMMH), CNRS, ESPCI Paris, PSL Research University; Sorbonne Université, Université de Paris, F-75005, Paris, France

[3] School of Mechanical Engineering and Automation, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, People's Republic of China

Published in *EPL (Europhysics Letters)* (Deng *et al.* (2021*b*)).

Generically, a local bifurcation only affects a single solution branch. However, branches that are quite different may nonetheless share certain eigenvectors and eigenvalues, leading to coincident bifurcations. For the fluidic pinball, two supercritical pitchfork bifurcations, of the equilibrium and the periodic solutions, occur at nearly the same Reynolds number. The mechanism of this kind of non-generic coincidence is modelled and explained.

3.1 Introduction

Vortex shedding, symmetry breaking, and self-sustained oscillations are very common in fluid flows (Strykowski & Sreenivasan, 1990; Crawford & Knobloch, 1991; Rowley *et al.*, 2002). In terms of dynamical systems theory, vortex shedding in the wake of an obstacle usually results from a Hopf bifurcation, either supercritical, as in the cylinder wake flow (Dušek *et al.*, 1994; Noack *et al.*, 2003) and in the wake of axisymmetric bodies (Fabre *et al.*, 2008), or subcritical, as in a stalling wing (Dimitriadis & Li, 2009). Symmetry-breaking pitchfork bifurcations may also occur, again either supercritical, as in bluff body wake flows (Grandemange *et al.*, 2012), or subcritical, as in a symmetric channel with an expanded and contracted section (Mizushima & Shiotani, 2001) and in spherical Couette flow (Mamun & Tuckerman, 1995). Many shear flows at high Reynolds numbers feature coherent structures similar to the patterns engendered by instabilities of the base flow at low Reynolds numbers. For instance, the von Kármán streets of vortices found in the wake of cylinders in cross-flow develop at Reynolds numbers as small as $O(10^2)$, where the flow is laminar, while the alley of vortices in the wake of islands in oceans are still observed for Reynolds numbers as large as $O(10^8)$, where the flow is fully turbulent. This is also true for the Kelvin-Helmholtz instability (Smyth & Moum, 2012) and Taylor-Couette flow (Grossmann *et al.*, 2016). New states may arise via secondary instabilities of the base flow as the Reynolds number is increased. Examples are the subcritical pitchfork bifurcation of turbulent vortex shedding (Schewe, 1983) and secondary transitions of the cylinder wake (Zhang *et al.*, 1995; Barkley & Henderson, 1996*b*).

Generically, a local bifurcation only affects one solution branch (as its name implies) and cannot affect the stability of other solution branches. Quite intriguingly however, a recent study of the fluidic pinball configuration has shown that two local bifurcations occur almost simultaneously, i.e. at nearly the same critical Reynolds number (Deng *et al.*, 2020). Both are supercritical pitchfork bifurcations; the first is that of the symmetric steady solution of the flow, while the second is that of the limit cycle associated with the cyclic release of vortices in the wake of the cylinders. Since it is non-generic, this coincidence should not have been observed in the wake flow, and the literature on this coincidence is surprisingly sparse (Shearer, 1981).

In this Letter, we model and explain in detail the non-generic coincidence of these two local pitchfork bifurcations in the fluidic pinball.

3.2 Flow configuration

The fluidic pinball configuration consists of three fixed cylinders of diameter D whose axes are located at the vertices of an equilateral triangle of side $3D/2$ in the (x, y) plane and which are oriented perpendicularly to this plane. The domain is the rectangle $[-6D, 20D] \times [-6D, 6D]$. One vertex of the triangle points upstream and the midpoint of the back two cylinders is chosen as the origin. The upstream flow, of uniform velocity U_∞ at the inlet of the domain, is in the x direction. A variety of flow patterns is found in this configuration as the Reynolds number and spatial arrangement are varied (Bao *et al.*, 2010; Zheng *et al.*, 2016; Chen *et al.*, 2020). We solve the incompressible Navier-Stokes equations

$$\partial_t \mathbf{u} + \nabla \cdot \mathbf{u} \otimes \mathbf{u} = \nu \Delta \mathbf{u} - \nabla p, \tag{3.1}$$

using a second-order finite-element discretization method of the Taylor-Hood type (Taylor & Hood, 1973), on an unstructured grid of 4 225 triangles and 8 633 vertices and implicit third-order time integration (Noack & Morzyński, 2017). The Reynolds number is defined by $Re = U_\infty D/\nu$, where ν is the kinematic viscosity of the fluid. A no-stress condition is applied at the outlet of the domain.

Reflection symmetry in y proves to be an important part of this scenario. For a velocity field $\mathbf{u} = (u, v)$ we define the y-reflection operator R via $R(u, v)(x, y) \equiv (u, -v)(x, -y)$. For a symmetric field, u is even and v is odd in y. The spanwise vorticity $\partial_x v - \partial_y u$ used to represent the flows in our visualizations, manifests y-reflection symmetry by being odd in y. Eigenvectors obtained by linearizing about a reflection-symmetric state are necessarily either symmetric or antisymmetric. Antisymmetric eigenvectors are associated with pitchfork bifurcations that lead to two symmetrically related asymmetric branches if they are real. If they are complex conjugate pairs, they are associated with Hopf bifurcations that lead to limit cycles satisfying the spatio-temporal symmetry $R\mathbf{u}(t) = \mathbf{u}(t + T/2)$. Symmetric real eigenvectors are associated with transcritical bifurcations and complex conjugate eigenvectors are associated with Hopf bifurcations leading to limit cycles that remain symmetric throughout: $R\mathbf{u}(t) = \mathbf{u}(t)$.

The bifurcation diagram is sketched in Fig. 3.1. For low Reynolds number, there is a unique solution $\mathbf{u}_b$, called the base flow. Depicted in Fig. 3.2(a), it is steady and reflection-symmetric with respect to $y = 0$. This solution undergoes a supercritical Hopf bifurcation at $Re_1 \approx 18$ to a pair of y-antisymmetric eigenmodes, leading to cyclic vortex shedding in the wake flow. The resulting T-periodic limit cycle, shown via the instantaneous visualization in Fig. 3.2(b), is the von Kármán vortex street and has the spatio-temporal symmetry $\mathbf{u}_{\mathrm{vk}}(t + T/2) = R\mathbf{u}_{\mathrm{vk}}(t)$ (Barkley, 2006). For this reason, we will call this limit cycle symmetric, despite the fact that the instantaneous flows are not symmetric.

At $Re = Re_2^{\mathrm{SS}} \approx 68$, the base flow undergoes another bifurcation, a supercritical pitchfork that breaks reflection symmetry by deflecting the base-bleeding jet up or down. Two asymmetric branches are generated: $\mathbf{u}^+$ with upwards deflection, shown in Fig. 3.2(c), and its y-reflection $\mathbf{u}^- = R\mathbf{u}^+$, with downwards deflection.

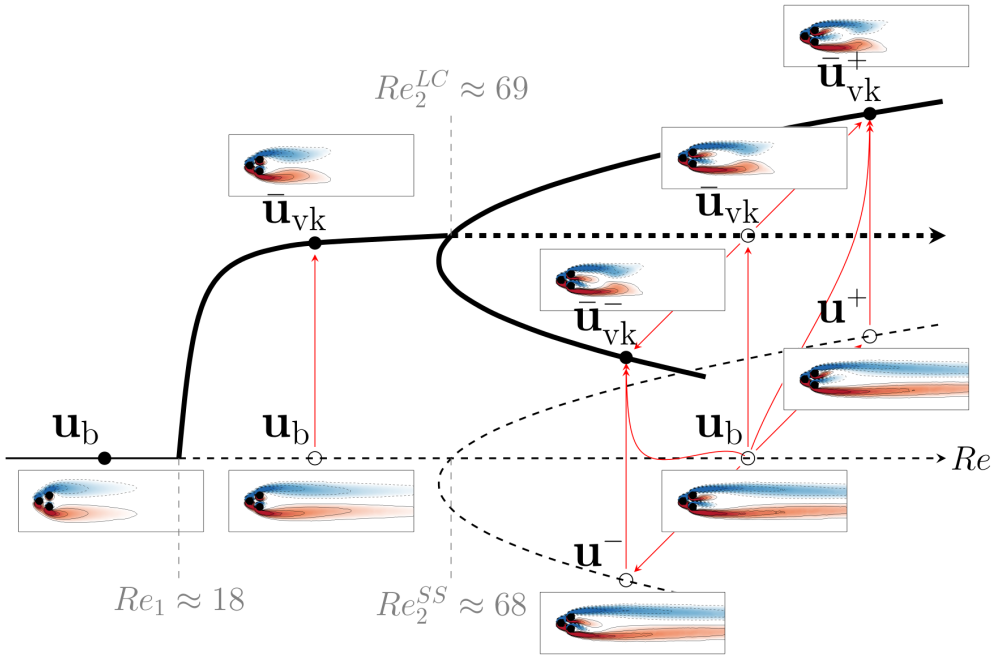


Figure 3.1: Bifurcation diagram for the fluidic pinball. Solid curves indicate stable branches and dashed curves indicate unstable branches. Bold curves indicate branches of periodic solutions and thin curves indicate steady solutions. The symmetric and asymmetric periodic solutions are presented by their time-averages $\bar{\mathbf{u}}_{\mathrm{vk}}$ and $\bar{\mathbf{u}}_{\mathrm{vk}}^{\pm}$. Vorticity fields are color-coded in the range $[-1.5, 1.5]$ from blue to red. The red arrows show the possible transitions between them. The critical Reynolds numbers are detected with linear or Floquet stability analysis of the corresponding solutions.

Like the steady base flow, the periodic von Kármán vortex street $\mathbf{u}_{\mathrm{vk}}$ also undergoes a supercritical pitchfork bifurcation, also involving deflection of the base-bleeding jet, as shown in Fig. 3.2(d). This bifurcation occurs at $Re_2^{\mathrm{LC}} \approx Re_2^{\mathrm{SS}}$ and leads to two limit cycles $\mathbf{u}_{\mathrm{vk}}^{\pm}$ that we call asymmetric because they lack the spatio-temporal symmetry. Before symmetry breaking, the vortex shedding is initiated downstream after the stagnation point of the jet. After symmetry breaking, the vortex shedding is initiated behind one of the back two cylinders, and the stagnation point disappears.

3.3 Stability analysis

In order to better understand this coincidence of $Re_2^{\mathrm{LC}} \approx Re_2^{\mathrm{SS}}$, we conducted linear stability analysis of the base flow. As shown in Fig. 3.3(a), a real eigenvalue has crossed the imaginary

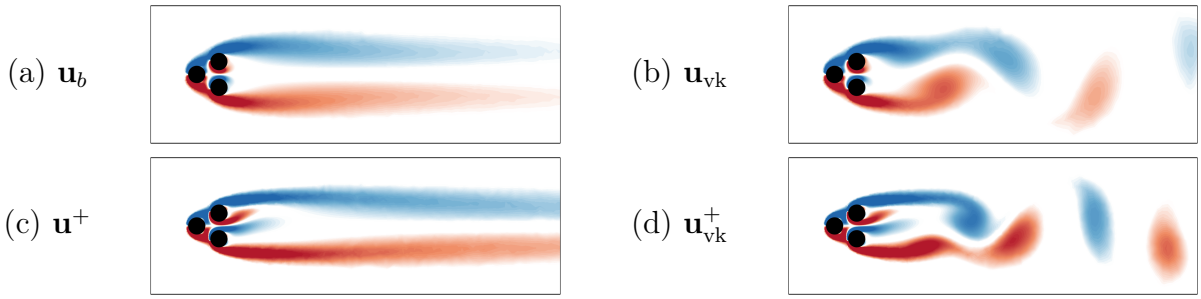


Figure 3.2: Vorticity fields in colour with $[-1.5, 1.5]$ of the base flow $\mathbf{u}_b$, a snapshot of the von Kármán vortex street $\mathbf{u}_{\mathrm{vk}}(t)$, at $Re = 30 > Re_1$, the asymmetric steady solution $\mathbf{u}^{+}$ and a snapshot of the asymmetric von Kármán vortex street $\mathbf{u}_{\mathrm{vk}}^{+}(t)$, , with the base-bleeding jet deflected to the top, at $Re = 80 > Re_2$, color-coded as in Fig. 3.1.

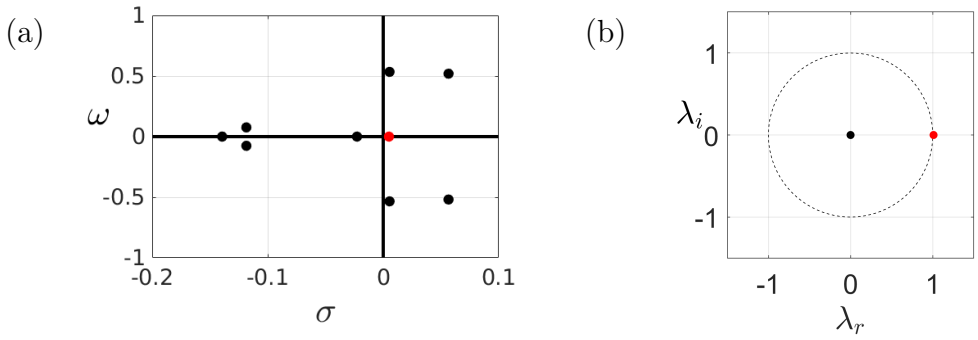


Figure 3.3: (a) Leading eigenvalues of the base flow and (b) leading Floquet multipliers of the von Kármán vortex street, both at $Re = 70$.

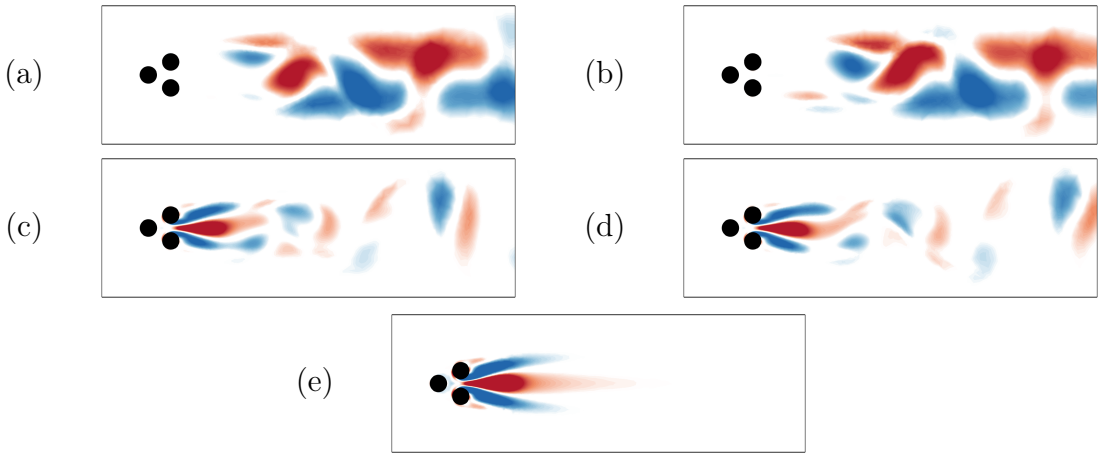


Figure 3.4: (a, b) Deviation of the von Kármán vortex street from the base flow $\mathbf{u}_{\rm vk}(t) - \mathbf{u}_b$ at two instants separated by $T/4$, (c, d) the corresponding Floquet modes, and (e) the eigenmode with real eigenvalue of the base flow, at $Re = 70$. For (a-d), the spatio-temporal symmetry $\mathbf{u}(t+T/2) = R\mathbf{u}(t)$ implies that the next two instants are obtained by y-reflection and sign change (color reversal) of these vorticity fields. In (e), the spanwise vorticity is even in y, corresponding to a flow which is antisymmetric with respect to y-reflection symmetry.

axis at $Re_2^{\rm SS} = 68 \pm 1$. We also performed Floquet analysis of the von Kármán vortex street. Figure 3.3(b) shows that a Floquet multiplier crosses the unit circle at $+1$ and the critical value determined is $Re_2^{\rm LC} = 69 \pm 1$, which is very close to $Re_2^{\rm SS}$.

Figures 3.4(a), (b) show the deviation of the von Kármán vortex street from the base flow at $Re = 70 > Re_2$ at two instants separated by a quarter-period, together with the corresponding Floquet modes in Fig. 3.4(c,d). Figure 3.4(e) is the eigenmode responsible for the pitchfork bifurcation of the base flow. These correspond to the positive real eigenvalue of Fig. 3.3(a) and the Floquet multiplier crossing the unit circle at $\lambda = +1$ of Fig. 3.3(b). The spatial structures of these modes resemble one another, indicating that the two pitchfork bifurcations are closely related and correspond to the same physical mechanism, i.e. the base-bleeding jet which is dominant in the near wake throughout the development of the von Kármán vortex street. The pitchfork bifurcations originate locally within the three cylinders of the fluidic pinball mechanism, while the von Kármán vortex street originates in the wake of the three cylinders taken as a whole and approximated by a single obstacle. Figures 3.4(a), (b) show that the difference between the von Kármàn vortex street and the base flow is nearly zero near the three cylinders, which makes it plausible that both flows undergo the same local instability within the pinball mechanism.

The real eigenvalue of the base flow $\mathbf{u}_b$ and the Floquet exponent of the von Kármán vortex street $\mathbf{u}_{\rm vk}$ are plotted in Fig. 3.5 as functions of the Reynolds number. The fact that the curves

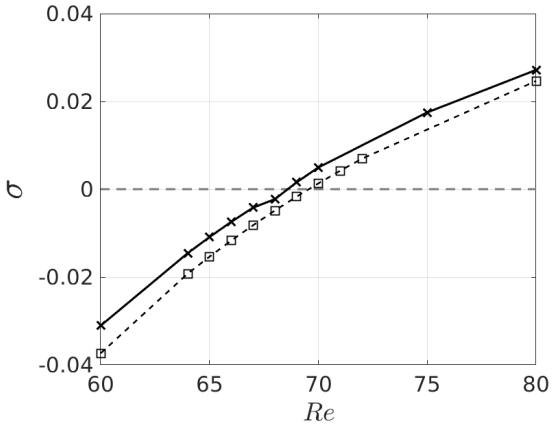


Figure 3.5: Evolution of the real eigenvalue of the base flow $\mathbf{u}_b$ (solid curve with crosses), and the Floquet exponent of the von Kármán vortex street $\mathbf{u}_{\text{vk}}$ (dashed curve with squares) with the Reynolds number.

in Fig. 3.5 remain parallel confirms that the jet's instability mechanism is independent of the downstream flow. Both eigenvalues increase with Re, eventually causing both $\mathbf{u}_{\text{vk}}$ and $\mathbf{u}_b$ to undergo symmetry-breaking pitchfork bifurcations at very close though distinct values of Re that deviate the central jet upwards or downwards.

3.4 A simple model for the coincidence

We now write down a simple model for this phenomenon. Consider a system that undergoes successively a supercritical Hopf bifurcation and, for a higher value of the control parameter μ, a supercritical pitchfork bifurcation. In the fluidic pinball, μ is the Reynolds number Re. The system involves three degrees of freedom, the two that are involved in the Hopf bifurcation, written in polar form as $re^{i\theta}$ and that involved in the pitchfork, z. The generic form of such a system reads:

$$\begin{cases} \dot{r} &= (\mu - \mu_1 - r^2 - \chi_r z^2) r \\ \dot{\theta} &= \omega_0 + r^2 - \chi_i z^2 \\ \dot{z} &= (\mu - \mu_2 - z^2 - \chi_z r^2) z \end{cases} . \tag{3.2}$$

The basic state $r = 0 = z$ undergoes a Hopf bifurcation at $\mu = \mu_1$, leading to a limit cycle with $r = \sqrt{\mu - \mu_1}$, $z = 0$, and a pitchfork bifurcation at $\mu = \mu_2$, leading to asymmetric steady states with $z = \pm\sqrt{\mu - \mu_2}$. Solutions with $r > 0$ and $z \neq 0$, i.e. asymmetric limit cycles, exist for $\mu > \mu_c$ where

$$\mu_c \equiv \mu_2 + \frac{\chi_z}{1 - \chi_z}(\mu_2 - \mu_1). \tag{3.3}$$

If $\chi_z = 0$, then $\mu_c = \mu_2$, i.e. the pitchforks of the limit cycle and of the basic state take place at the same critical value, and if $|\chi_z| \ll 1$, they occur almost simultaneously in μ. The non-generic property of coincident bifurcations is a direct consequence of $\chi_z \approx 0$.

The eigenvalues of system (3.2) are easily calculated from its Jacobian. The basic state has eigenvalues $\mu - \mu_1$ and $\mu - \mu_2$ in the r and z directions, respectively. The Floquet exponents of the symmetric limit cycle are the eigenvalues of system (3.2) with θ removed, that are $-2(\mu - \mu_1)$ and $\mu - \mu_2 - \chi_z(\mu - \mu_1)$. The second Floquet exponent will closely track eigenvalue $\mu - \mu_2$ of the basic state, as in Fig. 3.5, if $|\chi_z(\mu - \mu_1)|$ remains small compared to μ_2.

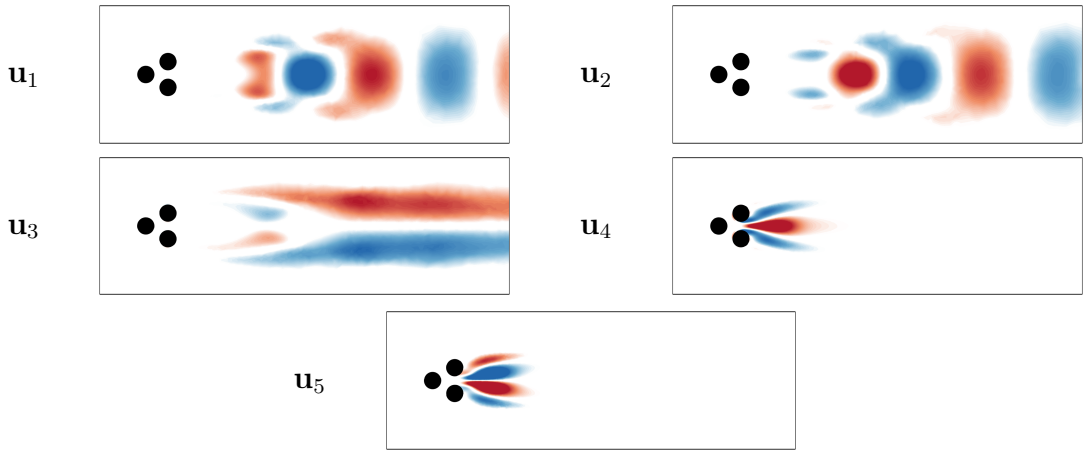


Figure 3.6: Elementary modes of the fluidic pinball at $Re = 70$. See text for details.

3.5 Least-order Galerkin model for the coincidence

We can derive system (3.2) from the corresponding least-order Galerkin model (Deng *et al.*, 2020) slaving the Reynolds-stress-related modes. As stated there, the dynamics for $Re > Re_2$ requires at least a 5th order Galerkin expansion. The ansatz for the velocity field reads:

$$\mathbf{u}(\mathbf{x}, t) = \mathbf{u}_b(\mathbf{x}) + \sum_{i=1}^{5} a_i(t)\mathbf{u}_i(\mathbf{x}), \tag{3.4}$$

where $\mathbf{u}_b(\mathbf{x})$ is the symmetric base flow of the Navier-Stokes equations. The elementary modes $\mathbf{u}_i(\mathbf{x})$ are shown in Fig. 3.6. Details on the identification of the individual modes can be found in (Deng *et al.*, 2020).

The first two modes $\mathbf{u}_{1,2}$, associated with the vortex shedding, are taken to be the two leading modes from Proper Orthogonal Decomposition (POD) in the permanent dynamical flow regime (Taira *et al.*, 2017). The shift mode $\mathbf{u}_3$ is the difference between the mean flow $\bar{\mathbf{u}}_{\mathrm{vk}}$ of the periodic regime and the base flow $\mathbf{u}_b$ (Noack *et al.*, 2003). Mode $\mathbf{u}_4$ is given by $\mathbf{u}_4 \propto \mathbf{u}^+ - \mathbf{u}^-$. Analogously to $\mathbf{u}_3$, mode $\mathbf{u}_5$ is the difference between the average $(\mathbf{u}^+ + \mathbf{u}^-)/2$ of the asymmetric steady solutions and the base flow $\mathbf{u}_b$.

The degrees of freedom a_1, a_2, a_4, are active, while a_3 and a_5 are slaved to a_1, a_2 and a_4. In the neighborhood of the second bifurcation threshold ($Re \approx Re_2$), the dynamical system reads:

$$\begin{aligned}
da_1/dt &= \sigma a_1 - \omega a_2, \qquad \sigma = \sigma_1 - \beta a_3 + \xi_r a_5 && \text{(3.5a)}\\
da_2/dt &= \sigma a_2 + \omega a_1, \qquad \omega = \omega_1 + \gamma a_3 + \xi_i a_5 && \text{(3.5b)}\\
da_3/dt &= \sigma_3 a_3 + \beta_3 \left(a_1^2 + a_2^2\right) && \text{(3.5c)}\\
da_4/dt &= \sigma_4 a_4 - \beta_4 a_4 a_5 + \xi_z a_4 a_3 && \text{(3.5d)}\\
da_5/dt &= \sigma_5 a_5 + \beta_5 a_4^2 && \text{(3.5e)}
\end{aligned}$$

which is equivalent to the system (3.2) if $\sigma_3 \ll 0$ and $\sigma_5 \ll 0$. Indeed, if these conditions are satisfied, the slaved modes $a_3 \propto a_1^2 + a_2^2$ and $a_5 \propto a_4^2$ provide the cubic nonlinearities of Eq. (3.2) with $a_1 + ia_2 = re^{i\theta}$ and $z \equiv a_4$. The coefficients of system (3.5) can be directly computed by a Galerkin projection of the Navier-Stokes equations on the bifurcation modes $\mathbf{u}_{1\ldots5}$. At $Re = 70$, close to Re_2, $\chi_z = 0.0168$, confirming that it is small.

3.6 Discussion

To check the robustness of the coincidence, we changed the distance between the cylinders from $L/D = 1.5$ to 1.4 and 1.6. The critical Reynolds numbers for the Hopf and pitchfork bifurcations,

L/D	Re_1	Re_2^{SS}	Re_2^{LC}	χ_z
1.4	18	79	80	0.0099
1.5	18	68	69	0.0168
1.6	20	63	63	0.0186

Table 3.1: Critical Reynolds numbers for the two bifurcations and associated χ_z when the cylinder gap is varied. For each gap, χ_z is determined from Galerkin projection close to the threshold at $Re = Re_2^{LC} + 1$.

together with the associated χ_z, are recorded in Table 3.1. In all three cases, the two pitchfork bifurcations still closely coincide, confirming the robustness of this phenomenon with respect to a change in one of the control parameters (here the cylinder gap).

For these cases, the three cylinders work as a single bluff body. The base-bleeding jet focuses on the near flow, which breaks the symmetry but does not separate the bluff body wake. When L/D is sufficiently large, however, the three cylinders no longer form one compact body. Two vortex streets appear and compete with each other. A new mechanism then changes the flow dynamics completely with varying gap distance; we refer interested readers to (Chen *et al.*, 2020).

We argue that coincident bifurcations like that observed in the fluidic pinball at $Re = Re_2$ should be observed in other flow configurations, and more generally in other nonlinear partial differential equations, when competing and independent instability mechanisms are present.

Acknowledgements : This work is supported by a public grant overseen by the French National Research Agency (ANR) by grant 'FlowCon' (ANR-17-ASTR-0022). N. D. has been supported by the China Scholarships Council (No.201808070123).

Chapter 4

Galerkin force model for transient and post-transient dynamics of the fluidic pinball

Nan Deng[1,2] , Bernd R. Noack[3,4], Marek Morzyński[5] and Luc R. Pastur[1]

[1] Institute of Mechanical Sciences and Industrial Applications, ENSTA-Paris, Institut Polytechnique de Paris, 828 Bd des Maréchaux, F-91120 Palaiseau, France

[2] LIMSI, CNRS, Université Paris-Saclay, Bât 507, rue du Belvédère, Campus Universitaire, F-91403 Orsay, France

[3] Center for Turbulence Control, Harbin Institute of Technology, Shenzhen, Room 312, Building C, University Town, Xili, Shenzhen 518058, People's Republic of China

[4] Institut für Strömungsmechanik und Technische Akustik (ISTA), Technische Universität Berlin, Müller-Breslau-Straße 8, D-10623 Berlin, Germany

[5] Chair of Virtual Engineering, Poznań University of Technology, Jana Pawla II 24, PL 60-965 Poznań, Poland

Published in *Journal of Fluid Mechanics* (Deng *et al.* (2021*a*)).

We propose an aerodynamic force model associated with a Galerkin model for the unforced fluidic pinball, the two-dimensional flow around three equal cylinders with one radius distance to each other. The starting point is a Galerkin model of a bluff-body flow. The force on this body is derived as a constant-linear-quadratic function of the mode amplitudes from first principles following the pioneering work of Noca (1997); Noca *et al.* (1999) and Liang & Dong (2014). The force model is simplified for the mean-field model of the unforced fluidic pinball (Deng *et al.*, 2020) using symmetry properties and sparse calibration. The model is successfully applied to transient and post-transient dynamics in different Reynolds number regimes: the periodic vortex shedding after the Hopf-bifurcation and the asymmetric vortex shedding after the pitchfork bifurcation comprising six different Navier-Stokes solutions. We foresee many applications of the Galerkin force model for other bluff bodies and flow control.

Contents

4.1 Introduction

The literature on aerodynamic forces on bodies associated with POD or any other Galerkin model is surprisingly sparse. On the one hand, force computations are at the heart of engineering fluid mechanics. On the other hand, systematic investigations and interpretations of the aerodynamic force in the Galerkin framework are mostly missing. Considering POD as a linear decomposition of the flow field realizations, Brunton & Rowley (2009) observed that

> "While POD modes and the low order model allow for accurate reconstruction of the flow field and preserve Lagrangian coherent structures, it is not clear that this model is directly useful for reconstructing body forces quickly and accurately, since lift and drag forces depend nonlinearly on the flow field, meaning that contributions from different POD modes cannot be added independently."

The pioneering early work of Noca (1997); Noca *et al.* (1999) reveals that the instantaneous fluid dynamic forces on the body can be expressed with only the velocity fields and their derivatives. Liang & Dong (2014) applied it to the velocity based POD modes, and derived a force expression in terms of the force of each POD mode and the force from the interaction between the POD modes. The Galerkin force model proposed in this work reveals that any force component is a constant-linear-quadratic function of the mode amplitudes.

The starting point of our investigation is a working Galerkin model based on a low-dimensional modal expansion of an incompressible viscous fluid flow around a stationary body. Intriguingly, mean-field theory (Stuart, 1958, 1971) was the first foundation of many Galerkin models, building on weakly nonlinear generalizations of stability analyses. Mean-field theory delivered the first derivation of the Landau model (see, e.g., Landau & Lifshitz, 1987) for super- and subcritical Hopf bifurcations. The Landau model is experimentally supported for the onset of vortex shedding behind the cylinder wake (Schumm *et al.*, 1994; Zielinska & Wesfreid, 1995). Generalizations explain

the cross-talk between different frequencies over the base flow (Luchtenburg *et al.*, 2009; Shaabani-Ardali *et al.*, 2020), special cases of 'quasi-laminar' interactions foreshadowed by Reynolds & Hussain (1972).

A few decades later, the pioneering wall turbulence POD model by Aubry *et al.* (1988) allows employing snapshot data far a low-dimensional encapsulation of the Navier-Stokes dynamics. Since then, numerous empirical reduced-order models have been proposed (Taira *et al.*, 2017; Kunisch & Volkwein, 2002; Bergmann *et al.*, 2009; Ilak & Rowley, 2006; Rempfer, 2000; Rowley *et al.*, 2004). Control-oriented versions have been developed by Rowley & Dawson (2017); Barbagallo *et al.* (2009); Bagheri *et al.* (2009*a*); Hinze & Volkwein (2005); Gerhard *et al.* (2003).

A working Galerkin model can predict the flow and thus the force. Theories for aerodynamic forces have a rich history documented in virtually every fluid mechanics textbook (see, e.g., Panton, 1984). There are several force formulae for different cases. Potential flow theory for finite bodies can only explain the force due to accelerations of the body and predicts vanishing drag (d'Alambert paradox). The Zhukovsky formula derives the lift for the potential flow around streamlined cylinders, while the drag computation is still excluded by the d'Alambert paradox. The lifting line theory by Prandtl (1921) extends Zukovsky's formula for finite wings and adds a drag estimate from the created trailing edge vortices. Kirchhoff (1869) laid the first practical foundation for bluff-body drag by allowing for a separation with infinitely thin shear layer. Until today, the drag and lift forces of a body are inferred from the downstream velocity profile (Schlichting & Gersten, 2016). These are arguably the most common force theories.

In the Galerkin modelling literature, unsteady forces have been formulated as functions of mode coefficients, like in Bergmann & Cordier (2008) and in Luchtenburg *et al.* (2009). The force formulae are generally calibrated from the reconstructed flow field. Noca *et al.* (1999) offered an expression of the unsteady forces on an immersed body in an incompressible flow, which only requires the knowledge of the velocity field and its time derivative. Based on this idea, Liang & Dong (2014) presented a velocity POD mode force survey method to measure the forces from POD modes on a flat plate. It has shown that the force superposition of each mode of a full POD model can accurately predict the instantaneous forces, and the leading six POD modes are enough to predict the drag force with 5% error.

In this study, we focus on the unforced "fluidic pinball", the flow around three equidistantly placed cylinders in crossflow (Bansal & Yarusevych, 2017). Following Chen *et al.* (2020), the gap distance between the cylinders is chosen one radius and the triangle formed by the centers of three cylinders points upstream. This distance allows for an interesting 'flip-flopping' dynamics. The advantage of the fluidic pinball is that already the two-dimensional laminar flow exhibits a surprisingly rich dynamics which has recently been accurately modeled (Deng *et al.*, 2020). As the Reynolds number increases, the flow behaviour changes from a globally stable fixed dynamics to a periodic symmetric vortex shedding after a Hopf bifurcation, to asymmetric vortex shedding after a subsequent pitchfork bifurcation, followed by quasi-periodic and chaotic behavior. Intriguingly, the post-pitchfork regime with three unstable steady solutions as well as two stable asymmetric limit cycles and one unstable symmetric limit cycle is adequately described by a single five-dimensional Galerkin model. Apparently, the force model for multiple transients of this pitchfork regime is already a challenge.

In the present work, we propose a Galerkin force model for the transient dynamics of the unforced fluidic pinball at different Reynolds numbers. We derive the unsteady forces from the Navier-Stokes equations yielding a constant-linear-quadratic expression of the mode amplitudes of the Galerkin expansion. The consistent form with Liang & Dong (2014) strengthens the theoretical basis of the force expression. Any known symmetric property of the modes is usually considered in the relative modal analysis (Rigas *et al.*, 2014; Podvin *et al.*, 2020), particularly advised for symmetry-breaking instabilities of flows around a symmetric configuration (Fabre *et al.*, 2008; Borońska & Tuckerman, 2010). Since the fluidic pinball exhibits a mirror-symmetry, we further investigate the force expression under the Z_2-symmetry. The drag and lift contributions must come

from the specific subsets of the constant-linear-quadratic polynomial functions, which is consistent with the drag- and lift-producing modes identified in Liang & Dong (2015).

The manuscript is organized as follows. § 4.2 derives the aerodynamic force from a Galerkin model. § 4.3 describes the simulation and Galerkin model of the fluidic pinball. In § 4.4, the force model for the transition of a simple Hopf bifurcation and for the transition of a simple pitchfork bifurcation are discussed. Next, the force model with the elementary modes of two successive bifurcations for the multi-attractor case is investigated in § 4.5, together with a optimization based on the correction of mean-field distortion. We summarize the results and outline future directions of research in § 4.6.

4.2 Galerkin force model

In this section, the derivation of a Galerkin force model is described and discussed. Based on the framework of a Galerkin expansion (§ 4.2.1), the drag and lift forces are expressed as constant-linear-quadratic functions of the mode amplitudes in § 4.2.2. Alternatively, the forces can consistently be derived from the momentum balance as elaborated in Appendix 4.A. The force model can be further simplified under symmetry considerations in § 4.2.3.

4.2.1 The Galerkin framework

The fluid flow satisfies the non-dimensionalized incompressible Navier-Stokes equations

$$\partial_t \boldsymbol{u} + \nabla \cdot \boldsymbol{u} \otimes \boldsymbol{u} = \nu \triangle \boldsymbol{u} - \nabla p, \tag{4.1}$$

where p and $\boldsymbol{u}$ are respectively the pressure and velocity flow fields, $\nu = 1/Re$, with the Reynolds number Re. Here, ∂_t, ∇, $\triangle$, $\otimes$ and $\cdot$ respectively denote the partial derivative in time, the Nabla and Laplace operator as well as the outer and inner tensor product. All the variables have been non-dimensionalized, with the cylinder diameter D, the oncoming velocity U, the time scale D/U, and the density ρ of the fluid.

It is assumed that there exists at least one steady solution $(\boldsymbol{u}_s, p_s)$, satisfying the steady Navier-Stokes equations

$$\nabla \cdot \boldsymbol{u}_s \otimes \boldsymbol{u}_s = \nu \triangle \boldsymbol{u}_s - \nabla p_s. \tag{4.2}$$

For the Galerkin framework, the space of the square-integrable vector fields $\mathcal{L}^2(\Omega)$ is introduced in the observation domain Ω. The associated inner product for two velocity fields $\boldsymbol{u}(\boldsymbol{x})$ and $\boldsymbol{v}(\boldsymbol{x})$ reads

$$(\boldsymbol{u}, \boldsymbol{v})_\Omega := \int_\Omega d\boldsymbol{x}\, \boldsymbol{u}(\boldsymbol{x}) \cdot \boldsymbol{v}(\boldsymbol{x}). \tag{4.3}$$

The velocity field is decomposed in a basic mode $\boldsymbol{u}_0$ and a fluctuating contribution. The basic mode may be the steady Navier-Stokes solution $\boldsymbol{u}_s$ or the time-averaged flow $\overline{\boldsymbol{u}}$. The fluctuation is represented by a Galerkin approximation of N orthonormal space-dependent modes $\boldsymbol{u}_i(\boldsymbol{x})$, $i = 1, \ldots, N$, with time-dependent amplitudes $a_i(t)$:

$$\boldsymbol{u}(\boldsymbol{x}, t) \;=\; \sum_{i=0}^{N} a_i(t) \boldsymbol{u}_i(\boldsymbol{x}), \tag{4.4}$$

where the basic mode $\boldsymbol{u}_0$ is associated with $a_0 \equiv 1$ following Rempfer & Fasel (1994*a*). The orthonormality condition reads $(\boldsymbol{u}_i, \boldsymbol{u}_j)_\Omega = \delta_{ij}, \quad i, j \in \{1, \ldots, N\}$.

The Galerkin expansion (4.4) satisfies the incompressibility condition and the boundary conditions by construction. The evolution equation for the mode amplitudes a_i is derived by a Galerkin

projection of the Navier-Stokes equation (4.1) onto the modes $\boldsymbol{u}_i$:

$$\frac{d}{dt}a_i = \nu \sum_{j=0}^{N} l_{ij}^{\nu} a_j + \sum_{j,k=0}^{N} q_{ijk}^{c} a_j a_k + \sum_{j,k=0}^{N} q_{ijk}^{p} a_j a_k, \tag{4.5}$$

with the coefficients $l_{ij}^{\nu} = (\boldsymbol{u}_i, \triangle \boldsymbol{u}_j)_\Omega$, $q_{ijk}^{c} = (\boldsymbol{u}_i, \nabla \cdot \boldsymbol{u}_j \otimes \boldsymbol{u}_k)_\Omega$ and $q_{ijk}^{p} = (\boldsymbol{u}_i, -\nabla p_{jk})_\Omega$ for the viscous, convective and pressure terms in the Navier-Stokes equations (4.1), respectively. Details are provided by Noack *et al.* (2005). Thus, a linear-quadratic Galerkin system (Fletcher, 1984) can be derived,

$$\frac{d}{dt}a_i = \nu \sum_{j=0}^{N} l_{ij}^{\nu}\, a_j + \sum_{j,k=0}^{N} \left[q_{ijk}^{c} + q_{ijk}^{p}\right]\, a_j\, a_k. \tag{4.6}$$

4.2.2 Drag and lift forces on a body

Let Γ be the boundary of the body in the flow domain Ω and $\boldsymbol{n}$ the unit normal pointing outward the surface element dS. The α-component F_α^ν ($\alpha = x, y, z$) of the viscous force vector $\boldsymbol{F}^\nu$ on the boundary is expressed by

$$F_\alpha^\nu = \boldsymbol{F}^\nu \cdot \boldsymbol{e}_\alpha = 2\nu \oint_\Gamma \sum_{\beta=x,y,z} S_{\alpha,\beta}\, n_\beta\, dS\,, \tag{4.7}$$

where $\boldsymbol{e}_\alpha$ is the unit vector in α-direction and $S_{\alpha,\beta} = \left(\partial_\alpha u_\beta + \partial_\beta u_\alpha\right)/2$ the strain rate tensor with indices $\alpha, \beta = x, y, z$.

Similarly, the α-component of the global pressure force, exerted on an immersed body, is defined as

$$F_\alpha^p = \boldsymbol{F}^p \cdot \boldsymbol{e}_\alpha = -\oint_\Gamma dS\, n_\alpha p. \tag{4.8}$$

Without external forces, the viscous and pressure forces in Ω counter-balance the inertial terms provided by the left-hand side of Eq. (4.1). The drag force is defined as the projection on $\boldsymbol{e}_x$ of the pressure and viscous forces exerted on the body

$$F_D(t) = F_x^p(t) + F_x^\nu(t). \tag{4.9}$$

The lift force is similarly defined as the projection on $\boldsymbol{e}_y$ of the resulting pressure and viscous forces exerted on the body

$$F_L(t) = F_y^p(t) + F_y^\nu(t). \tag{4.10}$$

The drag and lift coefficients read

$$C_D(t) = \frac{2F_D(t)}{\rho\, U^2}, \qquad C_L(t) = \frac{2F_L(t)}{\rho\, U^2}. \tag{4.11}$$

Employing the Galerkin approximation (4.4), the viscous force (4.7) can be re-written as

$$F_\alpha^\nu = \sum_{j=0}^{N} q_{\alpha;j}^{\nu} a_j, \tag{4.12}$$

where $q_{\alpha;j}^{\nu}$ can easily be derived from (4.7) with the corresponding $S_{\alpha,\beta}$ of the velocity mode $\boldsymbol{u}_j$, with the form

$$q_{\alpha;j}^{\nu} = 2\nu \oint_\Gamma \sum_{\beta=x,y,z} S_{\alpha,\beta}(\boldsymbol{u}_j)\, n_\beta\, dS. \tag{4.13}$$

Note that the contribution of the viscous force is linear with respect to the mode amplitudes a_j.

Similarly, from the pressure Poisson equation

$$\nabla^2 p = \nabla \cdot (-\nabla \cdot \boldsymbol{u} \otimes \boldsymbol{u}) = - \sum_{\alpha=x,y,z} \sum_{\beta=x,y,z} \partial_\alpha u_\beta \partial_\beta u_\alpha, \tag{4.14}$$

the expression of the pressure field is derived as

$$p(\boldsymbol{x}, t) = \sum_{j,k=0}^{N} p_{jk}(\boldsymbol{x})\, a_j(t)\, a_k(t), \tag{4.15}$$

with

$$\nabla^2 p_{jk} = \nabla \cdot (-\nabla \cdot \boldsymbol{u}_j \otimes \boldsymbol{u}_k) = - \sum_{\alpha=x,y,z} \sum_{\beta=x,y,z} \partial_\alpha u_\beta(\boldsymbol{u}_j)\, \partial_\beta u_\alpha(\boldsymbol{u}_k). \tag{4.16}$$

The boundary conditions for partial pressures p_{jk} are discussed by Noack *et al.* (2005). Integrating (4.8) with (4.15) shows that the pressure force is a quadratic polynomial of the a_j's

$$F^p_\alpha = \sum_{j,k=0}^{N} q^p_{\alpha;jk} a_j a_k, \quad \text{where} \quad q^p_{\alpha;jk} = - \oint_\Gamma dS\, n_\alpha p_{jk}. \tag{4.17}$$

Taking the steady solution as the basic mode $\boldsymbol{u}_0 = \boldsymbol{u}_s$ with $a_0 \equiv 1$ implies that $\boldsymbol{a}$ with $a_i = \delta_{0i}$ is a fixed point of Eq. (4.6) and the total force can be expressed as a constant-linear-quadratic expression in terms of the mode coefficients

$$F_\alpha = F^\nu_\alpha + F^p_\alpha = c_\alpha + \sum_{j=1}^{N} l_{\alpha;j} a_j + \sum_{j,k=1}^{N} q_{\alpha;jk} a_j a_k, \tag{4.18}$$

where

$$c_\alpha = q^\nu_{\alpha;0} + q^p_{\alpha;00}, \qquad l_{\alpha;j} = q^\nu_{\alpha;j} + q^p_{\alpha;j0} + q^p_{\alpha;0j}, \qquad q_{\alpha;jk} = q^p_{\alpha;jk}. \tag{4.19}$$

The force expression in Eq. (4.18) can be alternatively derived from the residual of the Navier-Stokes equations in the flow domain Ω, as demonstrated in Appendix 4.A.

With constant ρ and U, the drag and lift coefficients in (4.11) can be rewritten in the form

$$C_D = c_x + \sum_{j=1}^{N} l_{x;j} a_j + \sum_{j,k=1}^{N} q_{x;jk} a_j a_k, \tag{4.20a}$$

$$C_L = c_y + \sum_{j=1}^{N} l_{y;j} a_j + \sum_{j,k=1}^{N} q_{y;jk} a_j a_k. \tag{4.20b}$$

A crucial step relies on the choice of the $\boldsymbol{u}_i$ modes for the decomposition of Eq. (4.4). These could be the POD modes, as usually considered in fluid flows. However, a better choice could be to decompose the flow field on a basis of modes that are becoming active when the system is undergoing a bifurcation. This choice of the so-called *bifurcation modes* will be investigated in § 4.3.3.

4.2.3 The Navier-Stokes equations under the Z_2-symmetry

When the fluid flow configuration exhibits a mirror-symmetry, the Navier-Stokes equations (4.1) possess at least one symmetric steady solution $(\boldsymbol{u}_s, p_s)$, satisfying Eq. (4.2). The Z_2-symmetry of the velocity and pressure fields, with respect to the (x, z)-plane defined by $y = 0$, implies

$$\begin{aligned} u^s(x,-y,z) &= u^s(x,y,z), \quad v^s(x,-y,z) = -v^s(x,y,z), \\ p^s(x,-y,z) &= p^s(x,y,z), \end{aligned} \tag{4.21a}$$

$$\begin{aligned} u^a(x,-y,z) &= -u^a(x,y,z), \quad v^a(x,-y,z) = v^a(x,y,z), \\ p^a(x,-y,z) &= -p^a(x,y,z), \end{aligned} \tag{4.21b}$$

where the symmetric components $(u^s, v^s, p^s) \in \mathcal{U}^s$ and the antisymmetric components $(u^a, v^a, p^a) \in \mathcal{U}^a$, $\mathcal{U}^s$ and $\mathcal{U}^a$ being respectively the symmetric and antisymmetric subspaces of the system. Other steady solutions can exist, which break the symmetry of the system. We will consider the symmetric steady solution $(\boldsymbol{u}_s, p_s)$ as the reference point of Eq. (4.1) in the Reynolds decomposition of the flow field as Eq. (4.25).

The dynamics under consideration can include transient and post-transient regimes. Here, we introduce the T-averaged flow fields $\bar{\boldsymbol{u}}_T(\boldsymbol{x}, t)$ as

$$\bar{\boldsymbol{u}}_T(\boldsymbol{x}, t) = \frac{1}{T} \int_{t-T/2}^{t+T/2} \boldsymbol{u}(\boldsymbol{x}, \tau)\, d\tau, \tag{4.22}$$

where T is a time-scale to be chosen. When the flow field is oscillating in time, an appropriate choice for T is the period of the local oscillation. The mean flow field is further defined as

$$\bar{\boldsymbol{u}}(\boldsymbol{x}) = \lim_{T\to\infty} \bar{\boldsymbol{u}}_T(\boldsymbol{x}, t) \tag{4.23}$$

and only focuses on the post-transient limit.

When two mirror-conjugated attractors co-exist, it is convenient to introduce the ensemble-averaged flow field $\bar{\boldsymbol{u}}^\bullet_T(\boldsymbol{x}, t)$ as

$$\bar{\boldsymbol{u}}^\bullet_T(\boldsymbol{x}, t) = \frac{1}{2}(\bar{\boldsymbol{u}}^+_T(\boldsymbol{x}, t) + \bar{\boldsymbol{u}}^-_T(\boldsymbol{x}, t)). \tag{4.24}$$

where $\bar{\boldsymbol{u}}^\pm_T(\boldsymbol{x}, t)$ are the T-averaged flow field on the way to each individual attractor. This definition could be readily extended to more than two conjugated attractors. As an ensemble average on mirror-conjugated attracting sets, the ensemble-averaged flow field $\bar{\boldsymbol{u}}^\bullet_T(\boldsymbol{x}, t)$ belongs to the symmetric subspace $\mathcal{U}^s$.

At this point, it is most convenient to introduce the Reynolds decomposition of the flow field, in the form

$$\boldsymbol{u}(\boldsymbol{x}, t) = \bar{\boldsymbol{u}}^\bullet_T(\boldsymbol{x}, t) + \boldsymbol{u}'(\boldsymbol{x}, t) = \boldsymbol{u}_s(\boldsymbol{x}) + \boldsymbol{u}_\Delta(\boldsymbol{x}, t) + \boldsymbol{u}'(\boldsymbol{x}, t). \tag{4.25}$$

where the mean-field deformation $\boldsymbol{u}_\Delta(\boldsymbol{x}, t)$ accounts for the distortion of the flow field from the symmetric steady solution $\boldsymbol{u}_s(\boldsymbol{x})$ to the ensemble-averaged flow field $\bar{\boldsymbol{u}}^\bullet_T(\boldsymbol{x}, t)$ as

$$\boldsymbol{u}_\Delta(\boldsymbol{x}, t) = \bar{\boldsymbol{u}}^\bullet_T(\boldsymbol{x}, t) - \boldsymbol{u}_s(\boldsymbol{x}). \tag{4.26}$$

The fluctuation flow field $\boldsymbol{u}'(\boldsymbol{x}, t)$ has a vanishing time average, meaning that $\boldsymbol{u}(\boldsymbol{x}, t)$ is centered on $\bar{\boldsymbol{u}}^\bullet_T(\boldsymbol{x}, t)$. By construction, $\bar{\boldsymbol{u}}^\bullet_T(\boldsymbol{x}, t), \boldsymbol{u}_\Delta(\boldsymbol{x}, t), \boldsymbol{u}_s(\boldsymbol{x})$ belongs to the symmetric subspace $\mathcal{U}^s$ and $\boldsymbol{u}'(\boldsymbol{x}, t)$ to the anti-symmetric subspace $\mathcal{U}^a$. Thus, a symmetry-based decomposition of Eq. (4.1) results into a symmetric and an anti-symmetric part, yielding

$$\begin{aligned} \partial_t \boldsymbol{u}_\Delta + \nabla \cdot [\boldsymbol{u}_s \otimes \boldsymbol{u}_\Delta + \boldsymbol{u}_\Delta \otimes \boldsymbol{u}_s + \boldsymbol{u}_\Delta \otimes \boldsymbol{u}_\Delta + \boldsymbol{u}' \otimes \boldsymbol{u}'] &= \nu \triangle \boldsymbol{u}_\Delta - \nabla p_\Delta, && \text{(4.27a)} \\ \partial_t \boldsymbol{u}' + \nabla \cdot [\bar{\boldsymbol{u}}^\bullet_T \otimes \boldsymbol{u}' + \boldsymbol{u}' \otimes \bar{\boldsymbol{u}}^\bullet_T] &= \nu \triangle \boldsymbol{u}' - \nabla p'. && \text{(4.27b)} \end{aligned}$$

Integrating (4.27a) on the spatial domain Ω, both the left and right hand sides yield a time-evolving force vector aligned on $\boldsymbol{e}_y$, while integrating (4.27b) yields a time-evolving force vector aligned on $\boldsymbol{e}_x$. The former is the resulting lift force applying to the boundaries of the fluid domain, while the latter is the drag force. Thus, the Z_2-symmetry applied to equations (4.20a) and (4.20b) yields

$$C_D = C_D^\circ + \sum_{j=1}^{N} \underbrace{[l_{x;j} a_j]}_{\in \mathcal{U}^s} + \sum_{j,k=1}^{N} \underbrace{[q_{x;jk} a_j a_k]}_{\in \mathcal{U}^s}, \tag{4.28a}$$

$$C_L = \sum_{j=1}^{N} \underbrace{[l_{y;j} a_j]}_{\in \mathcal{U}^a} + \sum_{j,k=1}^{N} \underbrace{[q_{y;jk} a_j a_k]}_{\in \mathcal{U}^a}, \tag{4.28b}$$

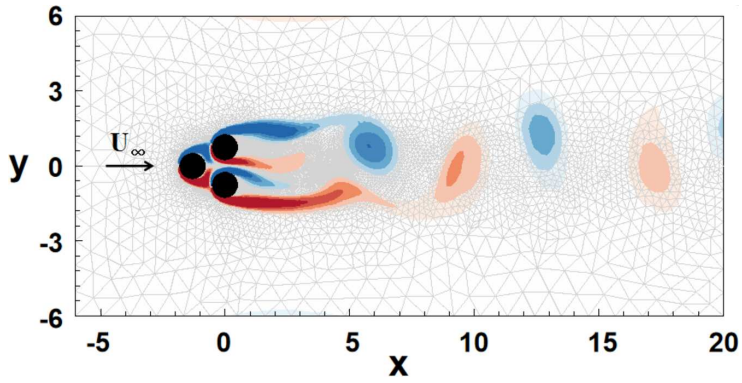


Figure 4.1: Configuration of the fluidic pinball and dimensions of the simulated domain. A typical field of vorticity is represented in color with $[-1.5, 1.5]$. The upstream velocity is denoted U_∞.

where C_D° is the drag coefficient of the symmetric steady solution.

The vanishing terms in (4.28) can be easily derived from the definition of $q^\nu_{\alpha;j}$ and $q^p_{\alpha;jk}$ in § 4.A as:

$$l_{x;j} = q^\nu_{x;j} + q^p_{x;0j} + q^p_{x;j0} = 0, \qquad \boldsymbol{u}_j \in \mathcal{U}^a, \tag{4.29a}$$
$$l_{y;j} = q^\nu_{y;j} + q^p_{y;0j} + q^p_{y;j0} = 0, \qquad \boldsymbol{u}_j \in \mathcal{U}^s, \tag{4.29b}$$
$$q_{x;jk} = q^p_{x;jk} = 0, \qquad \boldsymbol{u}_j \otimes \boldsymbol{u}_k \in \mathcal{U}^a, \tag{4.29c}$$
$$q_{y;jk} = q^p_{y;jk} = 0, \qquad \boldsymbol{u}_j \otimes \boldsymbol{u}_k \in \mathcal{U}^s. \tag{4.29d}$$

As a result, the drag contribution must come from the symmetric subsets of the constant-linear-quadratic polynomial functions, and from the antisymmetric subsets for the lift contribution.

4.3 Galerkin model of the fluidic pinball

The force model derived in § 4.2 is applied to a configuration of three equidistantly placed cylinders in a cross-flow, known as the "fluidic pinball" configuration (Noack & Morzyński, 2017). The flow configuration and the direct Navier-Stokes solver are described in § 4.3.1. As the Reynolds number is increased, the flow undergoes two subsequent supercritical Hopf and pitchfork bifurcations. The corresponding force dynamics at different Reynolds numbers are reported in § 4.3.2. The *bifurcation modes*, newly introduced by Deng *et al.* (2020), are defined in § 4.3.3. They provide the orthogonal basis for the Galerkin projection.

4.3.1 The fluidic pinball

The geometric configuration, shown in figure 4.1, consists of three fixed cylinders of unit diameter D mounted on the vertices of an equilateral triangle of side length $3D/2$ in the (x, y) plane. The flow domain is bounded with a $[-6, +20] \times [-6, +6]$ box. The upstream flow, of uniform velocity U_∞ at the input of the domain, is transverse to the cylinder axis and aligned with the symmetry axis of the cylinder cluster. All quantities will be non-dimensionalized with cylinder diameter D, the velocity U_∞, and the unit fluid density ρ. Considering the symmetry of this configuration, a Cartesian coordinate system will be used in the following discussion, with its origin in the middle of the rightmost two cylinders. In this study, no external force will be applied to these three cylinders. A no-slip condition is applied on the cylinders and the velocity in the far wake is assumed to be U_∞. Here, the Reynolds number is defined as $Re = U_\infty D/\nu$, where ν is the kinematic viscosity of the fluid. A no-stress condition is applied at the output of the domain.

The resolution of the Navier-Stokes equations (4.1) is based on a second-order finite-element discretization method of the Taylor-Hood type (Taylor & Hood, 1973), on an unstructured grid

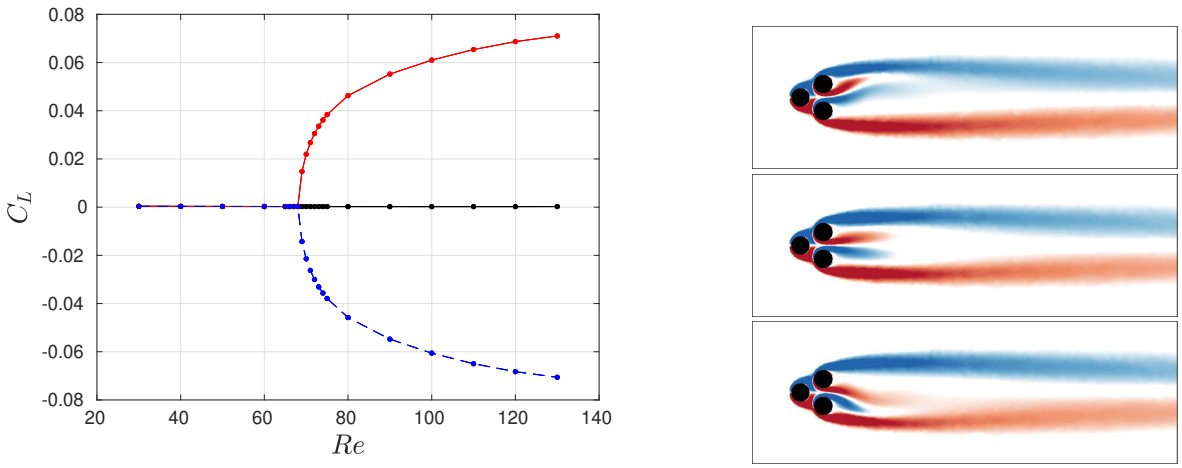


Figure 4.2: Lift coefficients at different Reynolds numbers (a) of the symmetric steady solutions $\boldsymbol{u}_s$(black curve),the asymmetric steady solutions $\boldsymbol{u}_s^-$(blue curve), the asymmetric steady solutions $\boldsymbol{u}_s^+$(red curve), exemplified with the vorticity field of $\boldsymbol{u}_s^+$, $\boldsymbol{u}_s$, $\boldsymbol{u}_s^-$ at $Re = 100$ from top to bottom (b).

of 4 225 triangles and 8 633 vertices, and an implicit integration of the third-order in time. The instantaneous flow field is calculated with a Newton-Raphson iteration until the residual reaches a tiny tolerance prescribed. This approach is also used to calculate the steady solution, which is derived from the steady Navier-Stokes equations (4.2). The Direct Navier-Stokes solver used herein has been validated in Noack *et al.* (2003) and Deng *et al.* (2020), with a detailed technical report (Noack & Morzyński, 2017). The grid used for the simulations was shown to provide a consistent flow dynamics, compared to a refined grid, see Deng *et al.* (2020).

4.3.2 Flow features and the corresponding force dynamics

Different from Deng *et al.* (2020), where the viscous contribution to the forces has been ignored, the lift C_L and drag C_D coefficients are here calculated from the resulting force $\boldsymbol{F}$ of pressure and viscous components exerted on the three cylinders.

The flow characteristics depend on the Reynolds number Re. Following the literature on clusters of cylinders (Chen *et al.*, 2020), the characteristic length scale is chosen to be the cylinder diameter D and not the transverse width $5D/2$ of the configuration. This width loses its dynamic significance for large distances considered in other studies.

For Reynolds numbers $Re < Re_{\mathrm{H}} \approx 18$, the symmetric steady solution $\boldsymbol{u}_s(\boldsymbol{x})$ was found to be stable and is the only attractor of the system. A supercritical Hopf bifurcation occurs at $Re = Re_{\mathrm{H}}$, associated with the cyclic release of counter-rotating vortices in the wake of the three cylinders from the shear-layers that delimit the configuration, forming a von Kármán street of vortices. The corresponding Reynolds number based on the transverse width of the fluidic pinball is 45, i.e., is well-aligned with typical onsets of vortex shedding behind bluff bodies. For the critical value $Re = Re_{\mathrm{PF}} \approx 68$, the system undergoes a supercritical pitchfork bifurcation. As a result, two additional (unstable) steady solutions occur, namely $\boldsymbol{u}_s^+(\boldsymbol{x})$ and $\boldsymbol{u}_s^-(\boldsymbol{x})$, which break the reflectional symmetry of the configuration, as shown with the lift coefficients of the steady solutions in figure 4.2. The mean-field inherits the asymmetry of the steady solutions, with the jet between the two downstream cylinders being deflected upward or downward. As reported in Deng *et al.* (2020), at $Re = Re_{\mathrm{PF}}$, the statistically symmetric limit cycle, associated with the statistically symmetric vortex shedding, becomes unstable with respect to two mirror-conjugated statistically asymmetric limit cycles, associated with statistically asymmetric von Kármán streets of vortices.

Figure 4.3 shows the time evolution of the lift and drag coefficients at $Re = 80$, when the initial condition is either the symmetric steady solution $\boldsymbol{u}_s$ (figure 4.3(a)) or the asymmetric steady

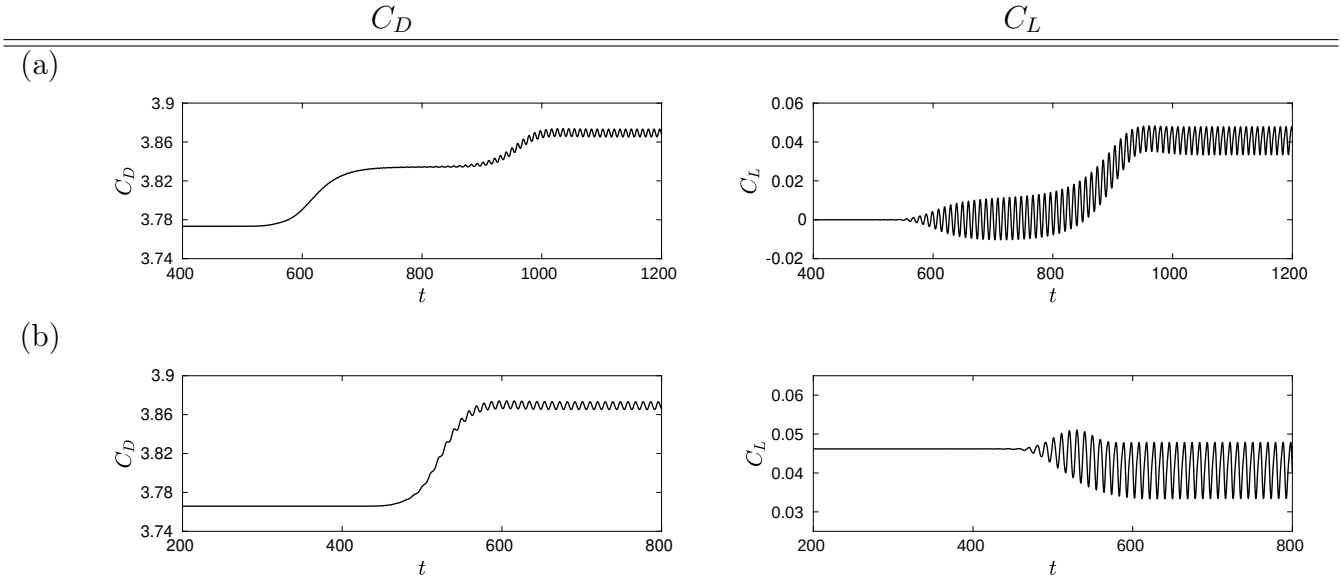


Figure 4.3: Time evolution of the drag (left) and lift (right) coefficients, starting (a) from the symmetric steady solution $\boldsymbol{u}_s$, (b) from the asymmetric steady solution $\boldsymbol{u}_s^+$, at $Re = 80$.

solution $\boldsymbol{u}_s^+$ (figure 4.3(b)). In both cases, the asymptotic regime is the same. However, when starting from the symmetric steady solution in figure 4.3(a), a long-living plateau of the drag coefficient is reached around time $t \approx 775$, which corresponds to the transient exploration of the unstable limit cycle, centered on the symmetric T-averaged flow field $\bar{\boldsymbol{u}}_{98}(\boldsymbol{x}, 775)$. Note that during the transient dynamics from the steady solution to the unstable limit cycle, the drag coefficient is monotonically increasing, before reaching the transient plateau. The drag coefficient is further increasing when leaving the unstable limit cycle towards the asymptotically stable limit cycle, the latter being centered on the asymmetric mean flow field $\bar{\boldsymbol{u}}^+$.

Figure 4.4 shows another representation of the transient dynamics for $Re = 30$, 80 and 100, starting from different initial conditions in the plane $(C_L, \Delta C_D)$, where $\Delta C_D = C_D - C_D^\circ$, C_D° being the drag associated with the symmetric steady solution at the Reynolds number under consideration. The black cross ($\times$) stands for the symmetric steady solution $\boldsymbol{u}_s$ while the asymmetric $\boldsymbol{u}_s^+$ and $\boldsymbol{u}_s^-$ steady solutions are respectively represented by a red circle and a blue square, when they exist, at $Re = 80$ and 100. As it can be observed in this figure, to the difference of what happens at $Re = 80$, the transient dynamics from the symmetric steady solution at $Re = 100$ first reaches one of the two asymmetric steady solutions, before evolving toward the stable attracting limit cycle.

4.3.3 The bifurcation modes of the fluidic pinball

In the case of two subsequent supercritical Hopf and pitchfork bifurcations, Deng *et al.* (2020) have shown that the reduced-order model must comprise 5 modes:

$$\boldsymbol{u}(\boldsymbol{x}, t) = \boldsymbol{u}_s(\boldsymbol{x}) + \sum_{j=1}^{5} a_j(t)\boldsymbol{u}_j(\boldsymbol{x}). \tag{4.30}$$

Hence, in the decomposition of Eq. (4.4), the number of modes is restricted to $N = 5$. For dynamic interpretability, the basic mode $\boldsymbol{u}_0(\boldsymbol{x})$ is chosen to be symmetric steady solution $\boldsymbol{u}_s(\boldsymbol{x})$. The first three modes $\boldsymbol{u}_{1,2,3}(\boldsymbol{x})$ are associated with the Hopf bifurcation, the last two modes $\boldsymbol{u}_{4,5}(\boldsymbol{x})$ with the pitchfork bifurcation. We will refer to these modes as the irreducible *bifurcation modes* of the system. Modes $\boldsymbol{u}_3(\boldsymbol{x})$ and $\boldsymbol{u}_5(\boldsymbol{x})$ are symmetric. The instability-related modes $\boldsymbol{u}_{1,2}(\boldsymbol{x})$ and $\boldsymbol{u}_4(\boldsymbol{x})$ are anti-symmetric. Modes $\boldsymbol{u}_{1,2}(\boldsymbol{x})$ span the subspace associated with the limit cycle of the Hopf

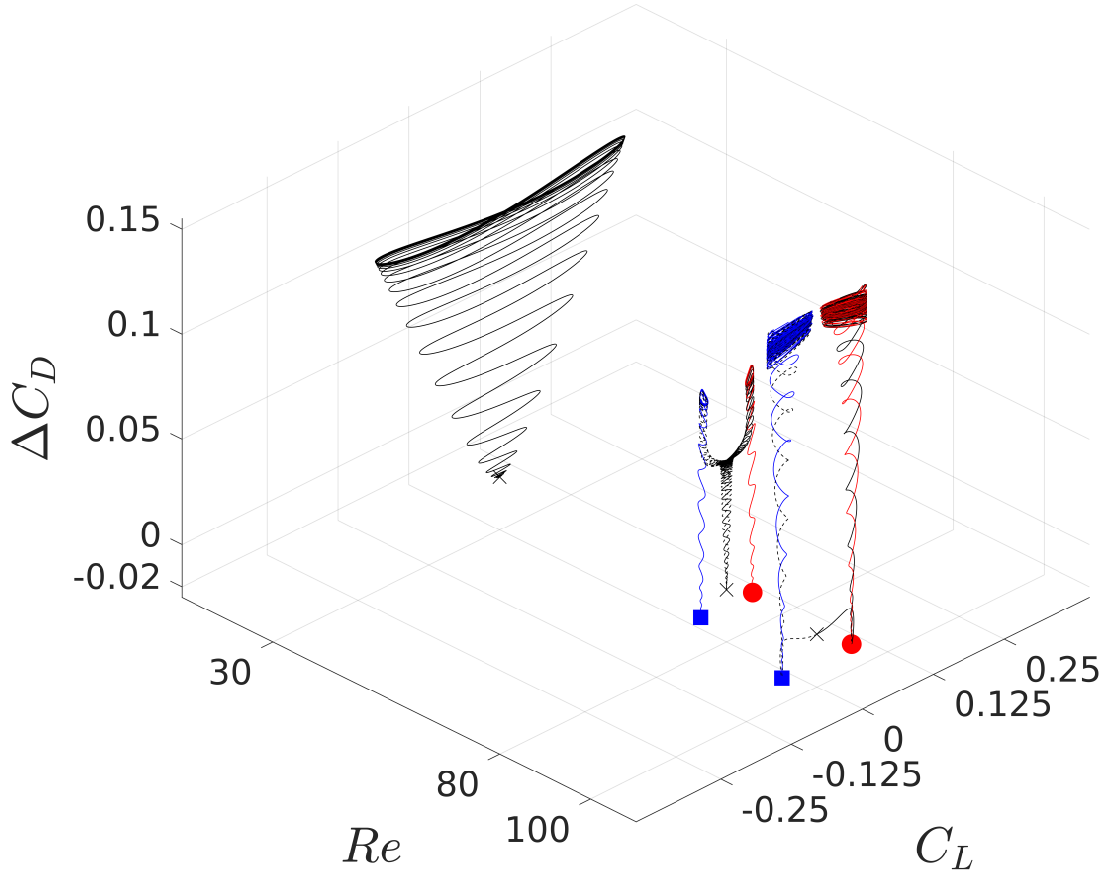


Figure 4.4: Trajectories in the $(C_L, \Delta C_D)$ plane, for Reynolds numbers $Re = 30$, 80 and 100, starting, for the black trajectories, close to the symmetric steady solution $\boldsymbol{u}_s$ ($\times$), for the red trajectories close to the asymmetric steady solution $\boldsymbol{u}_s^+$ ($\bullet$), and for the blue trajectories close to the asymmetric steady solution $\boldsymbol{u}_s^-$ ($\blacksquare$). $\Delta C_D = C_D - C_D^\circ$, where C_D° is the drag coefficient of the symmetric steady solution at the corresponding Reynolds number.

bifurcation, while $\boldsymbol{u}_4(\boldsymbol{x})$ accounts for the symmetry breaking of the pitchfork bifurcation. In Deng *et al.* (2020), modes $\boldsymbol{u}_{1,2}(\boldsymbol{x})$ are provided by the first two dominant POD modes, while mode $\boldsymbol{u}_4(\boldsymbol{x})$ is defined as

$$\boldsymbol{u}_4(\boldsymbol{x}) \propto \boldsymbol{u}_s^+(\boldsymbol{x}) - \boldsymbol{u}_s^-(\boldsymbol{x}), \tag{4.31}$$

where $\boldsymbol{u}_s^\pm(\boldsymbol{x})$ are the two additional (asymmetric) steady solutions arising from the supercritical pitchfork bifurcation. Mode $\boldsymbol{u}_3(\boldsymbol{x})$ is slaved to $\boldsymbol{u}_{1,2}(\boldsymbol{x})$ while $\boldsymbol{u}_5(\boldsymbol{x})$ is slaved to $\boldsymbol{u}_4(\boldsymbol{x})$. The mode $\boldsymbol{u}_3(\boldsymbol{x})$ is usually defined as the shift mode from $\boldsymbol{u}_s(\boldsymbol{x})$ to the asymptotic mean flow field, $\boldsymbol{u}_3(\boldsymbol{x}) \propto \bar{\boldsymbol{u}}(\boldsymbol{x}) - \boldsymbol{u}_s(\boldsymbol{x})$, before being ortho-normalized to $\boldsymbol{u}_1(\boldsymbol{x})$ and $\boldsymbol{u}_2(\boldsymbol{x})$. Here, $\bar{\boldsymbol{u}}(\boldsymbol{x})$ will be restricted to the *symmetric* mean flow field, associated with the statistically symmetric limit cycle, whether this limit cycle is stable or unstable. Similarly to $\boldsymbol{u}_4(\boldsymbol{x})$, mode $\boldsymbol{u}_5(\boldsymbol{x})$ is defined as

$$\boldsymbol{u}_5(\boldsymbol{x}) \propto (\boldsymbol{u}_s^+(\boldsymbol{x}) + \boldsymbol{u}_s^-(\boldsymbol{x})) - 2\boldsymbol{u}_s(\boldsymbol{x}), \tag{4.32}$$

These two modes, together with modes $\boldsymbol{u}_{1,2,3}$, are shown is figure 4.5 after orthonormalization by a Gram-Schmidt procedure, and the corresponding time-dependent amplitudes $a_i(t)$, $i = 1, \ldots, 5$, in the full-flow dynamics are shown in figure 4.6 when starting from either the symmetric steady solution (figure 4.6a) or the asymmetric steady solution (figure 4.6b).

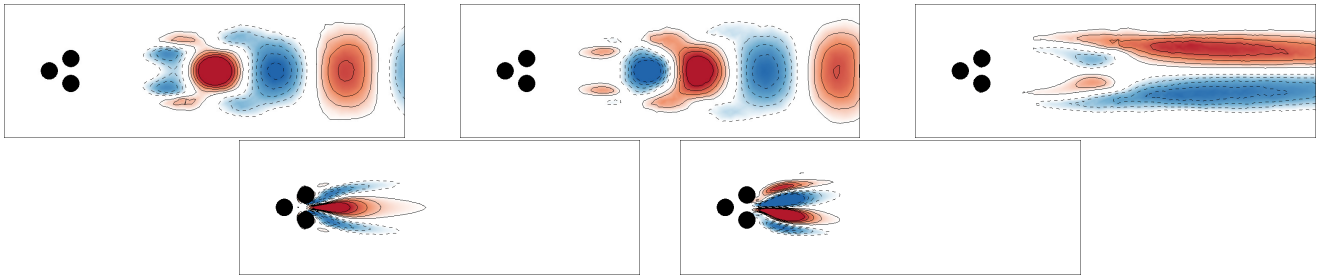


Figure 4.5: Spatial structures of the modes $\boldsymbol{u}_1(\boldsymbol{x})$, $\boldsymbol{u}_2(\boldsymbol{x})$, $\boldsymbol{u}_3(\boldsymbol{x})$ (top), $\boldsymbol{u}_4(\boldsymbol{x})$, $\boldsymbol{u}_5(\boldsymbol{x})$ (bottom), of the velocity field associated with the five elementary degrees of freedom $\{a_1(t) - a_5(t)\}$, at $Re = 80$.

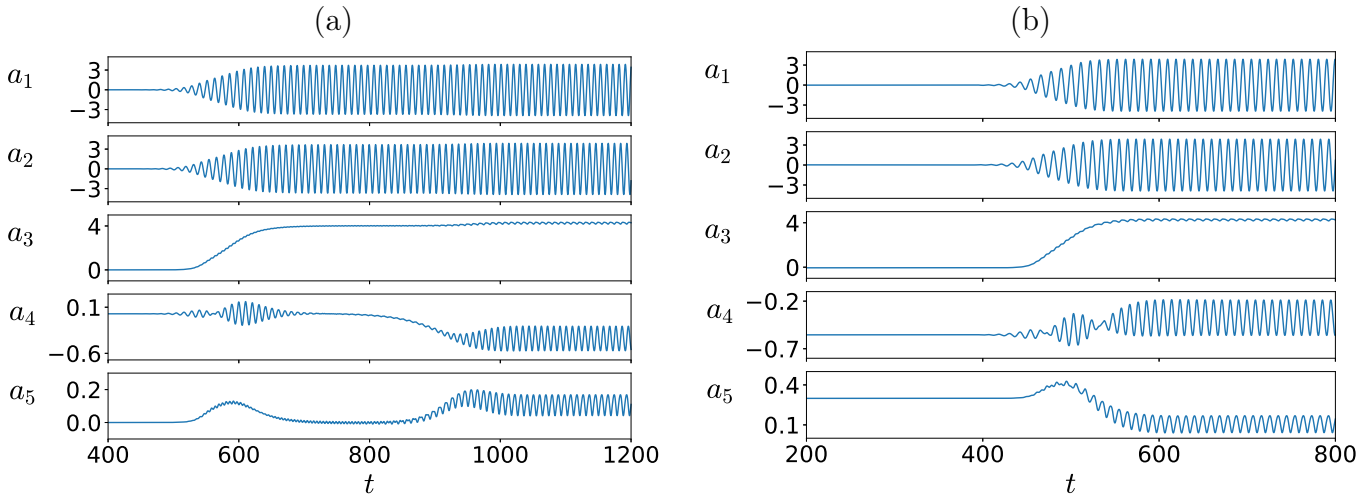


Figure 4.6: Mode amplitudes $a_i(t)$, $i = 1, \ldots, 5$ in the full-flow dynamics starting (a) from the symmetric steady solution $\boldsymbol{u}_s$, (b) from the asymmetric steady solution $\boldsymbol{u}_s^+$, at $Re = 80$.

4.4 Galerkin force model associated with the supercritical Hopf and pitchfork bifurcation

As already mentioned, the fluidic pinball undergoes a supercritical Hopf bifurcation at $Re = Re_{\mathrm{HP}}$ and a subsequent supercritical pitchfork bifurcation at $Re = Re_{\mathrm{PF}} > Re_{\mathrm{HP}}$. The Galerkin force models are derived for the supercritical Hopf bifurcation in § 4.4.1 and for the supercritical pitchfork bifurcation in § 4.4.2.

4.4.1 Force model associated with the supercritical Hopf bifurcation

The symmetric steady solution $\boldsymbol{u}_s \in \mathcal{U}^s$ is stable at low Reynolds numbers. At $Re \geq Re_{\mathrm{HP}}$, it undergoes a supercritical Hopf bifurcation. The resulting Galerkin expansion reads

$$\boldsymbol{u}(\boldsymbol{x},t) = \boldsymbol{u}_s(\boldsymbol{x}) + \underbrace{a_1(t)\,\boldsymbol{u}_1(\boldsymbol{x}) + a_2(t)\,\boldsymbol{u}_2(\boldsymbol{x})}_{\boldsymbol{u}'} + \underbrace{a_3(t)\,\boldsymbol{u}_3(\boldsymbol{x})}_{\boldsymbol{u}_\Delta}, \tag{4.33}$$

and the corresponding mean-field Galerkin system

$$da_1/dt = \sigma a_1 - \omega a_2, \tag{4.34a}$$
$$da_2/dt = \sigma a_2 + \omega a_1, \tag{4.34b}$$
$$da_3/dt = \sigma_3 a_3 + \beta_3\left(a_1^2 + a_2^2\right), \tag{4.34c}$$

with $\sigma = \sigma_1 - \beta a_3$ and $\omega = \omega_1 + \gamma a_3$, where σ_1 and ω_1 are the initial growth rate and frequency depending on the Reynolds number. For a direct supercritical Hopf bifurcation, $\sigma_1, \omega_1, \beta > 0$, $\sigma_3 < 0$ and $\beta_3 > 0$. We refer to Deng *et al.* (2020) for details.

Introducing (4.33) in equations (4.7) and (4.8), the total force can be written as (4.18) with $N = 3$ degrees of freedom. From symmetry considerations, as $\boldsymbol{u}_{1,2} \in \mathcal{U}^a$ and $\boldsymbol{u}_{0,3} \in \mathcal{U}^s$, the coefficients $l_{x;1}$, $l_{x;2}$, $q_{x;13}$, $q_{x;23}$, $l_{y;0}$, $l_{y;3}$, $q_{y;11}$, $q_{y;12}$, $q_{y;22}$, $q_{y;33}$ are vanishing. Finally, the drag formulae (4.28) simplify to

$$C_D = C_D^\circ + l_{x;3}\,a_3 + q_{x;11}\,a_1^2 + q_{x;12}\,a_1a_2 + q_{x;22}\,a_2^2 + q_{x;33}\,a_3^2, \tag{4.35a}$$
$$C_L = l_{y;1}\,a_1 + l_{y;2}\,a_2 + q_{y;13}\,a_1a_3 + q_{y;23}\,a_2a_3. \tag{4.35b}$$

Here again, C_D° is the drag coefficient associated with the symmetric steady solution The unknown parameters in the force model can be identified by a least-squares approach, according to the known force dynamics and the relevant mode amplitudes. However, for the mean-field Galerkin system (4.34), the slaving relation between the degree of freedom a_3 to the oscillating degrees of freedom a_1, a_2 imposes an additional sparsity in the force model. We employ the SINDy (Sparse Identification of Nonlinear Dynamics) algorithm (Brunton *et al.*, 2016*a*) to arrive at simpler and more interpretable models. A $L1$-regularization can be introduced in the LASSO (least absolute shrinkage and selection operator) regression process. Another option in the SINDy algorithm is the sequential thresholded least squares regression, which iteratively applies the least squares regression and eliminates terms with weight smaller than a given threshold. Both regression algorithms benefit from simplicity, only requiring one sparsity parameter λ. The optimal λ balances the accuracy and complexity of the identified model. To evaluate the performance of the identified model, the complexity is presented with the number of non-zero coefficients and the accuracy by the coefficient of determination,denoted as the r^2 *score* (Draper & Smith, 1998). A detailed review of this sparsity parameter can be found in Loiseau & Brunton (2018). A recent extension of the SINDy algorithm with physical constraints of energy-preserving quadratic nonlinearities successfully identifies the sparse model, benefiting from the Galerkin projection of the Navier-Stokes equations (Loiseau *et al.*, 2018*a*).

The LASSO algorithm is applied to a scenario starting with the unstable symmetric steady solution at $Re = 30$. The training data used for the sparse regression is provided by the force

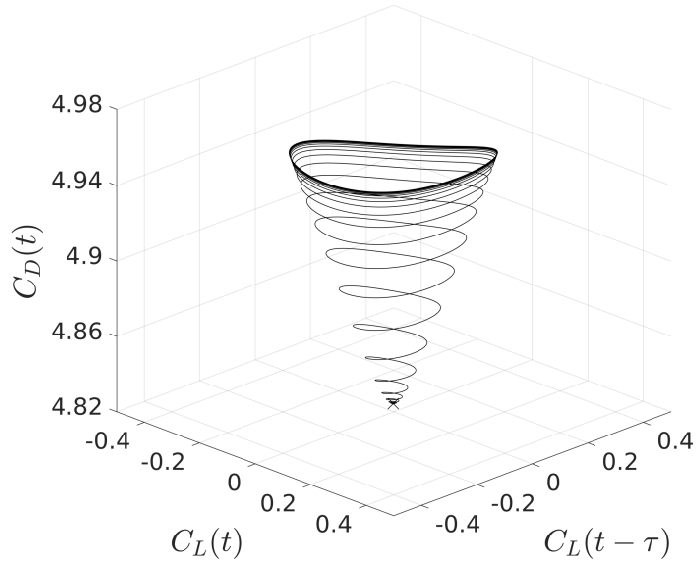


Figure 4.7: Transient dynamics from the unstable symmetric steady solution $\boldsymbol{u}_s$ ($\times$) to the asymptotic limit cycle (statistically symmetric vortex shedding), at $Re = 30$, in the time-delayed embedding space of the lift C_L and drag C_D coefficients, with $\tau = 2$.

coefficients and the mode amplitudes from the DNS starting with the symmetric steady solution to the final asymptotic regime. The resulting transient dynamics and the asymptotically attracting limit cycle are shown in the three-dimensional space of the time-delayed coordinates of C_L and C_D in figure 4.7.

The possible over-fitting terms, such as the slaving relation between a_3 and a_1^2, a_2^2, can be suppressed with a larger $L1$-penalty parameter for the LASSO algorithm. The choice of the $L1$-penalty parameter drives the sparsity of the identified model. A too small $L1$ will lead to a complex model with few eliminated terms; on the contrary, a too-large $L1$ can jeopardize accuracy. Both cases weaken the robustness of the identified model,and the same is observed for the sequential thresholded least squares regression. The influence of the sparsity parameter λ and the comparison of these two regression methods are presented in Appendix 4.B.

Gradually increasing the $L1$-penalty from 0 to nearly 1, the terms a_1a_2, a_3, a_2^2, a_1^2 are eliminated subsequently in the drag model, while a_3^2 is always retained. The sparsity parameter λ, here the $L1$-penalty, is chosen as the largest value without any known over-fitting term. Hence, according to the order of elimination, a_3 is the over-fitting term in the drag model due to the slaving relation between a_3 and a_1^2, a_2^2. The details of this choice can be found in Appendix 4.B. Finally, the identified force model reads

$$C_D = 4.82440448 - 0.00037484\, a_1^2 - 0.00098337\, a_2^2 + 0.01777408\, a_3^2, \tag{4.36a}$$

$$C_L = 0.00867623\, a_1 + 0.01397362\, a_2 + 0.0166239\, a_1a_3 - 0.01302317\, a_2a_3. \tag{4.36b}$$

The force model is highly accurate as corroborated by the r^2 scores of 0.9991 and 0.9942 for the drag and lift formulae, respectively. As shown in figure 4.8, the dynamics of the force model compares well with the real force transient dynamics, starting from the symmetric steady solution at $Re = 30$.

In the drag model (4.36a), the coefficient of a_3 is vanishing. Mode $\boldsymbol{u}_3$ actually contributes to the increase of the drag through a_3^2, as evidenced by the positive coefficient of the a_3^2 term. This is an interesting result, since the effect of the bifurcation mode $\boldsymbol{u}_3$ is to decrease the length of the recirculation bubble in the T-averaged flow field $\bar{\boldsymbol{u}}_T(\boldsymbol{x}, t) \approx \boldsymbol{u}_s(\boldsymbol{x}) + a_3(t)\boldsymbol{u}_3(\boldsymbol{x})$, resulting in an increase of the drag through the quadratic term a_3^2. This quadratic dependency is also reported in Loiseau *et al.* (2018*a*).

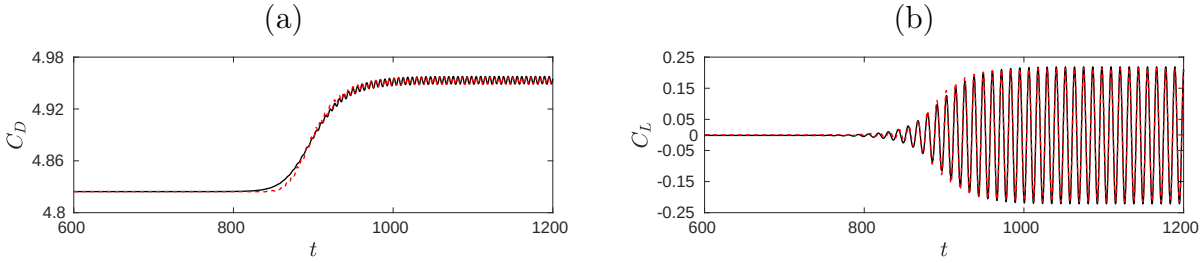


Figure 4.8: Performance of the force model with the three elementary modes of the Hopf bifurcation. Time evolution of the drag C_D (a) and lift C_L (b) coefficients, in the full flow dynamics (solid black line) and for the force model (red dashed line), at $Re = 30$. Initial condition: symmetric steady solution.

It is also worth noticing that a_3^2 contributes to the mean value of C_D while a_1^2, a_2^2 accounts for the instantaneous oscillations of C_D, as C_D oscillates at twice the vortex shedding frequency. For C_L, the oscillatory pair (a_1, a_2) fits well with the phase of the initial transient part, while the pair $(a_1 a_3, a_2 a_3)$ resolves the phase dependency of the post-transient part of the dynamics.

4.4.2 Force model associated with the supercritical pitchfork bifurcation

Next, we consider the supercritical pitchfork bifurcation, which breaks the symmetry of the symmetric steady solution $\mathbf{u_s}$ at $Re \geq Re_{\mathrm{PF}}$. In this case the antisymmetric mode $\boldsymbol{u}_4$ describes the antisymmetric instability, which corresponds to an unstable eigenmode with a real eigenvalue. The resulting Galerkin expansion reads

$$\boldsymbol{u}(\boldsymbol{x},t) = \boldsymbol{u}_s(\boldsymbol{x}) + \underbrace{a_4(t)\boldsymbol{u}_4(\boldsymbol{x})}_{\boldsymbol{u}'} + \underbrace{a_5(t)\boldsymbol{u}_5(\boldsymbol{x})}_{\boldsymbol{u}_\Delta}, \tag{4.37}$$

and the corresponding mean-field Galerkin system

$$da_4/dt = \sigma_4 a_4 - \beta_4 a_4 a_5, \tag{4.38a}$$
$$da_5/dt = \sigma_5 a_5 + \beta_5 a_4^2, \tag{4.38b}$$

where σ_4 is the positive initial growth rate, which depends on the Reynolds number. For a direct supercritical pitchfork bifurcation, $\sigma_4, \beta_4 > 0$, $\sigma_5 < 0$ and $\beta_5 > 0$, see Deng *et al.* (2020) for details.

Substituting (4.37) in equations (4.7) and (4.8), with $N = 2$ in (4.18), and with $\boldsymbol{u}_4 \in \mathcal{U}^a$ and $\boldsymbol{u}_s, \boldsymbol{u}_5 \in \mathcal{U}^s$, the force model becomes

$$C_D = C_D^\circ + l_{x;5}\ a_5 + q_{x;44}\ a_4^2 + q_{x;55}\ a_5^2, \tag{4.39a}$$
$$C_L = l_{y;4}\ a_4 + q_{y;45}\ a_4 a_5. \tag{4.39b}$$

Five parameters, namely $l_{x;0}$, $l_{x;5}$, $q_{x;44}$, $q_{x;55}$, $l_{y;4}$, $q_{y;45}$ need to be identified.

In the fluidic pinball, the pitchfork bifurcation occurs after the primary Hopf bifurcation as the Reynolds number is increased. However, the transient dynamics observed at $Re = 100$, when starting close to the symmetric steady solution, first exhibits the static symmetry breaking, which is typical of the pitchfork bifurcation, before developing the cyclic release of vortices, which is characteristic of the Hopf bifurcation. The early stage of the transient dynamics, starting from the symmetric steady solution and evolving toward one of the asymmetric steady solutions, is shown in figure 4.9. The time evolutions of the lift $C_L(t)$ and drag $C_D(t)$ coefficients are shown in figure 4.10.

Only the degrees of freedom associated with the pitchfork bifurcation are active in this early stage of the transient dynamics, as also shown in figure 4.6(a). The degrees of freedom associated

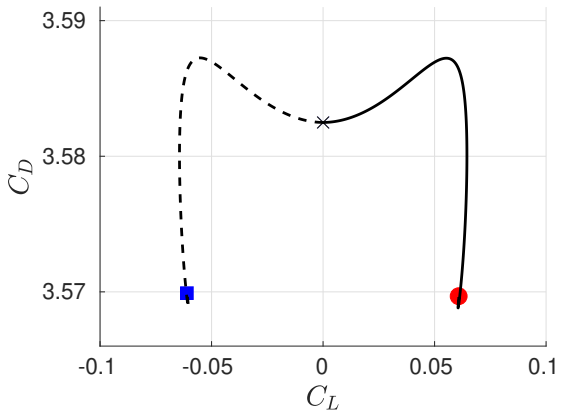


Figure 4.9: Transient trajectories (solid and dashed lines) starting from two initial conditions close to the symmetric steady solution, at $Re = 100$. Asymmetric steady solution $\boldsymbol{u}_s^+$ (•), asymmetric steady solution $\boldsymbol{u}_s^-$ (■).

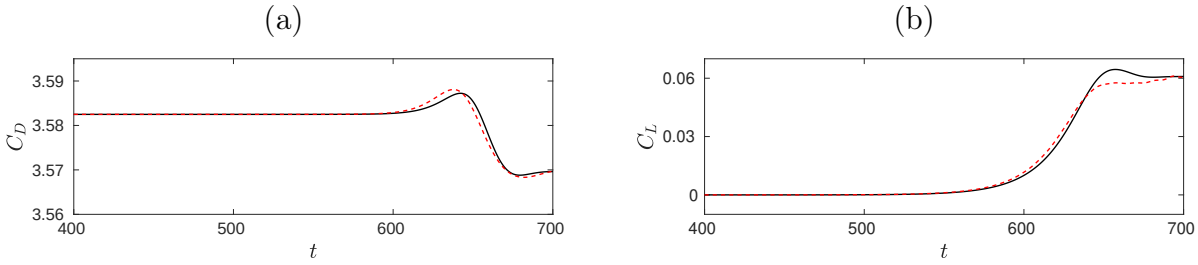


Figure 4.10: Performance of the force model with the two elementary modes of the pitchfork bifurcation. Time evolution of the drag C_D (a) and lift C_L (b) coefficients in the full flow dynamics (solid black line) and for the force model (red dashed line), at $Re = 100$. Initial condition: symmetric steady solution.

with the Hopf bifurcation will only become active further in time during the transient dynamics, which will be further discussed in § 4.5.4. Accordingly, a force model is derived for the transition after a simple pitchfork bifurcation. The training data are the lift $C_L(t)$ and drag $C_D(t)$ coefficients and the relevant mode amplitudes in Eq. (4.39) from the early to final stage of the transient dynamics. The observed slaving of a_5 in a_4^2 may reduce the robustness of the identified model. Gradually increasing the $L1$-penalty parameter in the LASSO regression, the optimized force model reads

$$C_D = 3.58248992 + 0.04367604\, a_5 - 0.08525184\, a_5^2, \tag{4.40a}$$
$$C_L = -0.13611053\, a_4 + 0.09194312\, a_4 a_5, \tag{4.40b}$$

with $r^2 = 0.9949$ for the drag model and $r^2 = 0.9992$ for the lift model. The over-fitting term a_4^2 has been eliminated in the sparse formula of the drag force. Note that the mode $\boldsymbol{u}_5$ contributes to the drag through a_5, while a_5^2 acts in decreasing the drag, as indicated by the sign of their associated coefficients in Eq. (4.40a).

Figure 4.10 compares the evolution of the drag and lift coefficients in the full flow dynamics (solid black line) to their prediction by the force model (4.40) (red dashed curve), during the early stage of the transient dynamics at $Re = 100$. The derived force model is well aligned with the real force dynamics using only two active degrees of freedom of the pitchfork bifurcation in the dynamics of the system.

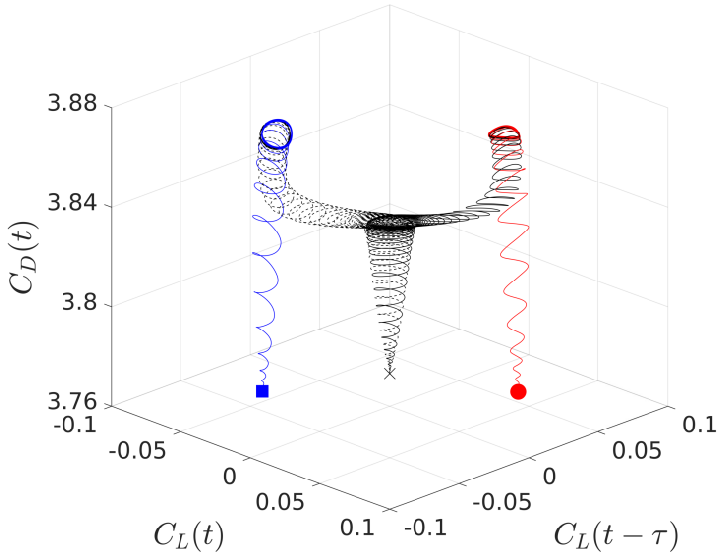


Figure 4.11: Trajectories in the time-delayed embedding space of the lift C_L and drag C_D coefficients, with $\tau = 2$, at $Re = 80$. Black trajectories starting close to the symmetric steady solution $\boldsymbol{u}_s$ ($\times$); red trajectory starting close to the asymmetric steady solution $\boldsymbol{u}_s^+$ ($\bullet$), blue trajectory starting close to the asymmetric steady solution $\boldsymbol{u}_s^-$ ($\blacksquare$).

4.5 Galerkin force model for multiple invariant sets

We focus on the regime after the pitchfork bifurcation $Re \geq Re_{\mathrm{PF}} = 68$ and before the quasi-periodic behaviour $Re \leq Re_{\mathrm{QP}} = 104$. This flow has 6 invariant sets: 3 unstable fixed points, 2 stable asymmetric mirror-conjugated periodic orbits, and one meta-stable symmetric limit cycle. § 4.5.1 investigates the dynamics of the fluidic pinball at $Re = 80$, when the degrees of freedom associated with the Hopf bifurcation are first activated before the degrees of freedom associated with the pitchfork bifurcation. The predictive power of the force model is assessed in § 4.5.2. § 4.5.3 introduces two additional degrees of freedom in the force model, in order to take into account the distortion of the shift mode when the attractor is reached. The robustness of the force model is emphasized in § 4.5.4 by considering the flow dynamics at $Re = 100$, where the pitchfork degrees of freedom are activated before the Hopf degrees of freedom during the transient dynamics.

4.5.1 Force model at $Re = 80$

At $Re = 80$, the system has already undergone a supercritical Hopf bifurcation and a supercritical pitchfork bifurcation. The trajectories issued from $\boldsymbol{u}_s$ and $\boldsymbol{u}_s^{\pm}$ are shown in the time-delayed embedding state space $(C_L(t), C_L(t-\tau), C_D(t))$ of figure 4.11. The force model will rely on five degrees of freedom at minimum, namely the three degrees of freedom associated with the Hopf bifurcation a_i, $i = 1, 2, 3$ and the two degrees of freedom a_i, $i = 4, 5$, associated with the pitchfork bifurcation. As a generalization of (4.35) and (4.39), the force model reads

$$\begin{aligned} C_D &= C_D^{\circ} + l_{x;3}\, a_3 + q_{x;11}\, a_1^2 + q_{x;12}\, a_1 a_2 + q_{x;22}\, a_2^2 + q_{x;33}\, a_3^2 \\ &\quad + l_{x;5}\, a_5 + q_{x;44}\, a_4^2 + q_{x;55}\, a_5^2 \\ &\quad + q_{x;14}\, a_1 a_4 + q_{x;24}\, a_2 a_4 + q_{x;35}\, a_3 a_5, \end{aligned} \tag{4.41a}$$

$$\begin{aligned} C_L &= l_{y;1}\, a_1 + l_{y;2}\, a_2 + q_{y;13}\, a_1 a_3 + q_{y;23}\, a_2 a_3 \\ &\quad + l_{y;4}\, a_4 + q_{y;45}\, a_4 a_5 \\ &\quad + q_{y;15}\, a_1 a_5 + q_{y;25}\, a_2 a_5 + q_{y;34}\, a_3 a_4. \end{aligned} \tag{4.41b}$$

Due to symmetry reasons, only 2 linear terms (a_3, a_5) and 9 quadratic terms (a_1^2, a_1a_2, a_1a_4, a_2^2, a_2a_4, a_3^2, a_3a_5, a_4^2, a_5^2) are left in Eq. (4.41b) for the drag coefficient. For the lift coefficient, only 3 linear terms, a_1, a_2, a_4, and 6 quadratic terms, a_1a_3, a_1a_5, a_2a_3, a_2a_5, a_3a_4, a_4a_5, are left in Eq. (4.41c). The training data is taken from the DNS starting from the three steady solutions, with the real force dynamics, see the black curves in figure 4.12, and the relevant mode amplitudes, see figure 4.6. The coefficients of the force models are identified by the sequential thresholded least-squares regression with the optimal sparsity parameter λ. We note that the LASSO regression can also be used here. See Appendix 4.B for the comparison of these two methods. The resulting force model reads

$$\begin{aligned} C_D &= 3.77331204 + 0.05888312\,a_5 - 0.01115552\,a_1^2 - 0.01088109\,a_2^2 \\ &\quad +0.01323449\,a_3^2 + 0.02949701\,a_3a_5 - 0.25910470\,a_5^2, \qquad (4.42a) \\ C_L &= \;\; 0.00953160\,a_1 \;\; + 0.00720164\,a_2 \;\; - 0.10179203\,a_4 \\ &\quad -0.00303677\,a_1a_3 - 0.00197075\,a_2a_3 - 0.00200840\,a_3a_4 \\ &\quad +0.05914386\,a_4a_5. \qquad (4.42b) \end{aligned}$$

The good accuracy of the identified drag model can be determined from the high r^2 score of 0.9816. The drag model of Eq. (4.42a) preserves both the basic forms of the drag model for the Hopf and pitchfork bifurcations and the signs of the coefficients. This indicates that the identified model is robust. The only remaining cross-term a_3a_5 provides the coupling relation between the degrees of freedom associated with both bifurcations.

A robust sparse formula for the lift model is more difficult to derive, due to the oscillating dynamics of the lift and the fact that a_4 and a_5 also oscillate at the fundamental frequency. With respect to the basic lift model of two bifurcations, a balanced method is used here to solve the difficulty of the identification. Starting with a large $L1$-penalty, the derived under-fitted system can figure out the most elementary features of the dynamics, eliminating a_1a_5, a_2a_5, a_4a_5. This is reasonable as a_3 is about ten times larger than a_5, which means that most of the mean-field distortion comes from $\boldsymbol{u}_3$. However, if the $L1$-penalty is too large, the term a_4a_5 can disappear from the lift model, making the resulting model non-consistent with Eq. (4.42b). In order to balance sparsity and robustness, a_4a_5 needs to be reintroduced into the library. The sparse formula of the lift model in Eq. (4.42b) is determined by least-squares regression, constraining the parameters of a_1a_5, a_2a_5 to zero. The r^2 score of the identified lift model is 0.9673.

The identified force dynamics in Eq. (4.42) (dashed red line) is compared to the real force dynamics (solid black line) at $Re = 80$ in figure 4.12. The force model based on the least-order model can reproduce the main features of the real force dynamics. The drag model of Eq. (4.42a) shows how the degrees of freedom of the Hopf (a_1^2, a_2^2, a_3^2) and pitchfork (a_5, a_5^2) bifurcations contribute to the drag force, as well as the coupling between these degrees of freedom (a_3a_5). The lift model of Eq. (4.42b) shows that the lift oscillations occur through the coupling of the oscillating degrees of freedoms a_1, a_2 to a_3, while the coupling between the degrees of freedoms a_4 and a_5 contribute to the mean value of C_L. Hence, the mean lift coefficient can be simplified with fewer terms, as $\overline{C_L} = l_{y;4}\ a_4 + q_{y;45}\ a_4a_5 + q_{y;34}\ a_3a_4$, which meets well with the Krylov-Bogoliubov assumption (Jordan & Smith, 1999).

4.5.2 Assessing the predictive power of the force model

The time-evolution of the drag and lift coefficients in the fluidic pinball are shown in figure 4.12 as solid black lines. The evolutions of the drag and lift coefficients in the model (4.42) are shown with dashed red lines. The model reproduces correctly the time scales of the force dynamics as well as the transient and asymptotic amplitudes of the forces. However, it is observed that the fine details of the transient dynamics, at the early stage of the linear instability, are not satisfactorily reproduced in the identification process (figure 4.12(a,b) at $t \approx 590$ and 475 respectively). The

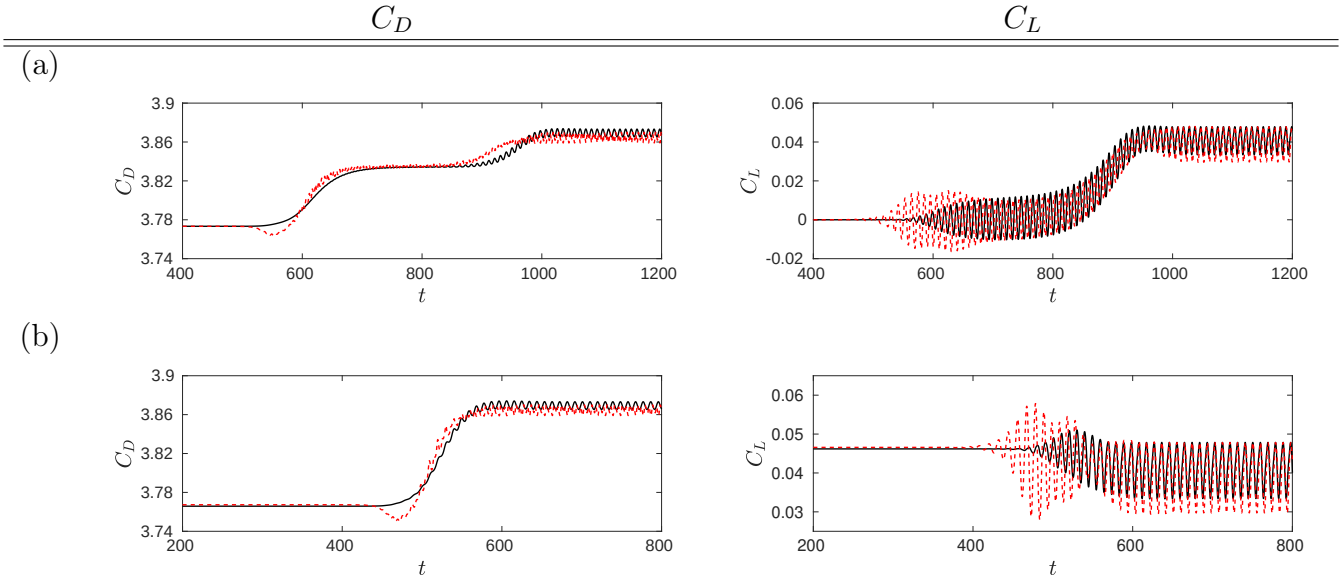


Figure 4.12: Performance of the force model with the five elementary modes. Time evolution of the drag C_D (left) and the lift C_L (right) coefficients in the full flow dynamics (solid black line) and for the force model (red dashed line), at $Re = 80$. Initial condition: (a) symmetric steady solution, (b) asymmetric steady solution.

ranges of time concerned, in both cases, are also associated with oscillations in a_4, as observed during the initial stage at $t \approx 590$ in figure 4.6(left) and $t \approx 475$ in figure 4.6(right). This strongly suggests that the oscillations of a_4 be triggered by the degrees of freedom associated with the Hopf bifurcation. This means that the degrees of freedom of the pitchfork bifurcation are affected by the degrees of freedom of the Hopf bifurcation, at least when the distance from the bifurcation point is large enough, which is the case at $Re = 80$.

In addition, as recalled in § 4.3.3, at $Re \approx 68$, both the steady symmetric solution and the symmetric-based limit cycle undergo a supercritical pitchfork bifurcation. We emphasize that this coincidence of two local pitchfork bifurcations might not occur by chance, as mentioned in Deng *et al.* (2020). As a result of these two simultaneous bifurcations, the degrees of freedom involved in the pitchfork bifurcation of the fixed point might not coincide with those involved in the pitchfork bifurcation of the limit cycle. For this reason, it is reasonable to introduce two distinct sets of degrees of freedom for each of them, namely a_4, a_5 at the fixed point and a_6, a_7 at the limit cycle. These two additional degrees of freedom will complete the mean-field model with more details and will take into account the mean-field distortion during the transition from the fixed point to the limit cycle. The new resulting mean-field Galerkin system, with seven degrees of freedom, is derived in appendix 4.E, while the new resulting force model is discussed in the next subsection.

4.5.3 The need for additional modes

All our attempts to smooth out the kicks observed at the beginning of the exponential growth, in both C_D and C_L in the frame of the force model (4.42), failed, even when over-fitting the model without any sparsity. This strongly indicates that five degrees of freedom might not be sufficient to account for the force evolution on the full-time range.

Digging into this idea, it becomes manifest that the way $\boldsymbol{u}_3(\boldsymbol{x})$ is built, namely as the difference between the statistically symmetric mean flow field, associated with the unstable limit cycle, and the symmetric steady solution $\boldsymbol{u}_s(\boldsymbol{x})$,

$$\boldsymbol{u}_3(\boldsymbol{x}) = \bar{\boldsymbol{u}}_T(\boldsymbol{x}, 775) - \boldsymbol{u}_s(\boldsymbol{x}), \tag{4.43}$$

see figure 4.12(a), does not allow to satisfactorily account for the complete dynamics of the lift and

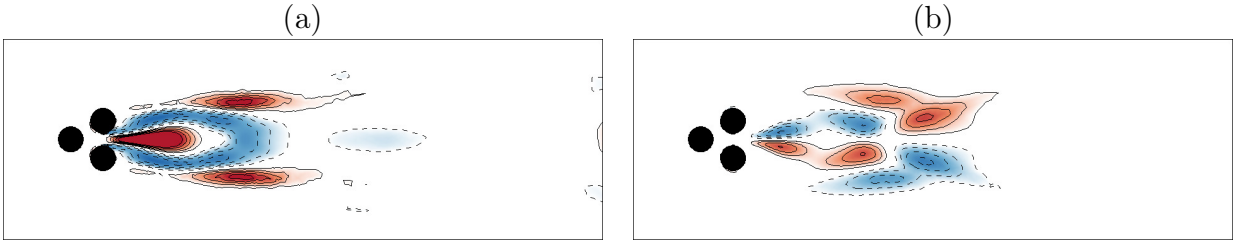


Figure 4.13: Vortical structure (color) of the modes $\boldsymbol{u}_6(\boldsymbol{x})$ (a), $\boldsymbol{u}_7(\boldsymbol{x})$ (b), at $Re = 80$.

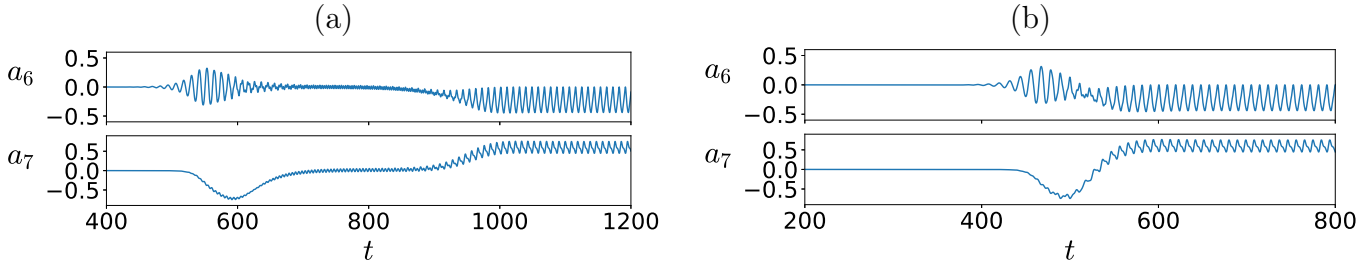


Figure 4.14: Mode amplitudes $a_{6,7}(t)$ in the full-flow dynamics starting (a) from the symmetric steady solution $\boldsymbol{u}_s$, (b) from the asymmetric steady solution $\boldsymbol{u}_s^+$, at $Re = 80$.

drag forces. This also indicates that $\boldsymbol{u}_3(\boldsymbol{x})$ gets *distorted* when the system is evolving along the manifold, which connects the unstable limit cycle to one of the two conjugated stable limit cycles. In other words, the mean-field distortion on the attractors associated with the two asymmetric mean flow fields $\bar{\boldsymbol{u}}^{\pm}$, namely

$$\boldsymbol{u}_3^{\pm}(\boldsymbol{x}) = \bar{\boldsymbol{u}}^{\pm}(\boldsymbol{x}) - \boldsymbol{u}_s^{\pm}(\boldsymbol{x}) \tag{4.44}$$

do not coincide exactly with $\boldsymbol{u}_3(\boldsymbol{x})$. The asymmetric mean flow fields $\bar{\boldsymbol{u}}^{\pm}$ only focus on the post-transient dynamics, as shown in figure 4.12(c), which can be expressed with $\bar{\boldsymbol{u}}_T^{\pm}(\boldsymbol{x}, 700)$. The difference between $\boldsymbol{u}_3^{\pm}(\boldsymbol{x})$ and $\boldsymbol{u}_3(\boldsymbol{x})$ is asymmetric and can be decomposed into a symmetric and an anti-symmetric part, respectively $\boldsymbol{u}_6(\boldsymbol{x})$ and $\boldsymbol{u}_7(\boldsymbol{x})$:

$$\boldsymbol{u}_3^{\pm}(\boldsymbol{x}) - \boldsymbol{u}_3(\boldsymbol{x}) = \pm\boldsymbol{u}_6(\boldsymbol{x}) + \boldsymbol{u}_7(\boldsymbol{x}). \tag{4.45}$$

As a result, the modes $\boldsymbol{u}_6(\boldsymbol{x})$ and $\boldsymbol{u}_7(\boldsymbol{x})$ can be defined as,

$$\boldsymbol{u}_6(\boldsymbol{x}) \propto \bar{\boldsymbol{u}}^{+}(\boldsymbol{x}) - \bar{\boldsymbol{u}}^{-}(\boldsymbol{x}), \tag{4.46a}$$
$$\boldsymbol{u}_7(\boldsymbol{x}) \propto (\bar{\boldsymbol{u}}^{+}(\boldsymbol{x}) + \bar{\boldsymbol{u}}^{-}(\boldsymbol{x})) - 2\bar{\boldsymbol{u}}_T(\boldsymbol{x}, 775). \tag{4.46b}$$

After orthogonal normalization by a Gram-Schmidt procedure, the resulting modes are shown in figure 4.13, with their mode amplitudes in figure 4.14. When comparing the definitions of $\boldsymbol{u}_6$ and $\boldsymbol{u}_7$ in Eq. (4.46) and of $\boldsymbol{u}_4$ and $\boldsymbol{u}_5$ in Eq. (4.31)–(4.32), it is not surprising that the spatial structure of $\boldsymbol{u}_6$, resp. $\boldsymbol{u}_7$ (see figure 4.13), be so similar to the spatial structure of $\boldsymbol{u}_4$, resp. $\boldsymbol{u}_5$ (see figure 4.5). To be mentioned, $\boldsymbol{u}_6$, $\boldsymbol{u}_7$ as defined in (4.46), would be equivalent to the pitchfork modes $\boldsymbol{u}_4$, $\boldsymbol{u}_5$ built on the periodic solutions instead of being built on the steady solutions. However, after the Gram-Schmidt procedure, the $\boldsymbol{u}_6$, $\boldsymbol{u}_7$ modes of figure 4.13 have been transformed into corrective modes of $\boldsymbol{u}_4$, $\boldsymbol{u}_5$ when departing from the steady solutions and approaching the asymptotic limit cycles. The corrective modes $\boldsymbol{u}_6$, $\boldsymbol{u}_7$ should be slaved to $\boldsymbol{u}_4$, $\boldsymbol{u}_5$ along the mean field distortion of $\boldsymbol{u}_3$. The corresponding slaving relation will not be discussed in this paper. Hence, the combination of $\boldsymbol{u}_i$, $i = 4, \ldots, 7$, works as a flexible pitchfork mode expansion, which adapts the whole phase space where all the invariant sets (steady/periodic) locate.

In figure 4.14, the transient dynamics of a_6, a_7 shows to be also similar to a_4, a_5 in figure 4.6. Not surprisingly, the opposite initial bump of a_6, a_7 helps to better fit the dynamics on the

manifold. Besides, a_6, a_7 show no contribution close to the steady solutions, as their role is to adapt the modes $\boldsymbol{u}_4$, $\boldsymbol{u}_5$ when approaching the stable limit cycle.

The force model identification is more challenging with these two additional modes. High robustness is required for our force model without losing the identified terms in § 4.5.1. Compared to the force formula (4.41) with five modes, 8 new terms are introduced in the drag formula, namely $a_7, a_1a_6, a_2a_6, a_4a_6, a_6^2, a_3a_7, a_5a_7, a_7^2$, and 7 additional terms are considered in the lift formula, namely $a_6, a_3a_6, a_5a_6, a_1a_7, a_2a_7, a_4a_7, a_6a_7$. Due to the similar transient dynamics of a_4, a_5 and a_6, a_7, the corrective degrees of freedom a_6, a_7 can easily replace a_4, a_5 in the identified model. Hence, the original structure of the force model with five modes could be lost. To avoid possible over-fitting, we need to free the active terms gradually and constraint the parameters of a_4, a_5 during the sparse regression to ensure the robustness of the result. In addition, the newly introduced terms should work as a corrective function to the original force model with five degrees of freedom. In other words, the new force model with seven degrees of freedom should inherit the original structure of Eq. (4.42).

Based on the structure of the drag model (4.42a), the terms a_7, a_7a_7, a_3a_7, a_4a_6 and a_5a_7 are introduced in the extended model. The terms a_1a_6, a_2a_6, a_6^2 are firstly set to zero because their corresponding terms a_1a_4, a_2a_4, a_4^2 in Eq. (4.42a) are vanishing. In order to improve the robustness of the regression results, the terms a_5 and a_5^2 are constrained with the values from Eq. (4.42a). Increasing the $L1$-penalty of the LASSO regression, $l_{x;7}$, $q_{x;46}$ and $q_{x;77}$ vanish successively, and an obvious under-fitting starts when losing $q_{x;35}$. The introduced terms $q_{x;37}$, $q_{x;57}$ are robust with few possibility of over-fitting. Eventually, the drag model reads

$$\begin{aligned} C_D &= 3.77331204 + 0.05888312a_5 - 0.00169970a_1^2 - 0.00156775a_2^2 \\ &\quad +0.00513885a_3^2 + 0.00786294a_3a_5 + 0.00950204a_3a_7 \\ &\quad -0.25910470a_5^2 - 0.06264888a_5a_7. \end{aligned} \tag{4.47}$$

Eq. (4.47) preserves the original form of Eq. (4.42a), with tiny changes of the coefficients. This extended model fits well the dynamics of the drag coefficient, with the r^2 score increasing to 0.9981, also can be seen with the red dashed curve of figure 4.15(left).

As already mentioned, the drag monotonously increases with the development of the vortex shedding. This is obvious, for instance, from figure 4.15, when the lift starts to oscillate and the drag to increase. The positive signs of $q_{y;33}$, $q_{y;35}$ and $q_{y;37}$, in the drag model of Eq. (4.47), are responsible for this monotonous increase of the drag. Compared to the drag model with only a_5 in § 4.5.1, the contribution to the drag of a_5 and a_7 is more subtle. They contribute to an increase of the drag through a_5, a_5a_3 and a_7a_3, while they promote a decrease of the drag through a_5^2 and a_5a_7. As a non-trivial result, the statistically asymmetric (stable) limit cycles have a larger drag than the statistically symmetric (unstable) limit cycle, while the asymmetric steady solutions have a lower drag than the symmetric steady solution. This is obvious in figure 4.11 when considering the relative positions of the three steady solutions and three limit cycles along the C_D axis. Note that the parameters $q_{x;11}$, $q_{x;22}$ and $q_{x;57}$ all own negative signs but are relatively small. The two parameters $q_{x;11}$, $q_{x;22}$ solely contribute to the oscillating dynamics, as discussed in §4.4.1, while $q_{x;57}$ optimizes the fitting result when evolving toward the attracting limit cycles.

Analogously, for the lift model, the values of $l_{y;4}$ and $q_{y;45}$ are taken from the identified lift model in Eq. (4.42b), while $q_{y;17}$, $q_{y;27}$ are set to zero for consistency with the structure of Eq. (4.42b), in which $q_{y;15}$, $q_{y;25}$ are absent. Based on the structure of model (4.42b), the terms a_6, a_6a_7, a_3a_6, a_5a_6 and a_4a_7 are introduced in the extended model. The final sparse form is identified by the LASSO regression with gradually increasing the $L1$-penalty. A sparse lift model, compatible with the structure of Eq. (4.42b), is derived as

$$\begin{aligned} C_L &= 0.00762433\,a_1 + 0.01102097\,a_2 - 0.10179203\,a_4 - 0.03129798\,a_6 \\ &\quad -0.00141416\,a_1a_3 - 0.00289952\,a_2a_3 + 0.00656293\,a_3a_4 - 0.01082375\,a_3a_6 \\ &\quad +0.05914386\,a_4a_5 + 0.02365784\,a_4a_7 - 0.03348935\,a_5a_6. \end{aligned} \tag{4.48}$$

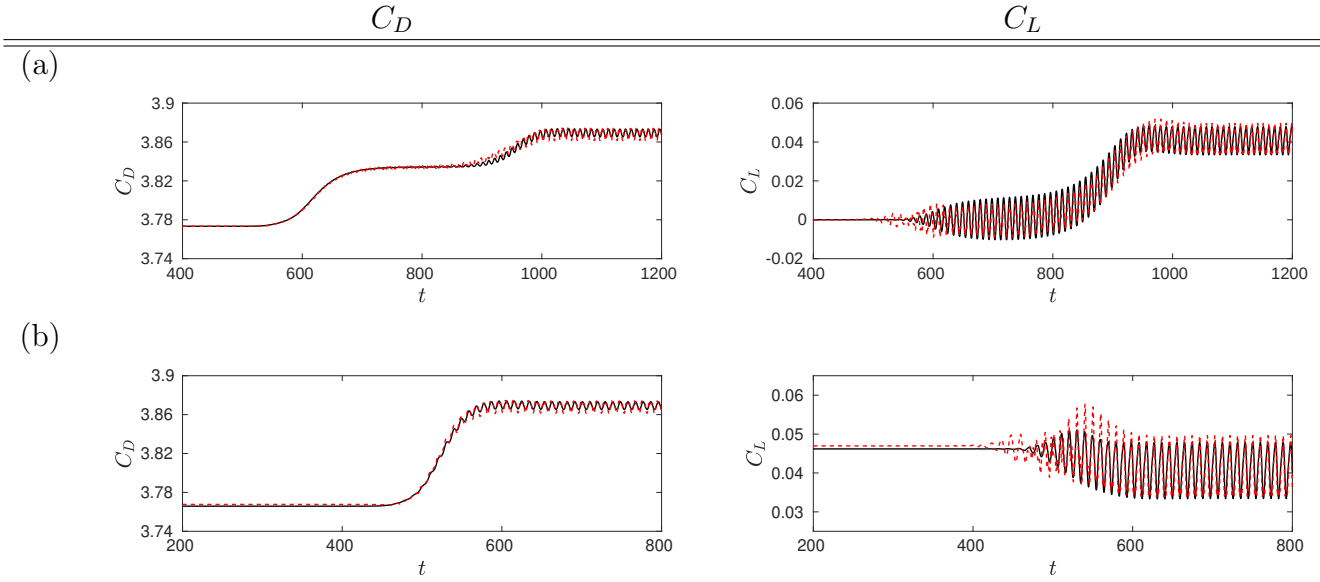


Figure 4.15: Performance of the force model with two additional slaved corrective modes. Time evolution of the drag C_D (left) and lift C_L (right) coefficients in the full flow dynamics (solid black line) and for the force model (red dashed line), at $Re = 80$. Initial condition: (a) symmetric steady solution $\boldsymbol{u}_s$, (b) asymmetric steady solution $\boldsymbol{u}_s^+$.

In addition to the lift model of Eq. (4.42b), the lift model of Eq. (4.48) contains the terms a_6, a_3a_6 and a_5a_6, as well as the coupling between a_4 to a_7. The r^2 score has increased to 0.9952. Both the oscillating dynamics in the early stage and the symmetry-breaking stage are better reproduced for the lift coefficient, as the red dashed curve of figure 4.15(right) proves.

With the two additional degrees of freedom a_6, a_7, the time evolution of the drag and lift coefficients are well reproduced, as shown in figure 4.15(a,b). Without notable changes of the original lift structure, the phase of the lift dynamics is now correctly caught along with the complete transient dynamics.

4.5.4 Force model at $Re = 100$

In § 4.4.2, we derived a basic force formula for the primary stage of the transient evolution at $Re = 100$, when only the degrees of freedom of the pitchfork bifurcation were involved. We now consider the complete force evolution at $Re = 100$. Figure 4.16 shows trajectories issued from the three different steady solutions in the three-dimensional time-delayed embedding space of C_L and C_D. The black trajectory, issued from the symmetric steady solution $\boldsymbol{u}_s$ (black cross $\times$ in figure 4.16) first approaches the asymmetric steady solution $\boldsymbol{u}_s^+$ (red point) before escaping out of it and eventually reaching the stable (statistically asymmetric) limit cycles around $\bar{\boldsymbol{u}}^+$.

The same mode decomposition strategy is proposed, resulting in a reduced-order model with 7 modes. The mode amplitudes from two DNS, starting from either the symmetric steady solution $\boldsymbol{u}_s$ (a) or the asymmetric steady solution $\boldsymbol{u}_s^+$ (b), are shown in figure 4.17.

As already observed in figure 4.10, the drag coefficient (solid black line) in figure 4.18(a) exhibits a minimal value for a transient state around $t \approx 700$. This transient state is the asymmetric steady solution $\boldsymbol{u}_s^+$ (red circle of figure 4.16). In the frame of our modal decomposition (4.30), $\boldsymbol{u}_s^+$ is approximated as

$$\boldsymbol{u}_s^+ \approx \boldsymbol{u}_s + a_4(700)\boldsymbol{u}_4 + a_5(700)\boldsymbol{u}_5, \tag{4.49}$$

with only a_4 and a_5 being active in the dynamics of the fluidic pinball, as can be seen in figure 4.17(a). From Eq. (4.40a), the drag coefficient only depends on a_5 and a_5^2, which actually contribute to the transitory increase and an overall decrease on the drag. This is fully consistent with the transition of the drag coefficient observed in figures 4.10(a) and 4.18(a) from $t = 300$ to 700; a_5

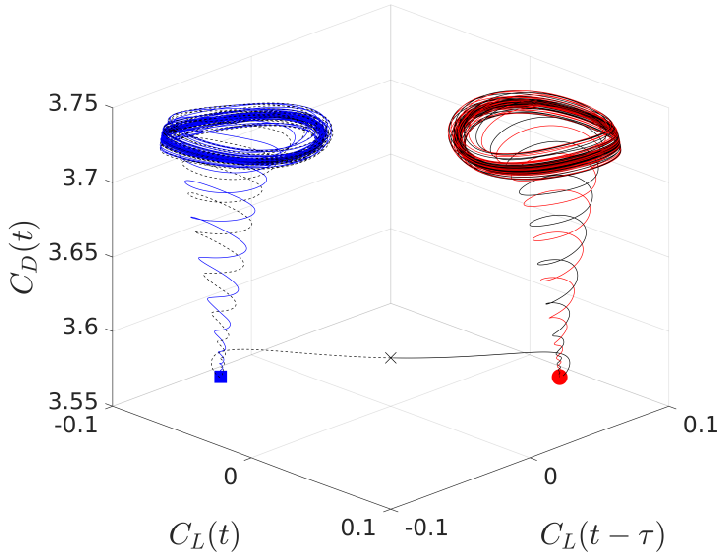


Figure 4.16: Trajectories in the time-delayed embedding space of the lift C_L and drag C_D coefficients, with $\tau = 2$, at $Re = 100$. Black trajectories starting close to the symmetric steady solution $\boldsymbol{u}_s$ ($\times$); red trajectory starting close to the asymmetric steady solution $\boldsymbol{u}_s^+$ ($\bullet$), blue trajectory starting close to the asymmetric steady solution $\boldsymbol{u}_s^-$ ($\blacksquare$).

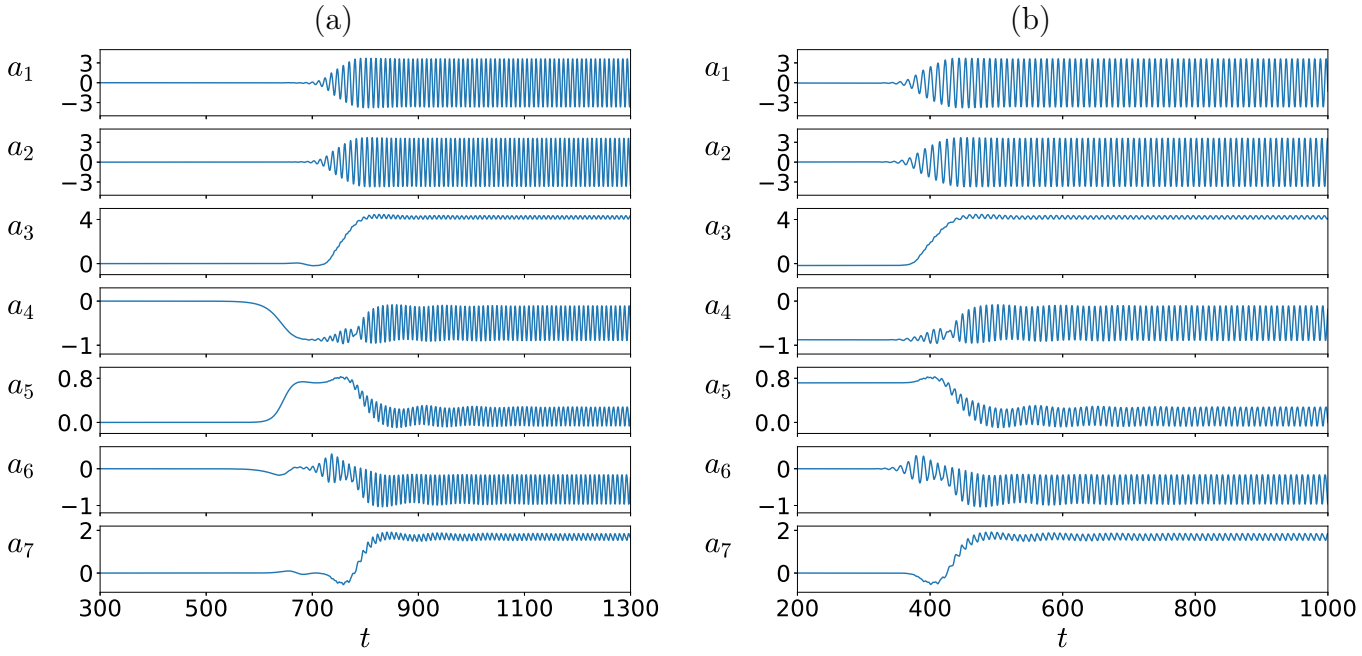


Figure 4.17: Mode amplitudes $a_{1,\dots,7}(t)$ in the full-flow dynamics starting (a) from the symmetric steady solution $\boldsymbol{u}_s$, (b) from the asymmetric steady solution $\boldsymbol{u}_s^+$, at $Re = 100$.

is found to contribute to the initial rising of C_D, around $t \approx 650$, while a_5^2 contributes to the subsequent decrease of the drag coefficient, around $t = 700$. The degrees of freedom associated with the Hopf bifurcation become active later during the transient dynamics, when the state space orbit leaves the unstable asymmetric steady solution $\boldsymbol{u}_s^+$ toward the stable attracting limit cycle around $\bar{\boldsymbol{u}}_s^+$.

The training data is the real force coefficients and the mode amplitudes taken from the DNS starting with the three different steady solutions to the final asymptotic regimes. Following the same calibration procedure as for $Re = 80$, we first apply the LASSO regression for the force model with the five leading degrees of freedom, and then introduce the two additional degrees of freedom a_6, a_7 into the regression for optimization. Performing the sparse regression in this way can prevent the elimination of a_4, a_5 and ensure the corrective effect of a_6, a_7, thereby improving the robustness of the identification. The force model at $Re = 100$ reads

$$\begin{aligned} C_D &= 3.58248992 + 0.04367604\, a_5 - 0.00302817\, a_1^2 - 0.00354079\, a_2^2 \\ &\quad +0.00158873\, a_3^2 + 0.02169661\, a_3 a_5 + 0.02223079\, a_3 a_7 \\ &\quad -0.08525184\, a_5^2 - 0.04763643\, a_5 a_7, \qquad (4.50a) \\ C_L &= 0.00346208\, a_1 + 0.00269236\, a_2 - 0.13611053\, a_4 + 0.05962648\, a_6 \\ &\quad +0.00029274\, a_1 a_3 - 0.00045784\, a_2 a_3 + 0.00389912\, a_3 a_4 - 0.02102284\, a_3 a_6 \\ &\quad +0.09194312\, a_4 a_5 + 0.02056288\, a_4 a_7 - 0.10980990\, a_5 a_6. \qquad (4.50b) \end{aligned}$$

with $r^2 = 0.9984$ for the drag model of Eq. (4.50a), and $r^2 = 0.9901$ for the lift model of Eq. (4.50b). As shown in figure 4.18, the force model fits well the time evolution of the drag and lift coefficients. Moreover, Eqs. (4.47), (4.48) and (4.50) own the same active terms. Henceforth, the drag force model preserves the same structure with the same signs of the active terms as the Reynolds number is increased. In addition, although the transient dynamics at $Re = 80$ and 100 are qualitatively very different, with the seven degrees of freedom differently activated during the transient, the force model of Eq. (4.47)–(4.48) is still consistent at $Re = 100$, with the correctly identified mean-field model. For the lift model (4.50b), we notice the same structure with the sign changes for the terms $a_1 a_3$ and a_6, compared to Eq. (4.48), which is acceptable for the oscillating dynamics. Compatible with the basic lift force model, the lift force model with seven degrees of freedom also correctly identifies the force transitions, as shown in figure 4.18(right).

4.6 Conclusions and outlook

We proposed aerodynamic force formulae complementing mean-field POD Galerkin models for the unforced fluidic pinball. The starting point is a general Galerkin method for unsteady incompressible viscous flow around a stationary body. First, the instantaneous force is derived as a constant-linear-quadratic function of the mode amplitudes from first principles. The viscous and pressure contributions to the force are directly obtained from the Galerkin expansion and lead to a constant-linear-quadratic force in terms of the mode amplitudes.

These terms lead to corresponding changes in the flow from which the force can also be derived. One contribution from the convective term describes the momentum flux contribution. The additional contribution from the local acceleration requires the Galerkin system to replace the time derivatives of the mode amplitudes by a state function. In contrast to the pioneering work by Noca *et al.* (1999), the derivation is valid for arbitrary multiply connected domains.

The drag and lift formula is simplified for the fluidic pinball model exploiting the symmetry of the modes. About half of the terms can be discarded on the grounds of symmetry. A second simplification is performed with a sparse calibration of the remaining coefficients. The sparsity parameter λ penalizes any non-vanishing term and yields sparse human-interpretable expressions. The challenges of the purely projection-based approach is discussed in Appendix 4.C, and the challenges of using standard POD modes is elaborated in Appendix 4.D.

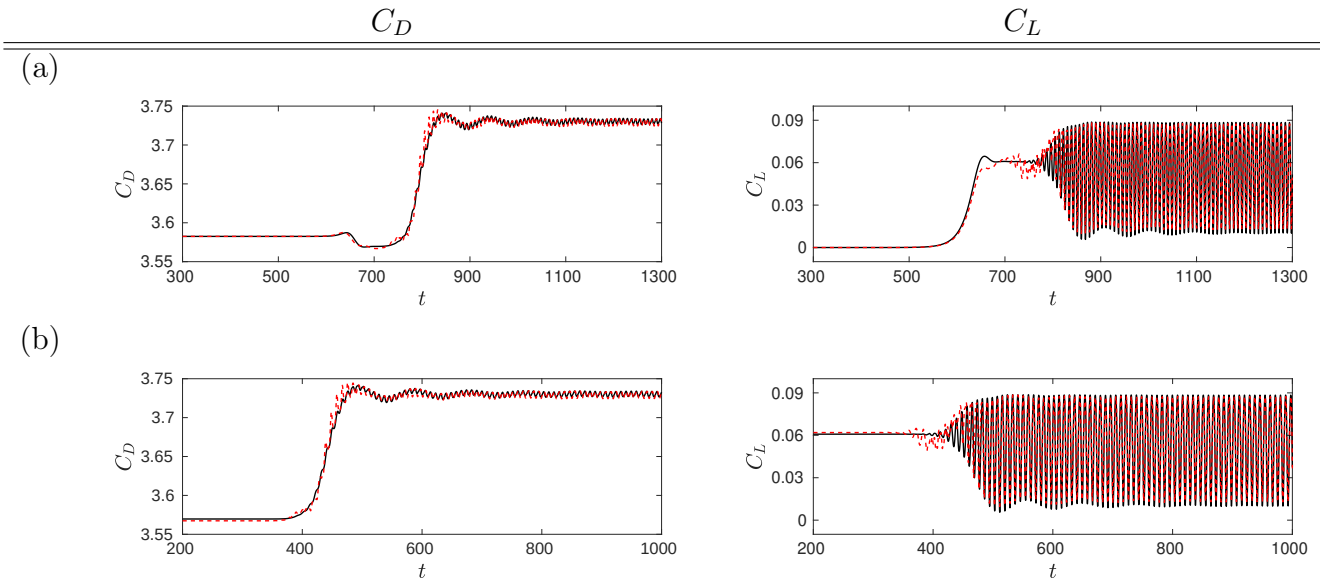


Figure 4.18: Performance of the force model with two additional slaved corrective modes. Time evolution of the drag C_D (left) and the lift C_L (right) coefficients in the full flow dynamics (solid black line) and for the force model (red dashed line), at $Re = 100$. Initial condition: (a) symmetric steady solution $\boldsymbol{u}_s$, (b) asymmetric steady solution $\boldsymbol{u}_s^+$.

The sparse force model methodology is applied to three transient dynamics: (1) the periodic regime of statistically symmetric vortex shedding at $Re = 30$, (2) the periodic regime of statistically asymmetric vortex shedding at $Re = 100$, and (3) the same regime at $Re = 80$ but with metastable statistically symmetric periodicity.

The transient dynamics at $Re = 30$ from the steady solution to the limit cycle is resolved by standard third-order mean-field Galerkin model with two oscillatory modes for vortex shedding and one shift mode for the mean-field distortion (Noack *et al.*, 2003). The drag formula includes the squares of all mode amplitudes consistent with the second harmonic fluctuations. The drag monotonically increases during the transient. The lift formula includes the amplitudes of the von Kármán modes and their products with the shift mode, consistent with expectations. Its oscillation increases until the limit cycle is reached.

The dynamics at $Re = 100$ after the Hopf and pitchfork bifurcation has three unstable fixed points, one symmetric steady solution and a mirror-symmetric pair of asymmetric ones. The transients from these fixed points terminate in one of the asymmetric limit cycles corresponding to the asymmetric shedding states. This dynamics is described by a fifth-order Galerkin model (Deng *et al.*, 2020), where the first three modes resolve the Hopf bifurcation and the next two modes the pitchfork bifurcation. The associated drag formula contains the terms of $Re = 30$. In addition, the drag is modified by linear and quadratic terms with the shift modes associated with the Hopf and pitchfork instabilities. These additional terms vanish without pitchfork bifurcation and do not introduce harmonics of vortex shedding. Similarly, the lift formula generalizes the expression at $Re = 30$.

The intermediate Reynolds number 80 leads to a more complex force model, as the transients may pass through a meta-stable symmetric limit cycle. The accuracy of the force model could significantly be increased by two additional Galerkin modes which resolve variations between symmetric and asymmetric limit cycles. The drag and lift formulae were correspondingly longer and good agreement with computational data is achieved.

Summarizing, the sparse force model describes multi-attractor behaviour of the unforced fluidic pinball even for complex dynamics with three steady and three periodic solutions. For this configuration, we have the advantage of a thorough understanding of the dynamics via a low-dimensional mean-field Galerkin model. We envision successful applications of sparse regression

for aerodynamic forces for turbulent flows, e.g., for the bi-stable behaviour of the Ahmed body wake (Grandemange *et al.*, 2013; Östh *et al.*, 2014; Barros *et al.*, 2017).

The force formula may be particularly instructive for drag reduction with active control (Choi *et al.*, 2008). Given a Galerkin model, the force formula indicates beneficial regions of the state space. Thus, an upfront kinematical insight is gained in which direction control needs to 'push' the attractor. For instance, the third-order mean-field model and the force formula implies that stabilization is required for drag reduction consistent with earlier studies of Protas (2004); Bergmann & Cordier (2008). Future generalizations may also profit from stochasticity (Sapsis & Lermusiaux, 2009).

Acknowledgements

N. Deng appreciates the support of the China Scholarships Council (No.201808070123) during his Ph.D. Thesis in the ENSTA Paris of Institut Polytechnique de Paris, and numerical supports from the laboratories LIMSI (CNRS-UPR 3251) and IMSIA (UMR EDF-ENSTA-CNRS-CEA 9219).

This work is supported by a public grant overseen by the French National Research Agency (ANR) by grant 'FlowCon' (ANR-17-ASTR-0022), and by Polish Ministry of Science and Higher Education (MNiSW) under the Grant No.: 0612/SBAD/3567.

We appreciate valuable discussions with Guy Cornejo Maceda, Franǫis Lusseyran, and Colin Leclercq. We thank the anonymous referees for their insightlful suggestions which have inspired some of our investigations.

Declaration of Interests. The authors report no conflict of interest.

4.A Forces from the momentum balance

The forces can be alternatively derived from the residual of the Navier-Stokes equations

$$\boldsymbol{R}(\boldsymbol{u},p) := \partial_t \boldsymbol{u} + \nabla\cdot\boldsymbol{u}\otimes\boldsymbol{u} - \nu\triangle\boldsymbol{u} + \nabla p \tag{4.51}$$

in the domain Ω. This domain is assumed to enclose the obstacle and extend sufficiently far away from the obstacle such that the free-stream condition $\boldsymbol{u} = \boldsymbol{e}_x$ can be applied on the left, top and bottom boundaries of the fluid domain Ω. The domain boundary $\partial\Omega$ contains the surface of the immersed body Γ and the outer surface S_∞. It should be noted that the surface element dS on the body points inside the body, i.e., opposite to the direction in § 4.2.2.

The force in direction $\boldsymbol{e}_\alpha$ is derived from the integrated momentum balance in that direction.

$$(\boldsymbol{e}_\alpha, \boldsymbol{R}(\boldsymbol{u},p))_\Omega = 0. \tag{4.52}$$

Four terms are obtained. The first contribution is the viscous term. This term can be converted into a skin friction integral over Γ and S_∞. The contribution over the outer integral vanishes under free-stream conditions. The remaining contribution is the viscous force applied to the immersed body:

$$(\boldsymbol{e}_\alpha, \nu\triangle\boldsymbol{u})_\Omega = \nu\boldsymbol{e}_\alpha \cdot \oint\limits_{\Gamma+S_\infty} \left(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^{\mathrm{T}}\right)\cdot\boldsymbol{n}\, dS \;= F^\nu_\alpha, \tag{4.53}$$

where $\boldsymbol{e}_\alpha \cdot \left(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^{\mathrm{T}}\right)\cdot\boldsymbol{n} = 2\sum_{\alpha,\beta=x,y,z} S_{\alpha,\beta}\, n_\beta$.

The second contribution is the pressure term which can analogously reduce to the pressure force on the immersed body:

$$(\boldsymbol{e}_\alpha, -\nabla p)_\Omega = -\oint\limits_{\Gamma+S_\infty} p\, n_\alpha\, dS = F^p_\alpha. \tag{4.54}$$

Not surprisingly, we arrive at the formula of § 4.2.2. The force exerted on the body is equal but opposite to the force exerted on the fluid.

The third term is the local acceleration:

$$\left(\boldsymbol{e}_\alpha, \partial_t \left[\sum_{j=0}^{N} a_j(t)\boldsymbol{u}_j(\boldsymbol{x})\right]\right)_\Omega = \sum_{j=1}^{N} m^t_{\alpha;j} \frac{da_j}{dt}(t), \tag{4.55}$$

where $m^t_{\alpha;j} = (\boldsymbol{e}_\alpha, \boldsymbol{u}_j)_\Omega$.

The fourth term arises from the convective acceleration:

$$\left(\boldsymbol{e}_\alpha, \nabla \cdot \left(\left[\sum_{j=0}^{N} a_j\boldsymbol{u}_j\right] \otimes \left[\sum_{k=0}^{N} a_k\boldsymbol{u}_k\right]\right)\right)_\Omega = \sum_{j,k=0}^{N} q^c_{\alpha;jk} a_j a_k \tag{4.56}$$

where $q^c_{\alpha;jk} = (\boldsymbol{e}_\alpha, \nabla \cdot [\boldsymbol{u}_j \otimes \boldsymbol{u}_k])_\Omega$. The volume integral over Ω can be converted into a momentum flux surface integral over the boundary.

Making use of the momentum balance (4.52), the third and fourth contributions from the acceleration terms equal the total force:

$$F_\alpha = \sum_{j=1}^{N} m^t_{\alpha;j} \frac{da_j}{dt} + \sum_{j,k=0}^{N} q^c_{\alpha;jk} a_j a_k. \tag{4.57}$$

This force formula contains constant, linear and quadratic terms of the mode amplitudes as well as their time derivatives. The state-dependent formula (4.18) may be obtained from (4.57) by replacing the time derivatives with (4.6). The total forces on the immersed body are here again represented by a constant-linear-quadratic expression.

The above mentioned formulae dresses Newton's second law $\boldsymbol{F} = m\boldsymbol{a}$ in a Galerkin framework for fluid flow. Eq. (4.57) corresponds to '$m\boldsymbol{a}$' and is purely based on the fluid motion. Eq. (4.18) corresponds to '$\boldsymbol{F}$' and allows distinguishing between the contribution of viscous and pressure stresses.

4.B Influence of the sparsity parameter and regression methods

In the SINDy algorithms, the sparsity parameter is either the $L1$-penalty for the LASSO regression or the threshold for the sequential thresholded least squares (STLS) regression. We denote the $L1$-penalty and the threshold as the sparsity parameter λ in both cases. These two methods can however lead to different results. We can choose the one with a better performance according to the actual needs.

In § 4.4.1, we derived the sparse drag model with three degrees of freedom at $Re = 30$. Benefit from the low cost of computation for the regression test, we can iteratively run the algorithm with changing the sparsity parameter λ and investigate the performance changes of the identified model. The performance of the identified drag model by these two regression methods when varying λ is illustrated in figure 4.19(a) and figure 4.20(a), together with a comparison with the real force dynamics for three typical values of λ.

The sparsity parameter starts with 0 (pure least square regression) and increases up to nearly 1. The structures of the resulting models at $\lambda = 0.95$ for the LASSO regression, see figure 4.19(d), and $\lambda = 0.45$ for the STLS regression, see figure 4.20(c), are identical, where only a_3^2 remains. However, the STLS regression is more sensitive to the sparsity parameter, as shown in figure 4.20(a). The terms a_3, a_1^2 and a_2^2 are eliminated at the same time. The remaining a_3^2 is replaced by a_3 with $\lambda > 0.48$, and the identified models are obviously under-fitted, as shown in figure 4.20(d). In

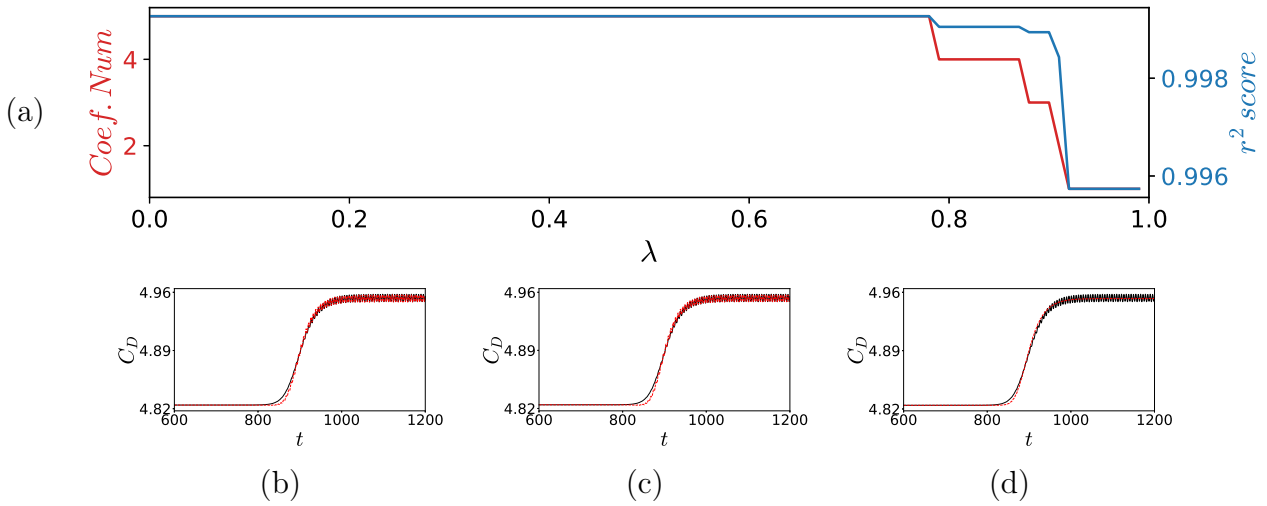


Figure 4.19: Illustration of the influence of the sparsity parameter λ on both the complexity and accuracy of the identified drag model by the LASSO regression with three degrees of freedom at $Re = 30$. (a) Evolution of the number of non-zero coefficients (red) and of the r^2 score (blue) as a function of the sparsity parameter λ. Performance of the identified drag model at $\lambda = 0.8$ (b), 0.9 (c), and 0.95 (d). Time evolution of the drag C_D coefficients in the full flow dynamics (solid black line) and for the force model (red dashed line). Initial condition: symmetric steady solution.

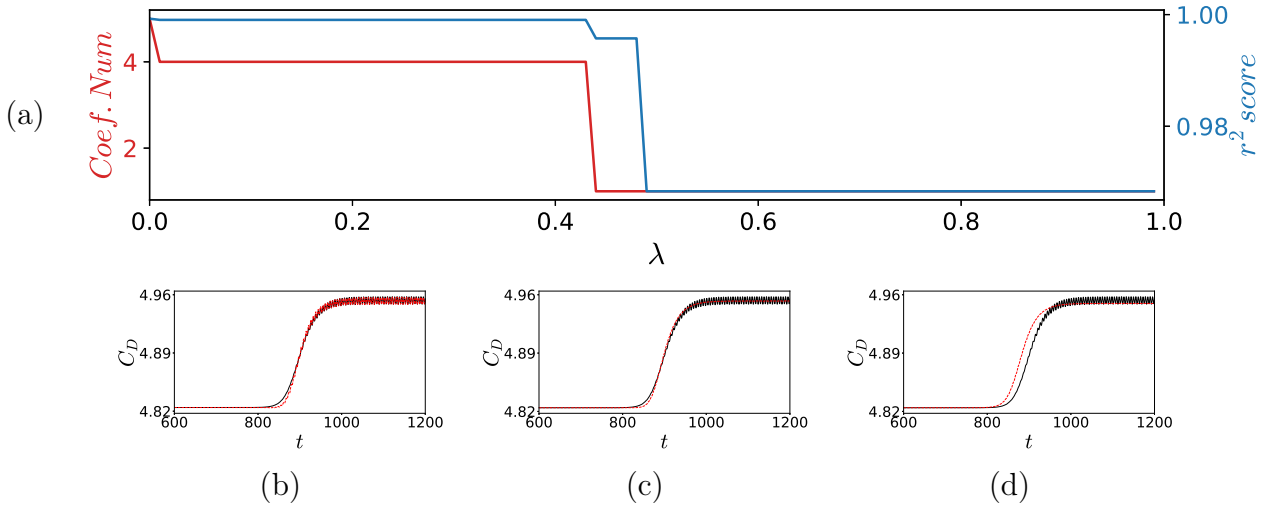


Figure 4.20: Illustration of the influence of the sparsity parameter λ on both the complexity and accuracy of the identified drag model by the sequential thresholded least square regression with three degrees of freedom at $Re = 30$. (a) Evolution of the number of non-zero coefficients (red) and of the r^2 score (blue) as a function of the sparsity parameter λ. Performance of the identified drag model at $\lambda = 0.3$ (b), 0.45 (c), and 0.9 (d). Time evolution of the drag C_D coefficients in the full flow dynamics (solid black line) and for the force model (red dashed line). Initial condition: symmetric steady solution.

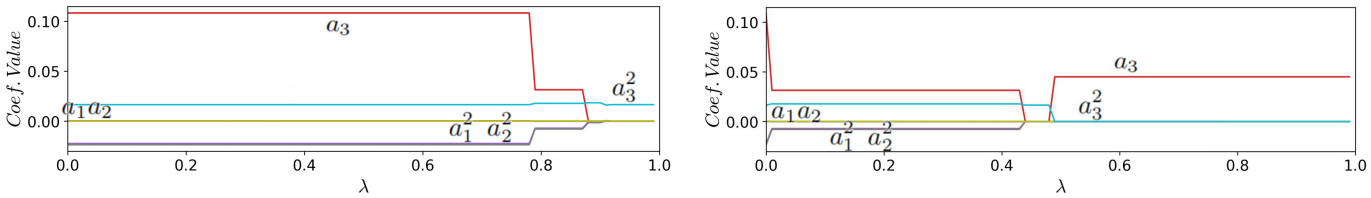


Figure 4.21: Evolution of the coefficients of the terms a_1a_2 (green), a_3 (red), a_3^2 (light blue), a_1^2 a_2^2 (purple), in the identified drag model as a function of the sparsity parameter λ for (a) the LASSO regression and (b) the sequential thresholded least square regression.

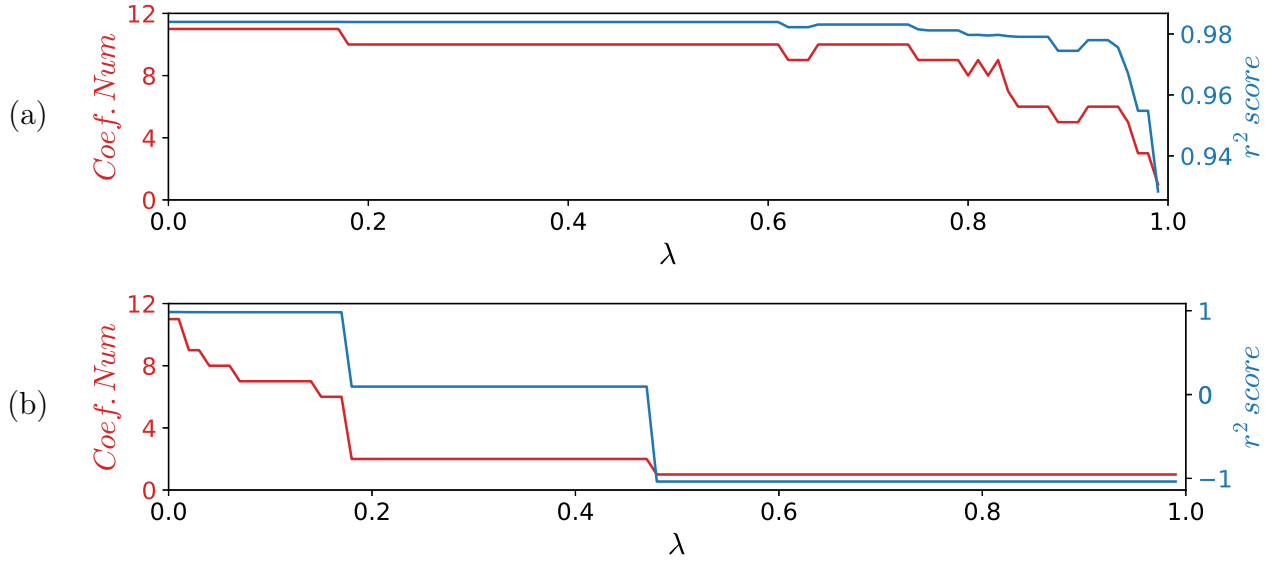


Figure 4.22: Illustration of the influence of the sparsity parameter λ on both the complexity and accuracy of the identified drag model by (a) the LASSO regression, (b) the sequential thresholded least square regression,with three degrees of freedom at $Re = 80$. Evolution of the number of non-zero coefficients (red) and of the r^2 score (blue) as a function of the sparsity parameter λ.

contrast, the LASSO regression eliminates the terms gradually, first a_1a_2, then a_3, and eventually a_1^2 together with a_2^2. In figure 4.19(a), the elimination of a_1^2 and a_2^2 only reduces the r^2 score by 0.003. But the loss of the fluctuating drag dynamics indicates an under-fitting. Hence, the optimal λ is found for 0.85.

To figure out the reason for the failure of the identification with the STLS regression when $\lambda > 0.48$, we compare the evolution of the coefficients with increasing λ in figure 4.21. a_1a_2 is the first eliminated term in both cases. The coefficients in the initial stage before the elimination of a_3 are almost the same. After the elimination of a_3 with the LASSO regression, as shown in figure 4.21(a), the coefficients of a_1^2 and a_2^2 become of order $O(10^{-3})$. Since the STLS regression algorithm thresholds the terms with smaller coefficients, the tiny coefficients of a_1^2 and a_2^2 will be set to zero simultaneously. When the STLS regression is used with a too large sparsity parameter λ, the term with larger coefficient can survive. As illustrated in figure 4.21(b), the remaining term a_3^2 is replaced by a_3 with a larger coefficient. This explains the reason why the STLS regression final converges to a_3, which is obviously the wrong term for the real drag force dynamics in figure 4.20(d).

We apply the same analysis for the sparse drag model with five degrees of freedom at $Re = 80$, as described in § 4.5.1. The evolution of the performances under the two regression methods are shown in figure 4.22. The STLS regression goes in the wrong direction as $\lambda > 0.17$. After checking the list of coefficients, the key term a_3^2 is deleted irretrievably, resulting in the inability of the

model to fit correctly. However, the regression result right before the critical value provides the most simplified and relevant drag model of Eq. (4.42a) with $r^2 = 0.9816$.

The LASSO regression is much safer on the elimination of terms. a_3^2 can survive during the regression in all the range of λ from 0 to almost 1. This further indicates that the key terms can own better robustness in the LASSO regression. From figure 4.22(a), the optimal λ is chosen at 0.85, involving six terms and $r^2 = 0.9791$. Although there are only five terms remaining when $\lambda = 0.9$, the resulting model is not stable. It returns to six terms and $r^2 = 0.9755$ at $\lambda = 0.95$, with different active terms compared to the model at $\lambda = 0.85$. At the optimal value, the identified drag model consists in terms a_5, a_1^2, a_1a_2, a_2^2, a_3^2, a_3a_5, where a_5^2 is missing. Since a_1a_2 is of order $O(10^{-4})$, we can directly apply the least square regression on the updated library with deleting a_1a_2 and adding a_5^2. The regression result is the same as for the STLS regression.

4.C Limitations of the purely projection-based approach

From the expression of pressure and viscous force on the body in § 4.2.2, the force contribution of each velocity mode in the Galerkin expansion can be numerically determined, as in Liang & Dong (2014).

The viscous force associated with mode $\boldsymbol{u}_j$ can be explicitly calculated through $q^{\nu}_{\alpha;j}$ in Eq. (4.13). However, solving $q^{p}_{\alpha;jk}$ in Eq. (4.16) needs a homogeneous Neumann boundary condition for the pressure, i.e. the normal derivative of p in the outward direction $\boldsymbol{n}$ must vanish on the whole domain boundary $\partial\Omega$,

$$\partial_n p = \boldsymbol{n} \cdot \nabla p = 0. \tag{4.58}$$

In this study, we apply a no-slip condition on velocity without the above-mentioned Neumann boundary condition on pressure. Hence, the partial pressure fields p_{jk} can not be determined to a constant pressure field. Analogously, $q^{p}_{\alpha;jk}$ can not be solved with an exact value. Even if we assume Neumann boundary conditions for the pressure field p, it is still a numerically challenging work since the pressure field are expanded to numerous partial pressure fields p_{jk}, see Eq. (4.15).

Without considering the pressure force contribution, we can reconstruct the viscous force from the viscous force contribution of the bifurcation modes. The resulting viscous force model only contains linear terms and reads

$$\begin{aligned} C_D^{\nu} &= 1.01814664 + 0.00159948\, a_3 - 0.0023798\, a_5 + 0.00601715\, a_7, \\ C_L^{\nu} &= 0.000267167\, a_1 + 0.00004522\, a_2 - 0.01409768\, a_4 - 0.0055717\, a_6. \end{aligned}$$

The viscous force contributions of each bifurcation mode is explicitly computed without any symmetry assumption, no sparsity can be expected in this model. Yet, after eliminating terms with a coefficient less than $O(10^{-5})$, the resulting force models (4.59) only involve the terms associated with the bifurcations modes with the appropriate symmetry, indicated as the symmertic modes $\boldsymbol{u}_3$, $\boldsymbol{u}_5$, $\boldsymbol{u}_7$ in C_D^{ν} and the symmertic modes $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\boldsymbol{u}_4$, $\boldsymbol{u}_6$ in C_L^{ν}. The performance of the force model using the real viscous force contribution of the seven bifurcation modes is illustrated in figure 4.23. The r^2 score for the viscous drag model is 0.9786 and 0.9183 for the viscous lift model. The accuracy and the predictive ability of the force model are acceptable for the drag model with only three items and the lift model with only four terms.

4.D Limitation of the POD-based force model

We apply POD on the fluctuating flow field $\boldsymbol{u}(\boldsymbol{x}, t) - \boldsymbol{u}_s(\boldsymbol{x})$, where $\boldsymbol{u}_s$ is the symmetric steady Navier-Stokes solution described in § 4.2.1. The snapshots used for the POD come from the two

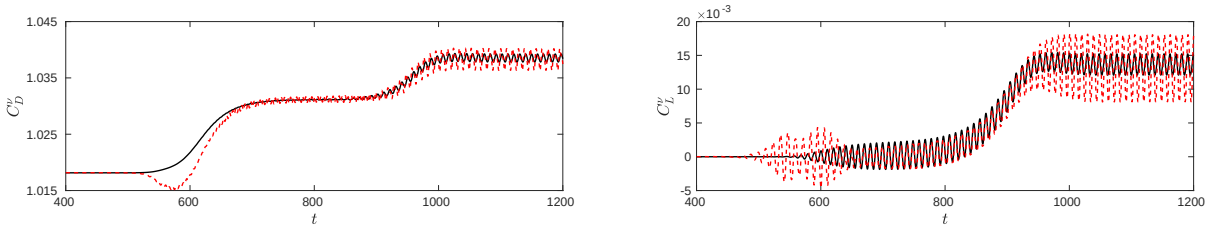


Figure 4.23: Performance of the force model with the real forces contribution of seven bifurcation modes. Time evolution of the viscous drag C_D^ν (left) and the viscous lift C_L^ν (right) coefficients in the full flow dynamics (solid black line) and for the force model (red dashed line) for DNS starting from the symmetric steady solution $\boldsymbol{u}_s$ at $Re = 80$.

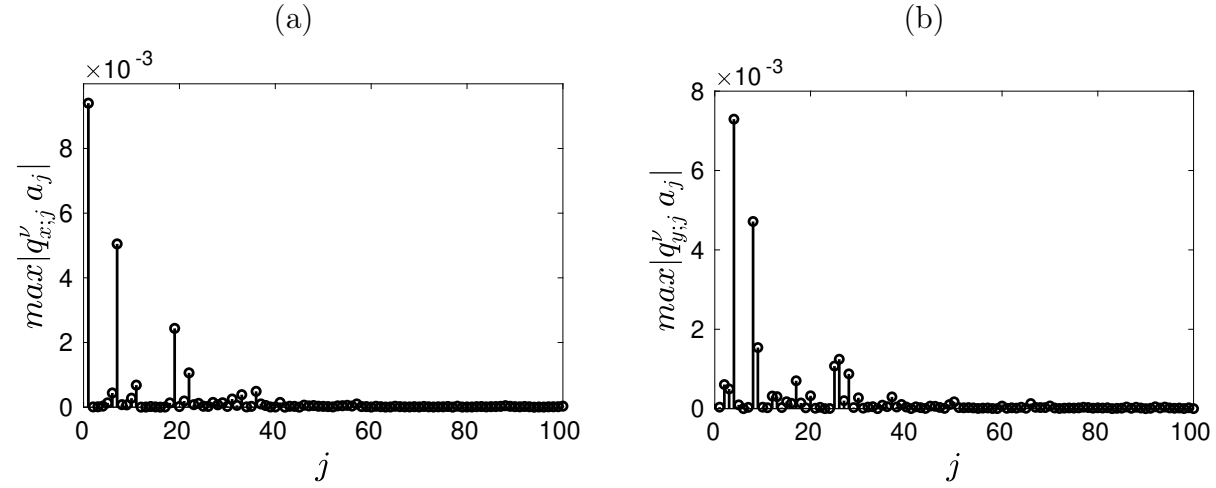


Figure 4.24: Contribution of the POD modes $\boldsymbol{u}_j$ to the viscous (a) drag and (b) lift forces for DNS starting from the symmetric steady solution $\boldsymbol{u}_s$ at $Re = 80$.

mirror-conjugated DNS trajectories started close by the symmetric steady solution. The POD mode expansion of the flow field reads:

$$\boldsymbol{u}(\boldsymbol{x},t) \quad = \quad \boldsymbol{u}_s + \sum_{j=1}^{N} a_j(t)\boldsymbol{u}_j(\boldsymbol{x}), \tag{4.60}$$

Due to the lack of boundary conditions for the pressure field contribution, we only focus on the reconstruction of the viscous force with the purely projection-based approach. The contribution to the viscous drag and lift forces, given by $\max|q_{\alpha;j}^\nu a_j|$, with $\alpha = x, y$, are shown in figure 4.24. The main force contribution comes from the leading 50 POD modes. The viscous force reconstructed with the N leading POD mode amplitudes reads

$$F_\alpha^\nu = c_\alpha^\nu + \sum_{j=1}^{N} q_{\alpha;j}^\nu a_j. \tag{4.61}$$

The viscous drag C_D^ν and lift C_D^ν coefficients reconstructed with different numbers of POD modes are compared to the real force dynamics in figure 4.25.

For a sequential N, the error of the reconstructed force coefficients with N leading POD modes can be also evaluated with the r^2 score. A higher r^2 score indicates less error in the reconstructed force. As expected, the error tends to decrease when the number of POD modes is increased. To achieve $r^2 > 0.999$, $N = 36$ leading POD modes are required for the drag force, and $N = 51$ modes for $r^2 > 0.9999$. For the lift force, these two critical numbers are respectively $N = 30$ and $N = 86$. In actual situations, the model with $r^2 > 0.999$ has enough accuracy. Note that no sparsity is

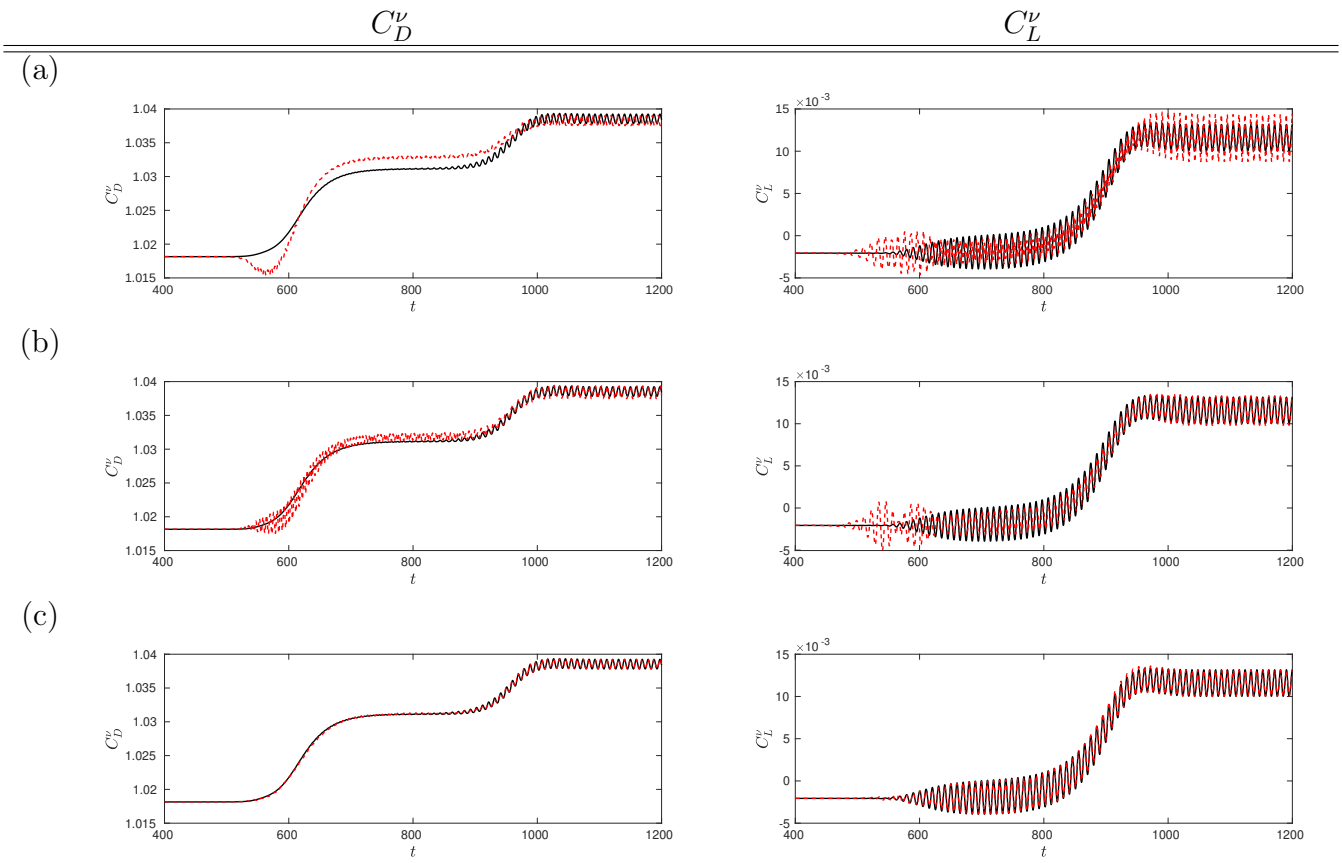


Figure 4.25: Viscous drag (left) and lift (right) force reconstruction with the (a) $N = 10$, (b) $N = 20$, (c) $N = 50$ leading POD modes starting from the symmetric steady solution $\boldsymbol{u}_s$ at $Re = 80$. Real force dynamics computed from the DNS (black curve), reconstructed forces from the N leading POD modes (dashed red line).

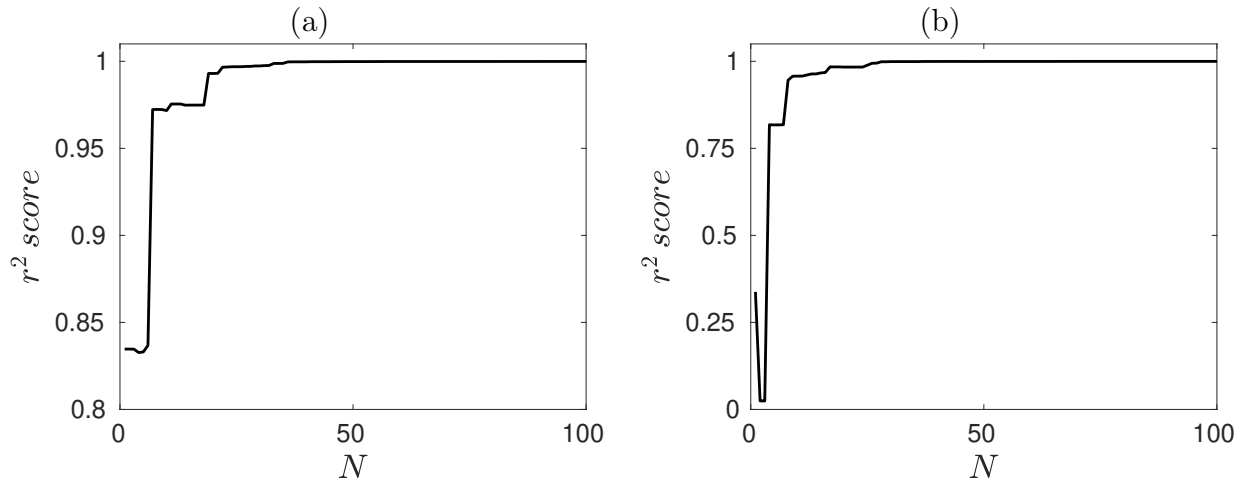


Figure 4.26: Error on the viscous (a) drag and (b) lift force reconstruction with the N leading POD modes starting the DNS starting from the symmetric steady solution $\boldsymbol{u}_s$ at $Re = 80$.

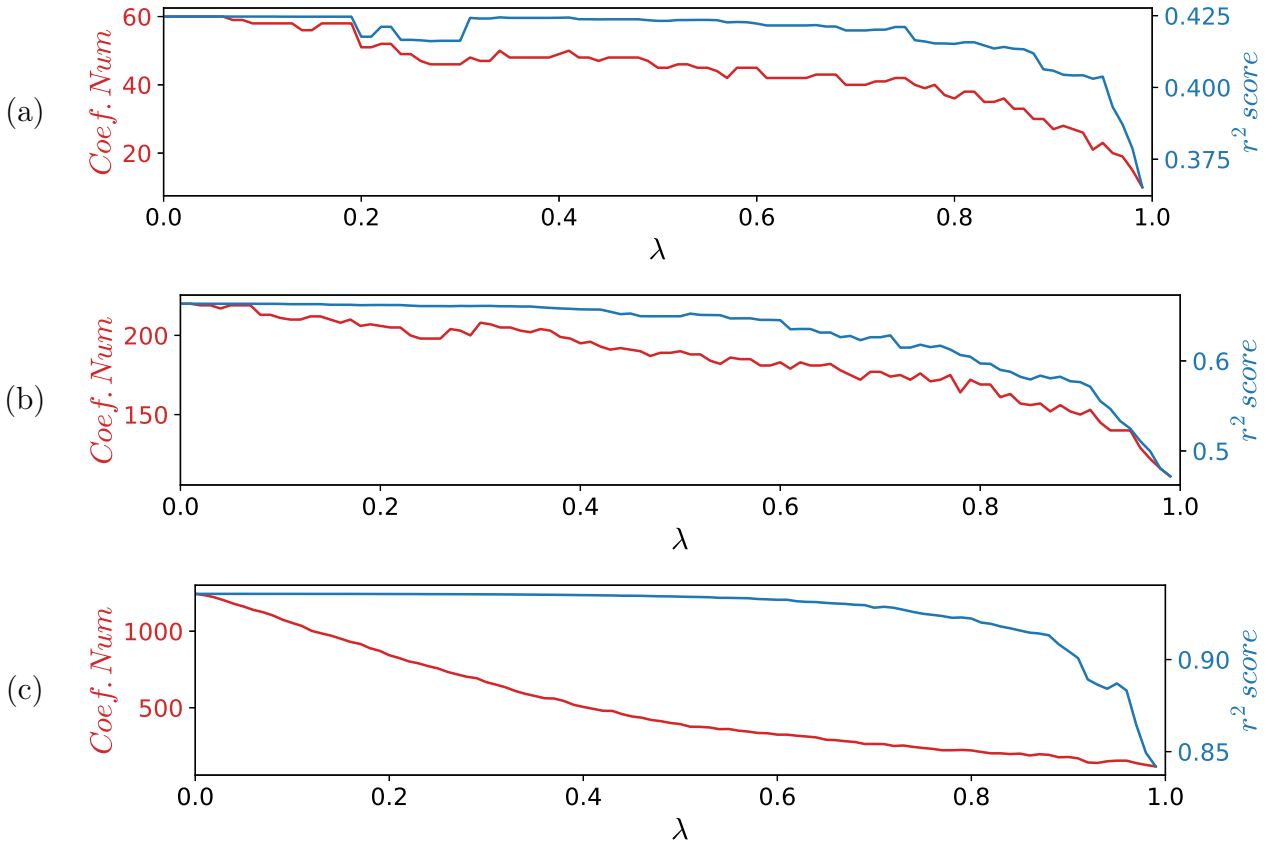


Figure 4.27: Illustration of the influence of the sparsity parameter λ on both the complexity and accuracy of the identified drag model by the LASSO regression with the (a) $N = 10$, (b) $N = 20$, (c) $N = 50$ leading POD modes at $Re = 80$. Evolution of the number of non-zero coefficients (red) and of the r^2 score (blue) as a function of the sparsity parameter λ.

involved in the model because the force contribution of each POD mode is computed explicitly.

We now focus on the regression-based approach, we set a truncation of the model with $N = 10, 20, 50$ leading POD modes, and try to use the sparse regression to find a drag model with a balance between accuracy and complexity. To be noted, the drag force considered here involves both the pressure and viscous contributions to the force. To reach the same r^2 score, it requires more POD modes due to the additional quadratic complexity of the pressure force contribution.

The library of mode amplitudes contains 66, 231 and 1326 candidate terms for the $N = 10, 20, 50$ leading POD modes. However, as shown in figure 4.27, the least square regression result for $N = 10, 20$ cannot reach a r^2 score higher than 0.7. Only the situation with $N = 50$ can start with $r^2 = 1$, but hundreds of terms are still required for an acceptable accuracy. The interpretability of the identified model is hopeless.

In summary, POD modes are decomposed and sorted according to energy criteria. The constant-linear-quadratic expression for the drag and lift forces can still be derived, but it requires a large number of POD modes. From a purely numerical approach, no sparsity is imposed in the model. For the regression-based approach, the library of sparse regression is polluted with harmonic modes and noise. Too many degrees of freedom and the harmonic relationships between them make it hard to derive a simple model from sparse regression. The most feasible solution is to find the dynamically related degrees of freedom between these modes — as we actually did in our approach. Another possible direction is to optimize the sparse regression process, in order to select the key degrees of freedom out of a polluted library of too many degrees of freedom.

4.E Reduced-order model with seven degrees of freedom

In Deng *et al.* (2020), the reduced-order model of the fluidic pinball dynamics was derived for five degrees of freedom at $Re = 80$, namely a_1 to a_5. Here we generalize the reduced-order model for seven modes, by adding a_6 and a_7 to the model. The new system reads

$$da_1/dt = a_1(\sigma_1 - \beta\ a_3 - \beta_{15}\ a_5) - a_2(\omega_1 + \gamma\ a_3 + \gamma_{15}\ a_5) + l_{14}\ a_4 + q_{134}\ a_3a_4, \tag{4.62a}$$
$$da_2/dt = a_2(\sigma_1 - \beta\ a_3 - \beta_{15}\ a_5) + a_1(\omega_1 + \gamma\ a_3 + \gamma_{15}\ a_5) + l_{24}\ a_4 + q_{234}\ a_3a_4, \tag{4.62b}$$
$$da_3/dt = \sigma_3\ a_3 + \beta_3\ r + l_{35}\ a_5 + q_{314}\ a_1a_4 + q_{335}\ a_3a_5 + q_{355}\ a_5^2, \tag{4.62c}$$
$$da_4/dt = \sigma_4\ a_4 - \beta_4\ a_4a_5 + a_1(l_{41} + q_{413}\ a_3 + q_{415}\ a_5) + a_2(l_{42} + q_{423}\ a_3 + q_{425}\ a_5), \tag{4.62d}$$
$$da_5/dt = \sigma_5\ a_5 + \beta_5\ a_4^2 + l_{53}\ a_3 + q_{514}\ a_1a_4 + q_{533}\ a_3^2 + q_{535}\ a_3a_5, \tag{4.62e}$$
$$da_6/dt = \sigma_6\ a_6 - \beta_6\ a_6a_7 + a_1(l_{61} + q_{613}\ a_3 + q_{617}\ a_7) + a_2(l_{62} + q_{623}\ a_3 + q_{627}\ a_7), \tag{4.62f}$$
$$da_7/dt = \sigma_7\ a_7 + \beta_7\ a_6^2 + l_{73}\ a_3 + q_{716}\ a_1a_6 + q_{726}\ a_2a_6 + q_{733}\ a_3^2 + q_{737}\ a_3a_7. \tag{4.62g}$$

The identified system coefficients are recorded in table 4.1, and the model performance is exemplified in figure 4.28.

σ_1	5.22×10^{-2}	β	1.31×10^{-2}	l_{14}	2.93×10^{-1}	l_{24}	-4.87×10^{-1}
ω_1	5.24×10^{-1}	γ	2.95×10^{-2}	q_{134}	-5.87×10^{-2}	q_{234}	1.18×10^{-1}
σ_3	-5.22×10^{-1}	β_3	1.53×10^{-1}	l_{41}	3.14×10^{-2}	l_{42}	-5.14×10^{-2}
σ_4	2.72×10^{-2}	β_4	5.78×10^{-2}	q_{413}	-7.56×10^{-3}	q_{423}	1.28×10^{-2}
σ_5	-2.72×10^{-1}	β_5	1.91×10^{-1}	q_{415}	2.99×10^{-2}	q_{425}	1.71×10^{-1}
		β_{15}	-2.42×10^{-2}	l_{35}	4.28	l_{53}	2.89×10^{-2}
		γ_{15}	1.70×10^{-2}	q_{335}	-1.11	q_{533}	-7.22×10^{-3}
				q_{355}	-5.13×10^{-1}	q_{535}	1.48×10^{-2}
				q_{314}	1.57×10^{-2}	q_{514}	-9.44×10^{-3}
σ_6	-7.6×10^{-2}	β_6	2.8×10^{-2}	q_{613}	-3.18×10^{-2}	q_{623}	3.23×10^{-2}
σ_7	-7.6×10^{-1}	β_7	6.27×10^{-1}	q_{617}	3.82×10^{-2}	q_{627}	-5.62×10^{-2}
		l_{61}	1.23×10^{-2}	q_{716}	-9.18×10^{-2}	q_{726}	-1.01×10^{-1}
		l_{62}	-1.33×10^{-2}	q_{733}	8.21×10^{-2}	q_{737}	1.37×10^{-1}
		l_{73}	-3.27×10^{-1}				

Table 4.1: Coefficients of the reduced-order model at $Re = 80$. See text for details.

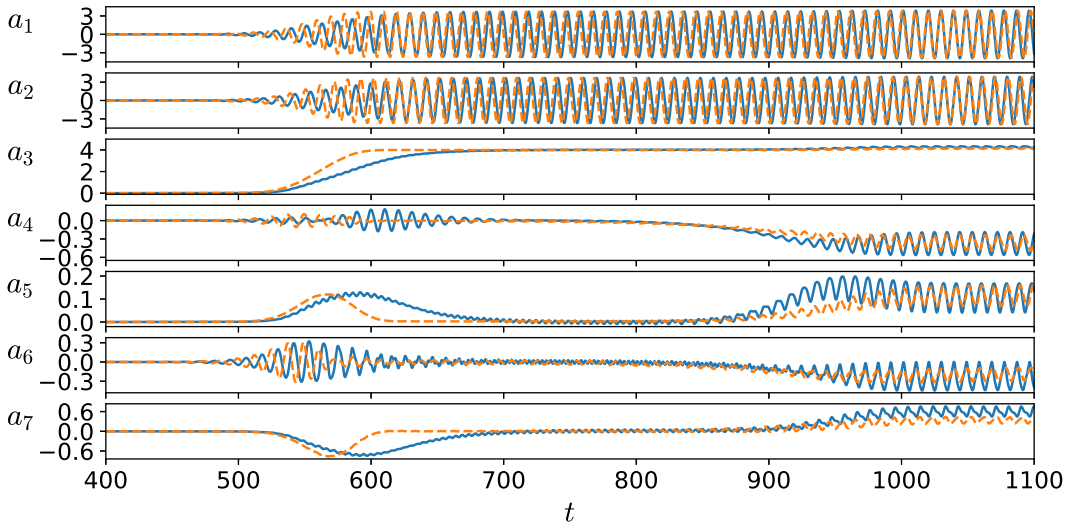


Figure 4.28: Performance of the reduced-order model with cross-terms. Time evolution of coefficients a_1 to a_7 in the full flow dynamics (solid blue line) and for the reduced-order model (red dashed line). The initial condition is the same for the reduced-order model and the full flow dynamics.

Chapter 5

Cluster-based hierarchical network model of the fluidic pinball — Cartographing transient and post-transient, multi-frequency, multi-attractor behaviour

Nan Deng[1,2] , Bernd R. Noack[3,4], Marek Morzyński[5] and Luc R. Pastur[1]

[1] Institute of Mechanical Sciences and Industrial Applications, ENSTA-Paris, Institut Polytechnique de Paris, 828 Bd des Maréchaux, F-91120 Palaiseau, France

[2] Université Paris-Saclay, CNRS, Laboratoire Interdisciplinaire des Sciences du Numérique, F-91400 Orsay, France

[3] Center for Artificial Intelligence and Aerodynamics, Harbin Institute of Technology, Shenzhen 518055, People's Republic of China

[4] Institut für Strömungsmechanik und Technische Akustik (ISTA), Technische Universität Berlin, Müller-Breslau-Straße 8, D-10623 Berlin, Germany

[5] Department of Virtual Engineering, Poznań University of Technology, Jana Pawla II 24, PL 60-965 Poznań, Poland

Submitted to *Journal of Fluid Mechanics*

We propose a self-supervised cluster-based hierarchical reduced-order modelling methodology to model and analyse the complex dynamics arising from a sequence of bifurcations for a two-dimensional incompressible flow of the unforced fluidic pinball. The hierarchy is guided by a triple decomposition separating a slowly varying base flow, dominant shedding and secondary flow structures. All these flow components are kinematically resolved by a hierarchy of clusters, starting with the base flow in the first layer, resolving the vortex shedding in the second layer and distilling the secondary flow structures in the third layer. The transition dynamics between these clusters is described by a directed network, called cluster-based hierarchical network model (HiCNM) in the sequel. Three consecutive Reynolds number regimes for different dynamics are considered: (i) periodic shedding at $Re = 80$, (ii) quasi-periodic shedding at $Re = 105$, and (iii) chaotic shedding at $Re = 130$, involving three unstable fixed points, three limit cycles, two quasi-periodic attractors and a chaotic attractor. The HiCNM enables identifying the transient and post-transient dynamics between multiple invariant sets in a self-supervised manner. Both the global trends and the local structures during the transition are well resolved by a moderate number of hierarchical clusters. The proposed reduced-order modelling provides a visual representation of transient and post-transient, multi-frequency, multi-attractor behaviour and may automate the identification and analysis of complex dynamics with multiple scales and multiple invariant sets.

Contents

5.1 Introduction

Fluid flows generally involve complex, high-dimensional and nonlinear dynamics, which makes them hard to understand. However, even at high Reynolds numbers, the flow dynamics keeps trace of the instabilities undergone at increasing Reynolds number (Huerre & Monkewitz, 1990). Stationary laminar flows are generally stable with respect to infinitesimal perturbations at sufficiently low Reynolds number. This steady state becomes unstable when the Reynolds number increases beyond a critical value Re_c, where a bifurcation occurs. On the way towards a fully turbulent regime, the flow may undergo a succession of bifurcations with increasing Reynolds number. Ruelle & Takens (1971) shows that the flow can reach a chaotic regime after a small number of bifurcations. The complex flow dynamics can be seen as the result of the interactions between the fundamental structures of different instabilities (Chomaz, 2005; Bagheri *et al.*, 2009*a*). A reduced-order model incorporating the underlying mechanisms is always the promising solution for flow analysis (Amsallem & Farhat, 2008; LeGresley & Alonso, 2000) and control (Choi *et al.*, 2008; Bagheri *et al.*, 2009*b*; Barbagallo *et al.*, 2009).

Numerous reduced-order models (ROMs) have been developed and applied (Taira *et al.*, 2017). The classical method starts with projecting the full system into a low-dimensional subspace, where the high-dimensional dynamics can be approximated with the optimal basis. This process is so-called Galerkin projection, which leads to a Galerkin system describing the dynamics in reduced-order ordinary differential equations (ODEs). According to the dimensionality reduction techniques and the model selection strategies, there exist many different projection-based ROMs. Proper orthogonal decomposition (POD) (Berkooz *et al.*, 1993; Holmes *et al.*, 2012*b*) is the most popular one, which has many empirical variations, for example, balanced POD (Rowley, 2005) with balanced truncation. The POD-Galerkin method can be optimized and extended with incorporating

the pressure term (Bergmann *et al.*, 2009), with numerical stabilizationis (Iollo *et al.*, 2000), with variational multiscale method (Iliescu & Wang, 2014) and with closure modelling strategies (Wang *et al.*, 2012). Based on first principles, the mean-field theory of Landau (1944) and Stuart (1958) is the lowest dimensional mean-field model to account for a supercritical Hopf bifurcation. Weakly nonlinear mean-field analysis has also been applied to more complex situations in which the flow has undergone two successive bifurcations, such as in the wake of axisymmetric bodies (Fabre *et al.*, 2008), the wake of a disk (Meliga *et al.*, 2009) or the wake of the fluidic pinball (Deng *et al.*, 2020). Gomez *et al.* (2016); Rigas *et al.* (2017*b*) included mean-field considerations in their resolvent analysis, decomposing the flow in time-resolved linear dynamics and a feedback term with the quadratic nonlinearity.

Alternatively, data-driven strategies show their advantage in pattern and system recognition without prior knowledge about flow dynamics (Brunton *et al.*, 2020), like Koopman analysis (Schmid, 2010; Mezić, 2013) using dynamic mode decomposition (DMD) (Tu *et al.*, 2014; Kutz *et al.*, 2016), data-driven Galerkin modelling (Noack *et al.*, 2016) using recursive DMD, and multi-scale Proper Orthogonal Decomposition (mPOD) (Mendez *et al.*, 2019) using a matrix factorization framework to enhance feature detection capabilities. Above mentioned methods still start with a modal decomposition of the original flow fields. The advances in machine-learning algorithms provide huge potential for data-driven ROMs, for example, using artificial neural network (ANN) to stabilize projection-based ROMs (San & Maulik, 2018) or to build the ANN ROMs (San *et al.*, 2019), turbulence modelling with deep neural networks (Kutz, 2017), feature-based manifold modelling (Loiseau *et al.*, 2018*b*) with sparse identification (Brunton *et al.*, 2016*b*).

Inspired with centroidal Voronoi tessellation ROMs in Burkardt *et al.* (2006), Kaiser *et al.* (2014) proposed the cluster-based reduced-order modelling (CROM) method to partition the flow data into clusters and analyze the flow dynamics with a cluster-based Markov model (CMM). CROM provides us with a novel modelling strategy, liberating us from the issue of choosing a low-dimensional space of the traditional projection method. Nair *et al.* (2019) applied CROM to the nonlinear feedback flow control and introduced the directed network (Newman, 2018) for the dynamical modelling. With the clusters being the nodes and the transitions between clusters being the edges, an extended Markov model with a directed network was built, emphasizing the non-trivial transitions between clusters. Fernex *et al.* (2021) and Li *et al.* (2021) further proposed the cluster-based network model (CNM) for time-resolved data by introducing local interpolations between clusters with the pre-specified transition times. The CNM can be seen as an extension of the traditional CMM, using the network model instead of the standard Markov model to describe the transient dynamics. Networks of complex dynamical systems have attracted a great deal of interest, forming an increasingly important interdisciplinary field known as network science (Watts & Strogatz, 1998; Albert & Barabási, 2002; Barabási, 2013). The network-based approaches have been used in fluid mechanics to describe the interactions among vortical elements (Nair & Taira, 2015), detect the Lagrangian vortex motion (Hadjighasem *et al.*, 2016), and model and analyze turbulent flows (Taira *et al.*, 2016; Yeh *et al.*, 2021). Together with the clustering approaches, networks have been also used to extract key features of complex flows (Bollt, 2001; Schlueter-Kuck & Dabiri, 2017; Murayama *et al.*, 2018; Krueger *et al.*, 2019). The critical structures modifying the flow can be identified by the intra- and inter-cluster interactions using community detection (Gopalakrishnan Meena *et al.*, 2018; Gopalakrishnan Meena & Taira, 2021). Theories and techniques in the field of network science may play a crucial role in the modelling, analysis and control of fluid systems.

The accuracy of the cluster-based model depends on the number of clusters. However, too many clusters will increase the complexity of the Markov/network model. A high level of human experience is required to achieve a good compromise between resolution and a simple model. The focus of this paper is to optimise the data-driven cluster analysis by introducing a hierarchical structure and a systematic self-supervised way to model the transient and post-transient flows in the case of multiple unstable solutions and multiple attractors. The hierarchical modelling strategy

shows good consistency with the Reynolds decomposition from the mathematical foundation. The systematic data treatment process shows its great potential for multiscale and multi-frequency modelling.

Inspired by the hierarchical Markov model (Fine *et al.*, 1998), we apply a scale-dependent hierarchical clustering to the classic network modelling under the mean-field consideration. The time-scales of the different flow components provide a good indicator for figuring out the typical structures in multiscale flows, and enable the hierarchical model to address the complex dynamics of multiscale problems. The resulting cluster-based hierarchical network model (HiCNM) can systematically identify complex dynamics involved in the case of multiple attractors. Both the global trends and the local structure during the transition can be well preserved by a fewer number of clusters in the hierarchical structure, which leads to a better understanding of the physical mechanisms involved in the flow dynamics.

We consider the two-dimensional incompressible flow configuration of Bansal & Yarusevych (2017), defined as the (unforced) "fluidic pinball" in Deng *et al.* (2020). With increasing Reynolds number, the wake undergoes a first instability leading to a periodic vortex shedding, then a static symmetry breaking, and finally a transition to a quasi-periodic regime before transiting to a chaotic regime. HiCNMs are built for these flow regimes, which have multiple invariant sets and exhibit different transient dynamics. We provide a principle sketch of our HiCNM framework in figure 5.1.

The manuscript is organised as follows: § 5.2 describes the numerical plant of the fluidic pinball and the flow features at different Reynolds number. § 5.3 discusses the different perspectives on the cluster-based hierarchical network modelling strategy. In § 5.4, we discuss the HiCNMs applied to the transient and post-transient dynamics of a flow configuration involving six invariant sets, for three different Reynolds numbers, respectively associated with a periodic, a quasi-periodic and a chaotic dynamics. § 5.5 summarises the main findings and gives some suggestions for improvement and future directions.

5.2 Flow configuration and flow features

We consider two-dimensional incompressible flows in the fluidic pinball (Noack & Morzyński, 2017) as the benchmark configuration for our hierarchical modelling strategy. The flow configuration and the direct Navier-Stokes solver are described in § 5.2.1. The transient and post-transient dynamics at different Reynolds numbers are illustrated in § 5.2.2.

5.2.1 Flow configuration and direct Navier-Stokes solver

Figure 5.2 shows the geometric configuration of the fluidic pinball, consisting of three fixed cylinders of unit diameter D. Their axes are placed on the vertices of an equilateral triangle of side $3D/2$ in the (x, y) plane. The upstream flow is in the x-axis direction with a uniform velocity U_∞ at the inlet of the domain. The computational domain Ω is bounded by a rectangular box of size $[-6D, +20D] \times [-6D, +6D]$. A Cartesian coordinate system is used for description, and its origin is placed in the middle of the back two cylinders considering the symmetry of this configuration. Since no external force is applied to these three cylinders, a no-slip condition is applied on the cylinders, and the velocity in the far wake is assumed to be U_∞. The Reynolds number is defined as $Re = U_\infty D/\nu$, where ν is the kinematic viscosity of the fluid. A no-stress condition is applied at the outlet of the domain.

The fluid flow is governed by the non-dimensionalized incompressible Navier-Stokes equations in scales with the cylinder diameter D and the velocity U_∞, which read

$$\partial_t \boldsymbol{u} + \nabla \cdot \boldsymbol{u} \otimes \boldsymbol{u} = \nu \triangle \boldsymbol{u} - \nabla p, \quad \nabla \cdot \boldsymbol{u} = 0, \tag{5.1}$$

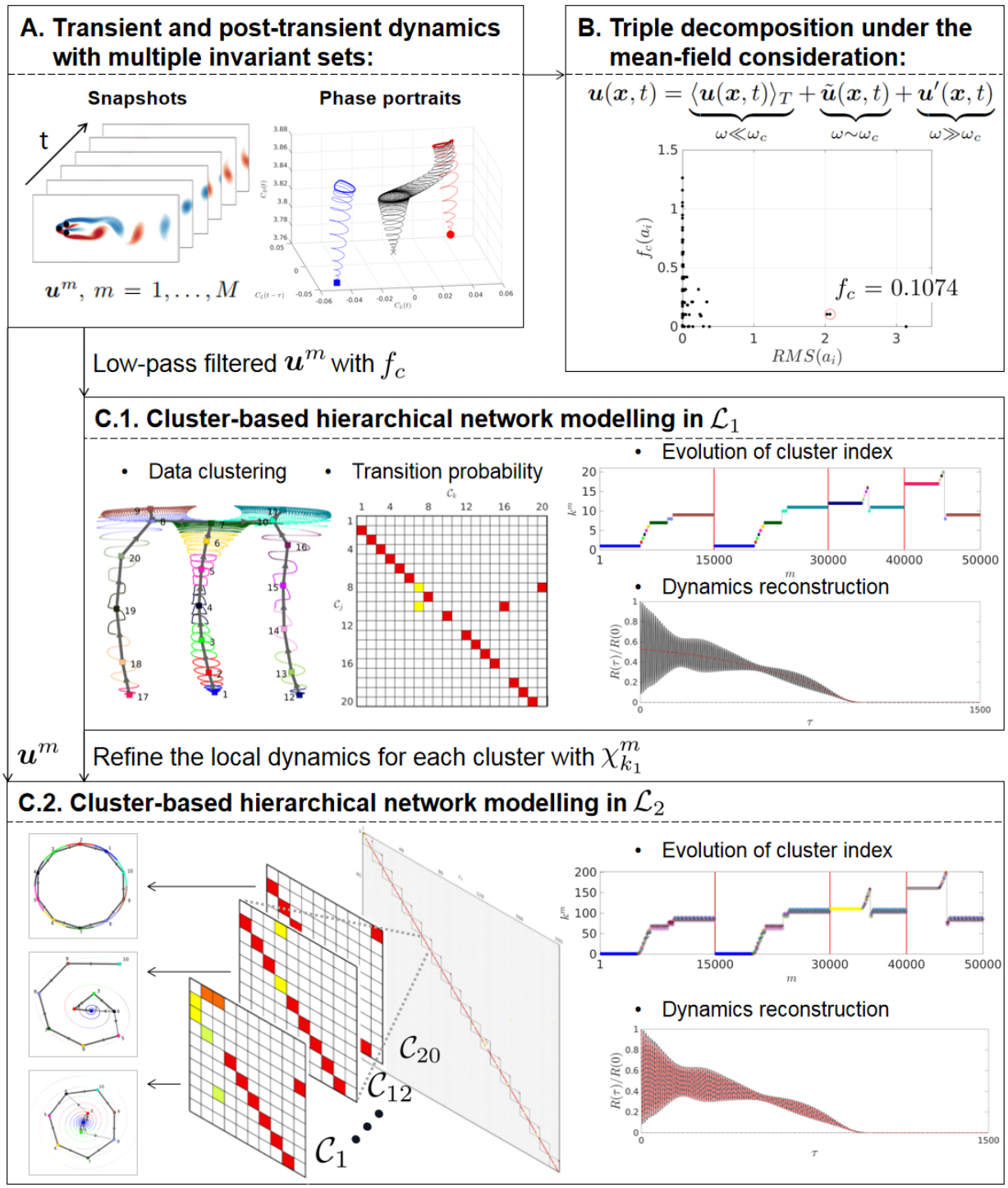


Figure 5.1: Overview of the cluster-based hierarchical network modelling framework exemplified at $Re = 80$. (A) The flow dynamics involves six invariant sets associated with three unstable fixed points, three limit cycles, as shown in the 3D phase portrait of the drag and lift forces. (B) Under the mean-field consideration, the flow can be decomposed into a slowly-varying mean flow, the coherent and incoherent components, separated by the dominant frequency of the coherent part. The non-coherent fluctuating component is weak in this case, and the third term of the triple decomposition can be ignored. (C) Therefore, a HiCNM with two layers is enough to extract the global trend and the local dynamics of the varying mean-flow field. The transient and post-transient dynamics, characterized by multiple frequencies and multiple invariant sets, are introduced in § 5.2.2. The hierarchical network modelling strategy is discussed in 5.3.2 under the mean-filed consideration in § 5.3.1. The dynamics reconstruction of the resulting hierarchical network model is given in § 5.3.2.

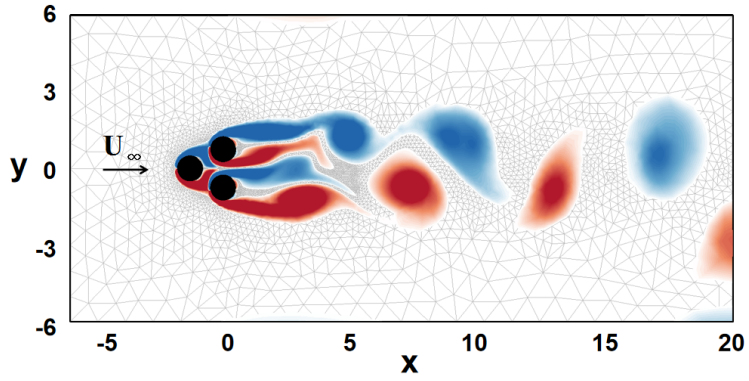


Figure 5.2: Configuration of the fluidic pinball and computational grid for the simulated domain. The upstream velocity is denoted U_∞. An example vorticity field at $Re = 130$ is colour-coded in the range $[-1.5, 1.5]$ from blue to red.

where p and $\boldsymbol{u}$ are respectively the pressure and velocity flow fields and $\nu = 1/Re$. The advection time scale is D/U_∞ and the pressure scale is ρU_∞^2, where ρ is the unit fluid density for the incompressible flow. It is assumed that there exists a solution $(\boldsymbol{u}_s, p_s)$ satisfying the steady Navier-Stokes equations

$$\nabla \cdot \boldsymbol{u}_s \otimes \boldsymbol{u}_s = \nu \triangle \boldsymbol{u}_s - \nabla p_s, \quad \nabla \cdot \boldsymbol{u}_s = 0. \tag{5.2}$$

The inner product of two square-integrable velocity fields $\boldsymbol{u}(\boldsymbol{x})$ and $\boldsymbol{v}(\boldsymbol{x})$ in the computational domain Ω reads

$$(\boldsymbol{u}, \boldsymbol{v})_\Omega := \int_\Omega d\boldsymbol{x}\, \boldsymbol{u}(\boldsymbol{x}) \cdot \boldsymbol{v}(\boldsymbol{x}). \tag{5.3}$$

The associated norm of the velocity field $\boldsymbol{u}(\boldsymbol{x})$ is defined as

$$||\boldsymbol{u}||_\Omega = \sqrt{(\boldsymbol{u}, \boldsymbol{u})_\Omega}. \tag{5.4}$$

The direct numerical simulation (DNS) of the Navier-Stokes equations (5.1) is based on a second-order finite-element discretization method of the Taylor-Hood type (Taylor & Hood, 1973), on an unstructured grid of 4 225 triangles and 8 633 vertices, and an implicit integration of the third-order in time. The unsteady flow field is calculated by an unsteady solver with Newton-Raphson iteration until the residual is less than a prescribed tolerance. This approach is also employed to calculate the steady solution by a steady solver for the steady Navier-Stokes equations (5.2). The direct Navier-Stokes solver used herein has been validated in Noack *et al.* (2003); Deng *et al.* (2020), and the grid used for the simulations provides a consistent flow dynamics comparing to a refined grid. A relevant numerical investigation for this kind of equilateral-triangle configuration can also be found in Chen *et al.* (2020). The data-driven HiCNM method is exemplified on this benchmark configuration with a blockage ratio $B = 0.21$, defined as the ratio of the cross-section length of the cluster of three cylinders $5D/2$ to the width of the computational domain $12D$, which mimics the experimental setups in Raibaudo *et al.* (2020). The numerical results with the current computational domain remain similar compared with a larger domain with $B = 0.025$, as detailed in appendix 5.A.

5.2.2 Flow features

As shown in figure 5.3, the flow undergoes a supercritical Hopf bifurcation at $Re_1 \approx 18$, a supercritical pitchfork bifurcation at $Re_2 \approx 68$ and a Neimark-Säcker bifurcation at $Re_3 \approx 105$, before entering the chaotic regime beyond $Re_4 \approx 115$ with increasing the Reynolds number (Deng *et al.*, 2020). Depending on the Reynolds number, the wake flow may present rich transient dynamics

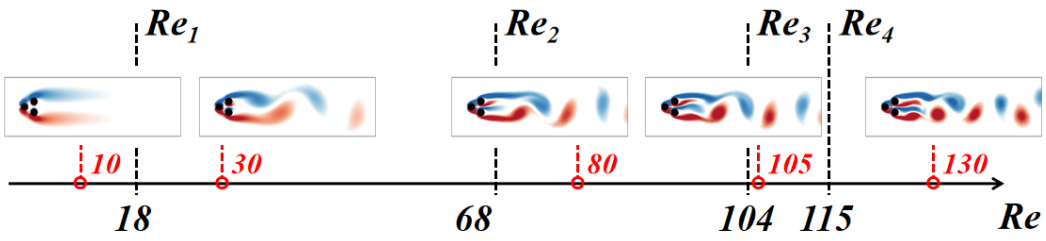


Figure 5.3: Post-transient flow state for different flow regimes at the Reynolds numbers marked in red. The critical values of the supercritical Hopf bifurcation Re_1, the supercritical pitchfork bifurcation Re_2, and the Neimark-Säcker bifurcation Re_3 before the system entering into chaos at Re_4 are marked in black on the Re-axis.

due to multiple exact solutions of the Navier-Stokes equations, assciated with the co-existing invariant sets in the state space. For instance, at $Re_1 < Re < Re_2$, the symmetric steady solution $\boldsymbol{u}_s$ is the only fixed point of the system. This exact solution of the Navier-Stokes equations is unstable. The only attractor in the state space is a symmetric limit cycle, associated with the cyclic release of vortices in the wake of the cylinders, forming a von Kármán street of regular vortices. At $Re > Re_2$, three fixed points are solutions of the steady Navier-Stokes equations, one symmetric $\boldsymbol{u}_s$ and two asymmetric steady solutions $\boldsymbol{u}_s^{\pm}$, and all three points are unstable. Meanwhile, the unsteady Navier-Stokes equations have three periodic solutions. The symmetric limit cycle, associated with symmetric vortex shedding, is unstable. The two mirror-conjugated asymmetric limit cycles, associated with asymmetric vortex sheddings, co-exist as attractors of the flow dynamics in the state space. For $Re_3 < Re < Re_4$, the two attracting asymmetric limit cycles thicken into torii by introducing an additional low frequency, which modulates the vortex shedding quasi-periodically. Beyond Re_4, the vortex shedding dynamics is chaotic. The interested reader can find more details on the route to chaos in the fluidic pinball in Deng *et al.* (2020).

The flow features can be illustrated by the forces exerted on the body. The drag F_D and lift F_L forces are the projection on $\boldsymbol{e}_x$ and $\boldsymbol{e}_y$ of the resultant force $\boldsymbol{F} = F_D\,\boldsymbol{e}_x + F_L\,\boldsymbol{e}_y$, obtained by integrating the viscous and pressure forces over the cylinder surfaces. The flow dynamics is analyzed with the lift coefficient C_L,

$$C_L(t) = \frac{2F_L(t)}{\rho\, U_\infty^2}. \tag{5.5}$$

We apply the DNSs at $Re = 30$, 80, 105 and 130 respectively, starting close to the symmetric steady solution (for $Re > Re_1$) until $t = 1\,500$ and the asymmetric steady solutions (for $Re > Re_2$) until $t = 1\,000$. The time evolutions of the lift coefficient C_L are shown in figure 5.4, where different transient dynamics are observed, from the steady solutions to the asymptotic regimes.

At $Re = 30$, as shown in figure 5.4(a), the lift coefficient C_L starts to oscillate visibly at $t \approx 800$, indicating that the flow leaves the neighborhood of the symmetric steady solution. Then, C_L oscillates around a vanishing value with increasing amplitude until converging to a fixed amplitude. This state refers to a symmetric vortex shedding, as the instantaneous flow is oscillating around a geometrical symmetric mean-flow field.

At $Re = 80$, as shown in figure 5.4(b), the primary transition is the same as at $Re = 30$. Next, the slowly-varying mean lift coefficient $\langle C_L \rangle_T$, averaged over the oscillation period T, leaves from 0 to 0.04. This indicates that the oscillatory dynamics in the permanent regime has lost the statistical symmetry, and the flow state refers to an asymmetric vortex shedding. Starting nearby either one of the two asymmetric steady solutions, C_L directly evolves to the asymmetric vortex shedding regime thats shares the same asymmetry.

At $Re = 105$, C_L visibly increases at $t \approx 580$ starting with the vanishing value of the symmetric steady solution, as illustrated with the black curve in figure 5.4(c). However, the initial transition

reaches a non-oscillating value equal to the initial value of the red curve, which refers to one of the two asymmetric steady solutions. It eventually enters a quasi-periodic state, the vortex shedding oscillations being modulated at a low frequency. Starting from the other two asymmetric steady solutions will directly evolve into the permanent quasi-periodic state with the same asymmetry.

At $Re = 130$, the initial transition of the black curve in figure 5.4(d) is similar to the initial transition at $Re = 80$, but the dynamics enters a chaotic regime shortly after the symmetric vortex shedding has started. Simulations converge to the same chaotic attracting set, starting with all the three different steady solutions.

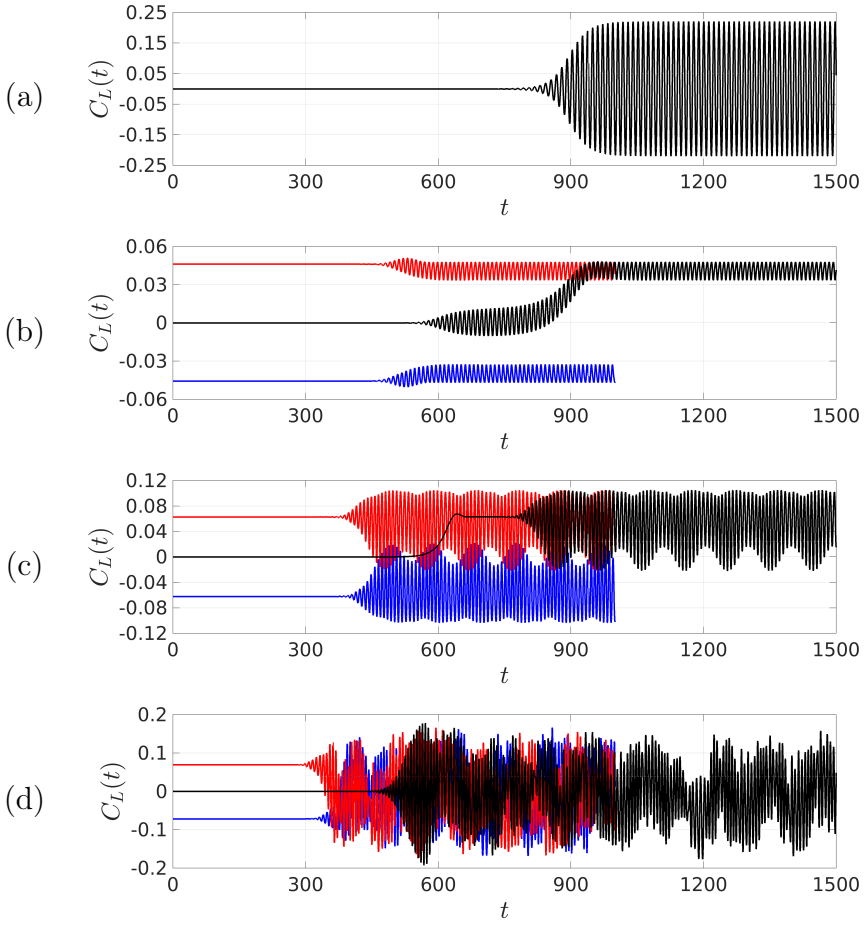


Figure 5.4: Transient and post-transient dynamics starting with different steady solutions, illustrated with the time evolution of the lift coefficient C_L at $Re = 30$ (a), 80 (b), 105 (c), 130 (d).

5.3 Cluster-based hierarchical reduced-order modelling

In this section, the general approach of the cluster-based hierarchical reduced-order modelling is described and discussed. In § 5.3.1, we present the relevant background on the flow decomposition and the standard cluster-based reduced-order model (CROM). The cluster-based reduced-order modelling with hierarchical structure is described in § 5.3.2, as well as the relevant analysis of the cluster-based hierarchical network model (HiCNM).

5.3.1 Background

The standard CROM is obtained in two steps: the snapshots are first clustered into coarse-grained representative states before building either a Markov or a network model for the analysis of the

dynamics. Clustering all transients and post-transients at once can suffer from the inability to accurately capture the dynamics at different scales. Under the mean-field consideration, we introduce a hierarchical structure of clusters for the flow dynamics of different time scales, resulting in a cluster-based hierarchical reduced-order model (HiCROM).

Flow decomposition with mean-field consideration

The starting point of the HiCROM is the triple decomposition of the flow field similar to Reynolds & Hussain (1972)

$$\boldsymbol{u}(\boldsymbol{x},t) = \underbrace{\langle \boldsymbol{u}(\boldsymbol{x},t)\rangle_T}_{\omega \ll \omega_c} + \underbrace{\tilde{\boldsymbol{u}}(\boldsymbol{x},t)}_{\omega \sim \omega_c} + \underbrace{\boldsymbol{u}'(\boldsymbol{x},t)}_{\omega \gg \omega_c}, \tag{5.6}$$

where the dominant angular frequency ω_c is defined as the dominant peak in the Fourier spectrum of the velocity field. Here, the velocity field is decomposed into a slowly-varying mean-flow field $\langle \boldsymbol{u}\rangle_T$, a coherent component on time-scales of order $2\pi/\omega_c$, involving coherent structures $\tilde{\boldsymbol{u}}$, and the remaining non-coherent small scale fluctuations $\boldsymbol{u}'$. This kind of decomposition can also be found in the low-order Galerkin models of Tadmor *et al.* (2011) and the weakly nonlinear modelling of Rigas *et al.* (2017*a*).

The slowly-varying mean-flow field $\langle \boldsymbol{u}\rangle_T$ can be defined as the average of the velocity field $\boldsymbol{u}$ over one local period $T \approx 2\pi/\omega_c$ of the coherent structures,

$$\langle \boldsymbol{u}(\boldsymbol{x},t)\rangle_T := \frac{1}{T}\int_{t-T/2}^{t+T/2} d\tau\, \boldsymbol{u}(\boldsymbol{x},\tau), \tag{5.7}$$

which eliminates both the coherent contribution from $\tilde{\boldsymbol{u}}$ and the non-coherent contribution from $\boldsymbol{u}'$. Unlike the mean-flow field defined by the post-transient limit,

$$\bar{\boldsymbol{u}}(\boldsymbol{x}) = \lim_{T\to\infty} \frac{1}{T}\int_0^T \boldsymbol{u}(\boldsymbol{x},\tau)d\tau, \tag{5.8}$$

the finite-time averaged-flow field considered in this study owns a slowly varying dynamics. From the mean-field theory of Stuart (1958), the slowly-varying mean-flow field evolves out of the steady solution under the action of the Reynolds stress associated with the most unstable eigenmode(s). The mean-flow field deformation $\boldsymbol{u}_\Delta$ is used to describe the difference between the slowly-varying mean-flow field and the invariant steady solution $\boldsymbol{u}_s(\boldsymbol{x})$, which reads

$$\langle \boldsymbol{u}(\boldsymbol{x},t)\rangle_T = \boldsymbol{u}_s(\boldsymbol{x}) + \boldsymbol{u}_\Delta(\boldsymbol{x},t). \tag{5.9}$$

Clustering algorithm

We consider the state vectors, for instance, the velocity fields $\boldsymbol{u}(\boldsymbol{x},t)$ in the computational domain Ω, which is sampled at times $t^m = m\Delta t$ with a time step Δt, where the superscript $m = 1,\ldots,M$ is the snapshot index. The clustering process aims at partitioning the M time-discrete states (snapshots) $\boldsymbol{u}^m = \boldsymbol{u}(\boldsymbol{x},t^m)$ into K clusters $\mathcal{C}_k$, $k = 1,\ldots,K$. Snapshots of a given cluster share similar attributes featured by its cluster centroid $\boldsymbol{c}_k$. The distance between the snapshot $\boldsymbol{u}^m$ and the centroid $\boldsymbol{c}_k$ is defined as

$$D_k^m := ||\boldsymbol{u}^m - \boldsymbol{c}_k||_\Omega. \tag{5.10}$$

Each snapshot is partitioned to the cluster of the closest centroid by $\underset{k}{\operatorname{argmin}}\, D_k^m$, and the characteristic function is defined as

$$\chi_k^m := \begin{cases} 1, & \text{if } \boldsymbol{u}^m \in \mathcal{C}_k, \\ 0, & \text{otherwise.} \end{cases} \tag{5.11}$$

A cluster index k^m, $m = 1, \ldots, M$, indicates the cluster assignment of the corresponding snapshot with $\boldsymbol{u}^m \in \mathcal{C}_k$, and records the visited clusters consecutively. The number of snapshots n_k in cluster k is given by

$$n_k := \sum_{m=1}^{M} \chi_k^m. \tag{5.12}$$

The cluster centroids $\boldsymbol{c}_k$ are defined as the average of the snapshots belonging to the cluster $\mathcal{C}_k$:

$$\boldsymbol{c}_k = \frac{1}{n_k} \sum_{m=1}^{M} \chi_k^m \boldsymbol{u}^m. \tag{5.13}$$

The performance of clustering is judged by the within-cluster variances:

$$J\left(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_K\right) = \sum_{k=1}^{K} \sum_{m=1}^{M} \chi_k^m \left\|\boldsymbol{u}^m - \boldsymbol{c}_k\right\|_\Omega^2. \tag{5.14}$$

The clustering algorithm minimizes J and determines the optimal centroid positions,

$$\boldsymbol{c}_1^{\text{opt}}, \ldots, \boldsymbol{c}_K^{\text{opt}} = \underset{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_K}{\operatorname{argmin}} J\left(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_K\right), \tag{5.15}$$

by iteratively updating the characteristic function and the centroid positions.

To solve the optimization problem (5.15), we use the k-means++ algorithm (Arthur & Vassilvitskii, 2006). Comparing to the traditional k-means algorithm, the k-means++ algorithm selects the initial centroids as far away as possible to avoid any bias from the initial conditions. The remaining steps of the two algorithms are the same. At each iteration, the snapshots are divided into clusters of the nearest newly determined centroids. The optimal centroids are obtained by iterating until either convergence or when the maximum number of iterations is reached.

Cluster-based network model

Based on the clustering result, Kaiser *et al.* (2014) derived a cluster-based Markov model (CMM), which provides a probabilistic representation of the system using a Markov process, with the assumption that the fluid system is memoryless. Nair *et al.* (2019) removed the transitions residing in the same cluster and emphasized the non-trivial transitions between two different clusters. In these two works, the transitions are only characterised by probabilities. The cluster-based network model (CNM) proposed in Fernex *et al.* (2021) and Li *et al.* (2021), inherited the idea of focusing on the non-trivial transitions, and further introduced time-scale characteristics by recording the transition times. We here briefly review some concepts of the CNM, as they will be used in our benchmark of HiCNM.

The M consecutive snapshots define $M - 1$ transitions, containing trivial transitions staying in the same cluster and non-trivial transitions between two different clusters. The number of transitions from $\mathcal{C}_j$ to $\mathcal{C}_i$ reads

$$n_{ij} := \sum_{m=1}^{M-1} \chi_j^m \chi_i^{m+1}. \tag{5.16}$$

Considering the non-trivial transitions, n_j is the total number of departing snapshots from $\mathcal{C}_j$, with $n_j = \sum_{i=1}^{K} (1 - \delta_{ij}) n_{ij}$. The direct transition probability P_{ij} reads

$$P_{ij} = \frac{(1 - \delta_{ij}) n_{ij}}{n_j}, \quad i, j = 1, \cdots, K \tag{5.17}$$

where the non-migrating transition n_{jj} is eliminated. All the non-trivial transitions are identified with the direct transition matrix $\mathbf{P}$.

The residence time matrix $\mathbf{T}$ relies on the time information of the snapshots. After clustering, each snapshot $\boldsymbol{u}^m = \boldsymbol{u}(t^m)$ is associated with the closest centroid with a cluster index k^m. Assuming that N ($N < M$) non-trivial transitions occur along the trajectory, the moments of transition t_n, $n = 1, \ldots, N$ — including the initial time $t_0 = t^1$ — are defined as the time entering into a new cluster,

$$t_n = t^m \text{ if } \boldsymbol{u}^{m-1} \in \mathcal{C}_k \ \ \& \ \ \boldsymbol{u}^m \notin \mathcal{C}_k, \tag{5.18}$$

with ascending order $t_0 < t_1 < \ldots < t_N$. The cluster index k^m remains unchanged in a time range $[t_n, t_{n+1})$. The sequence of visited clusters over time can be simplified with the first entering snapshot for each non-trivial transition with k^m taken from the moments of transition t_n.

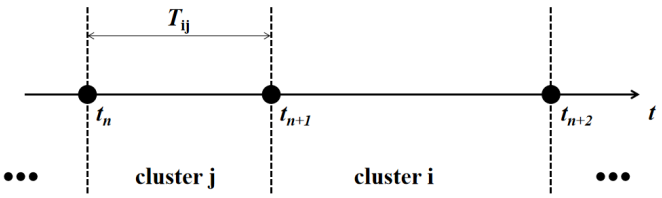


Figure 5.5: An illustration of the residence time in the cluster-based network model. • remarks the entering time into new clusters.

For a simple transition from $\mathcal{C}_j$ to $\mathcal{C}_i$ as illustrated in figure 5.5, where the trajectory first enters in $\mathcal{C}_j$ at time t_n, and leaves $\mathcal{C}_j$ for $\mathcal{C}_i$ at time t_{n+1}, the residence time in $\mathcal{C}_j$ is defined as

$$T_{ij} = t_{n+1} - t_n. \tag{5.19}$$

In the case of multiple trajectories of transition from $\mathcal{C}_j$ to $\mathcal{C}_i$, the residence time will be averaged according to the number of trajectories.

At this point, the time-resolved snapshots $\boldsymbol{u}^m$ can be represented by cluster centroids $\boldsymbol{c}_{k^m}$ with the time evolution of the cluster index k^m. The transient dynamics is described by both the direct transition matrix of non-migrating transitions $\mathbf{P}$ and the residence time matrix $\mathbf{T}$.

5.3.2 Hierarchical modelling with mean-field consideration

As introduced in figure 5.1, the framework of cluster-based hierarchical network modelling contains the following three steps. The hierarchical clustering with mean-field consideration is introduced in § 5.3.2. Based on the identified clusters, the network modelling with hierarchical structure is derived in § 5.3.2 for the mean-field model of Eq. (5.6) under a small number of general assumptions. § 5.3.2 introduces the autocorrelation function and its root mean square error of the rebuilt flow for the validation of the HiCNM. An introductory example is introduced in Appendix 5.B to clarify the primary form of the HiCNM.

Hierarchical clustering inspired by the triple decomposition

To better understand the global and local properties of the data, we present a novel hierarchical clustering algorithm inspired by the triple decomposition introduced in § 5.3.1. The principle of hierarchical clustering is to divide the snapshots into layers of clusters. Snapshots belonging to clusters of the parent layer are further partitioned into clusters of the child layer. Hierarchical clustering algorithms are generally divided into two categories:

(a) The agglomerating ("bottom-up") hierarchical clustering begins with the smallest clusters at the bottom, each snapshot being an elementary cluster. The two closest clusters are merged to generate a new cluster according to certain criteria, introducing an additional layer from the bottom. This merging is repeated until all snapshots belong to one cluster at the top of the hierarchy.

(b) The divisive ("top-down") hierarchical clustering starts with only one cluster, which owns all the snapshots. In our case, the centroid of the top cluster would be the mean-flow field from by ensemble averaging. From the top to the bottom, the snapshots in each cluster of the parent layer are divided into multiple clusters in the child layer, according to certain criteria. The bottom layers will be associated with the small scale fluctuations of the flow field. The division can be continued until each snapshot is a cluster.

We employ a divisive hierarchical clustering to distil the different features in a hierarchy, which is consistent with the triple decomposition of Eq. (5.6). For instance, fluid flows characterised by multiple frequencies require only a finite number of layers to describe the different components bounded by frequency.

Transient and post-transient dynamics are statistically non-homogeneous due to the existence of multiple invariant sets. If so, a scale subdivision of the flow-field decomposition like in Eq. (5.6) is used during the clustering process. Accounting for the Reynolds stress contribution, the slowly-varying mean-flow field $\langle \boldsymbol{u} \rangle_T$ is enough to describe the global trend. Next, the local dynamics around $\langle \boldsymbol{u} \rangle_T$ can be zoomed in, considering the coherent structures involved in $\tilde{\boldsymbol{u}}$. This scale subdivision can still be extended to a hierarchical structure with more layers, which involve secondary frequencies in the case of quasi-periodic dynamics or turbulence from $\boldsymbol{u}'$, as illustrated in figure 5.6.

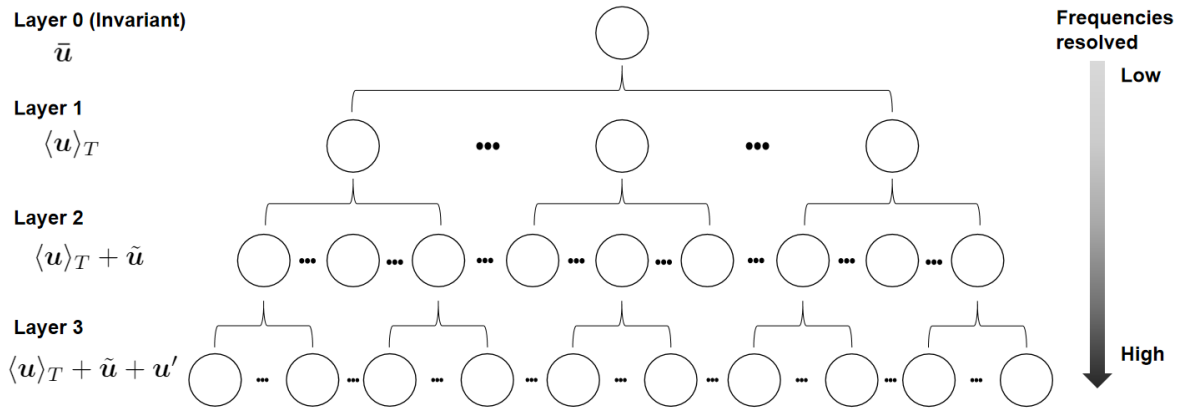


Figure 5.6: An illustration of the hierarchical structure with different scales in the triple flow decomposition in § 5.3.1. Layer 0: The top layer is characterized by the invariant mean flow $\bar{\boldsymbol{u}}$. Layer 1: The global trend is described by the slowly-varying mean-flow field $\langle \boldsymbol{u} \rangle_T$. Layer 2: The coherent part $\tilde{\boldsymbol{u}}$ is added for the local dynamics around the varying mean-flow field. Layer 3: The non-coherent part $\boldsymbol{u}'$ is considered in the case of turbulent flow.

In order to clearly describe the clusters in the hierarchy, we systematically name the clusters from top to bottom. The sole cluster on the top $\mathcal{L}_0$ contains the ensemble of input data, and we define it symbolically as $\mathcal{C}_0$. The sub-division of this cluster leads to K_1 subclusters in the first layer $\mathcal{L}_1$, named as $\mathcal{C}_{0,k_1}$, $k_1 = 1, \cdots, K_1$. The first subscript $k_0 = 0$ can be ignored because there is only one cluster in $\mathcal{L}_0$, and the second subscript k_1 indicates the index of the subcluster in $\mathcal{L}_1$. The second subdivision works on each cluster $\mathcal{C}_{k_1}$ separately, and generates refined K_2 subclusters for each of them. The cluster index in the current layer $\mathcal{L}_2$ is presented by an additional subscript $k_2 = 1, \cdots, K_2$, which is written as $\mathcal{C}_{k_1,k_2}$. For a higher layer number $\mathcal{L}_{L\in\mathbb{N}}$, $L \geqslant 3$, more subscripts $k_L, l = 1, \cdots, L$, are needed to record the cluster index in each layer $\mathcal{L}_l$ from the top to the bottom, written as $\mathcal{C}_{k_1,\ldots,k_L}$. This naming method can clearly trace out all clusters in the hierarchy, and also works for other properties of clusters, e.g., the centorids $\boldsymbol{c}_{k_1,\ldots,k_L}$ and the characteristic function $\chi^m_{k_1,\ldots,k_L}$. In this work, two or three layers ($L \leqslant 3$) will be enough to extract the transient dynamics out of multiple invariant sets.

Hierarchical network modelling

The starting point is the hierarchical Markov model of Fine *et al.* (1998), which introduces the hierarchical structure to describe the stochastic processes, comparing to the standard Markov model. Each state of a Markov model in the parent layer is considered separately, and a new Markov model of the sub-states of a state is built in the child layer. As the layer increases, the state is continuously subdivided. Therefore, the hierarchical Markov model records a sequence of states in different layers. In our case, each cluster is seen as a state. When a cluster in the parent layer is activated, its subclusters in the child layer turns activated recursively. Meanwhile, the refined dynamics between the subclusters can be described by a Markov model. Hence, the hierarchical Markov model can more effectively solve the problem of subsets.

In this work, we derive the HiCNM by replacing the Markov model by the network model. The hierarchical structure is identical, and the only change is the way to describe the transient dynamics between clusters.

A typical structure between the parent and child layers is shown in figure 5.7. We start with

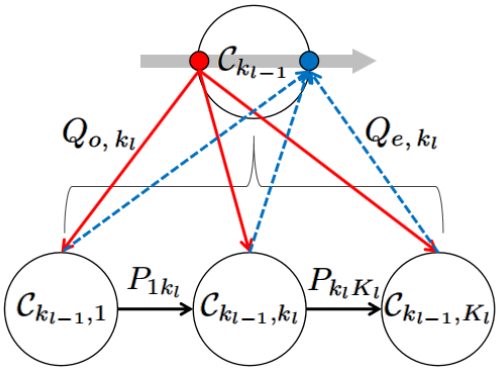


Figure 5.7: An illustration of the transitions between the parent and child layers in the hierarchical network model. The trajectories pass through the cluster in the parent layer: the entering and exiting snapshots are marked with red dot and the blue dot. After clustering, a classic network model is built between N subclusters, with transition probability P. The vertical transitions indicates the ports of entry and exit of the subclusters with probability $Q_{o,j}$ and $Q_{e,j}$.

cluster $\mathcal{C}_{k_1,\ldots,k_{l-1}}$ in the parent layer $\mathcal{L}_{l-1}$, where the leading subscripts $k_1, \ldots, k_{l-2}$ refers to the cluster number in each upper layer. For convenience, when the context will be unambiguous, the cluster will be only referenced by its number in the current layer, e.g. $\mathcal{C}_{k_{l-1}}$. As indicated with the sequence of cluster numbers in its complete name, this cluster comes from the sub-division of the cluster $\mathcal{C}_{k_1,\ldots,k_{l-2}}$ in the parent layer $\mathcal{L}_{l-2}$. We suppose that M snapshots $\boldsymbol{u}^m$, $m = 1, \ldots, M$, exist in this cluster and are divided into K_{l-1} subclusters by a sub-division clustering algorithm. The sub-cluster $\mathcal{C}_{k_{l-1}}$ contains $n_{k_{l-1}}$ snapshots, calculated from Eq. (5.12) with the characteristic function $\chi^m_{k_{l-1}}$. A standard network model for the cluster $\mathcal{C}_{k_1,\ldots,k_{l-2}}$ can be derived with the direct transition matrix $\mathbf{P}_{k_1,\ldots,k_{l-2}}$ and the residence time matrix $\mathbf{T}_{k_1,\ldots,k_{l-2}}$, as recorded in § 5.3.1, which describe the dynamics between the subclusters $\mathcal{C}_{k_{l-1}}$.

In the following, we focus on the trajectories passing through cluster $\mathcal{C}_{k_{l-1}}$. The snapshots entering and leaving from $\mathcal{C}_{k_{l-1}}$ are marked out for each trajectory, with the following characteristic function

$$\begin{aligned}
\chi^m_{o,k_{l-1}} &:= \begin{cases} 1, & \text{if } \boldsymbol{u}^{m-1} \notin \mathcal{C}_{k_{l-1}} \ \ \& \ \ \boldsymbol{u}^m \in \mathcal{C}_{k_{l-1}}, \\ 0, & \text{otherwise.} \end{cases} \\
\chi^m_{e,k_{l-1}} &:= \begin{cases} 1, & \text{if } \boldsymbol{u}^m \in \mathcal{C}_{k_{l-1}} \ \ \& \ \ \boldsymbol{u}^{m+1} \notin \mathcal{C}_{k_{l-1}}, \\ 0, & \text{otherwise.} \end{cases}
\end{aligned} \tag{5.20}$$

The entering snapshots are denoted by the subscript "o", and the exiting snapshots by the subscript

"e". The number of entering snapshots n_o and of exiting snapshots n_e read

$$n_o = \sum_{m=1}^{M} \chi^m_{o,k_{l-1}}, \quad n_e = \sum_{m=1}^{M} \chi^m_{e,k_{l-1}}. \tag{5.21}$$

In the child layer $\mathcal{L}_l$, the n_k snapshots in the cluster $\mathcal{C}_{k_{l-1}}$ have been divided into the subclusters $\mathcal{C}_{k_1,\ldots,k_l}$, $k_l = 1,\ldots,k_L$. Without loss of generality, a standard network model for the cluster $\mathcal{C}_{k_{l-1}}$ can be built with its subclusters with the direct transition matrix $\mathbf{P}_{k_1,\ldots,k_{l-1}}$ and the residence time matrix $\mathbf{T}_{k_1,\ldots,k_{l-1}}$.

The snapshots $\boldsymbol{u}^m$ are approximated by the time evolution of the cluster centroids $\boldsymbol{c}_{k_1^m,\ldots,k_l^m}$ in $\mathcal{L}_l$,

$$\hat{\boldsymbol{u}}^m_{\mathcal{L}_l} = \boldsymbol{c}_{k_1^m,\ldots,k_l^m}. \tag{5.22}$$

The residence time elements of $\mathbf{T}_{k_1,\ldots,k_l}$ can be assembled in order to determine the moments of transition based on the sequence of visited clusters in $\mathcal{L}_l$.

The entering and exiting snapshots defined in Eq. (5.20) can be used to describe the vertical transitions. Although they are not necessary to describe the dynamics of the fluidic pinball, the probability of the vertical transitions $Q_{o,j}$ and $Q_{e,j}$, described in appendix 5.B, completes all possible transitions in our hierachical structure and make it consistent with the classic hierarchical Markov model of Fine *et al.* (1998).

Dynamics reconstruction of the hierarchical network model

The reconstructed flow in Eq. (5.22) is a statistical representation of the original snapshot sequence by a few representative centroids, which is a highly discretized description compared to the full dynamics. The approximations in the different layers provide different metrics for the flow dynamics.

The cluster-based hierarchical model uses the centroids $\boldsymbol{c}_{k_1^m,\ldots,k_l^m}$ in the original data space together with the time evolution of the cluster index k_l^m to simplify the description of the original flow, which is more intuitive and closer to the original flow than the POD reconstruction. For input data with I-dimensional state vectors of the velocity field and M snapshots, a POD reconstruction truncated to R modes will lead to a $I \times R$ matrix of POD modes and a $R \times M$ matrix of mode amplitudes. A HiCNM with K centroids leads to a $I \times K$ matrix of centroids and a sequence of the N visited clusters of length $N \ll M$. In this sense, the compressive ability of HiCNM is more powerful for the large amount of continuously sampled data, as $I \times K < (I + M) \times R$. In addition, the hierarchical clustering works as a sparse sampling technique, extracting the representive states according to the clustering subspace.

We use the unbiased auto-correlation function (Protas *et al.*, 2015),

$$R(\tau) = \frac{1}{T-\tau} \int_{\tau}^{T} \left(\boldsymbol{u}(\boldsymbol{x}, t-\tau) - \boldsymbol{u}_s(\boldsymbol{x}), \boldsymbol{u}(\boldsymbol{x}, t) - \boldsymbol{u}_s(\boldsymbol{x})\right)_\Omega \mathrm{d}t, \quad \tau \in [0, T), \tag{5.23}$$

after normalization with respect to $R(0)$ to check the accuracy of the dynamics reconstruction of the HiCNM in Eq. (5.22). The autocorrelation function without delay $R(0)$ is twice the time-averaging kinetic energy. The modeled autocorrelation function $\hat{R}_{\mathcal{L}_l}(\tau)$ in layer $\mathcal{L}_l$ is based on the rebuilt flow $\hat{\boldsymbol{u}}_{\mathcal{L}_l}$ in Eq. (5.22) instead of $\boldsymbol{u}$ in Eq. (5.23).

For the discrete snapshots, the root mean-square error (RMSE) of the autocorrelation function $R(\tau)$ of the reference data and that of the model $\hat{R}_{\mathcal{L}_l}(\tau)$ is defined as

$$R^l_{\mathrm{rms}} = \sqrt{\frac{1}{M} \sum_{m=1}^{M} \left(R(\tau) - \hat{R}_{\mathcal{L}_l}(\tau)\right)^2}, \tag{5.24}$$

where M is the number of snapshots $\boldsymbol{u}^m$ and $\hat{\boldsymbol{u}}^m_{\mathcal{L}_l}$ at $\tau = m\Delta t$.

5.4 Hierarchical network modelling of the fluidic pinball

In this section, we apply the hierarchical modelling strategy to the fluidic pinball at different Reynolds numbers. With increasing Reynolds number, the flow dynamics is undergoing successive instabilities and bifurcations, introducing multiple exact solutions of the Navier-Stokes equations and multiple invariant sets for the dynamics. In § 5.4.1, the modelling strategy dealing with multiple invariant sets is introduced. We derive the HiCNMs for the transient dynamics involving six invariant sets at $Re = 80$ in § 5.4.2, for the quasi-periodic regime at $Re = 105$ in § 5.4.4, and for the chaotic regime at $Re = 130$ in § 5.4.5.

5.4.1 Hierarchical modelling with multiple invariant sets

The flow field is computed with the direct numerical simulation (DNS) described in § 5.2.1. The resulting flow field is an ensemble of time-resolved snapshots starting with some given initial condition. The transient and post-transient dynamics of the flow constitute a time-resolved trajectory sampled with a fixed time step. Different invariant sets and multiple attractors can co-exist in the state space, only part of them being explored by each individual trajectory from the initial condition to the asymptotic regime. All the cases of interest in this paper are such that $Re > Re_2$, i.e. beyond the supercritical pitchfork bifurcation. The data set consists of the snapshots computed from four different trajectories: two mirror-conjugated trajectories starting in the vicinity of the symmetric steady solution, the two others starting from the two mirror-conjugated asymmetric steady solutions. The simulations are respectively run until $t = 1\,500$ and $t = 1\,000$ for the symmetric steady solution and the asymmetric steady solutions.

Sampled with time step $\Delta t = 0.1$, the input data basis is an ensemble of $M = 50\,000$ snapshots $\boldsymbol{u}^m(\boldsymbol{x})$ from four transient trajectories, where the superscript m is the snapshot index for the successive instants $t^m = m\Delta t$. In order to distinguish the different trajectories, the snapshot index m is sorted as:

(a) $m = 1, \ldots, 15000$ and $m = 15001, \ldots, 30000$ for the two mirror-conjugated trajectories starting in the vicinity of the symmetric steady solution,

(b) $m = 30001, \ldots, 40000$ and $m = 40001, \ldots, 50000$ for the two others starting from the two mirror-conjugated asymmetric steady solutions.

The time continuity is critically important during the dynamical analysis. Snapshots in each trajectory are time-resolved but have no time relationship in different trajectories. For a trajectory of M snapshots, it exists $M - 1$ transitions as described in § 5.3.1. Hence, $M - 4$ transitions occur in the four individual trajectories mentioned above.

In our case, a standard network model with 200 clusters are still not enough to distinguish the transitions starting from three different steady solutions, as shown in § 5.4.3. An optimal way to achieve correct classification is to hierarchically cluster the ensemble of snapshots. The hierarchical clustering is performed with the slaving assumption under the mean-field consideration in § 5.3.1, by applying an unsupervised clustering algorithm (k-means++) for different time scales. The clusters in the parent layer are split into subclusters in the child layer, where the clustering result in the parent layer works as a pre-classified indicator in the child layer. The number of clusters is preset to 10 for each clustering algorithm, but can be adjusted to the minimal number for an accurate dynamics reconstruction, typically for the network model in the first layer. The clustering algorithm in the first layer is meant to distinguish different invariant sets together with the transitions between them with a limited number of clusters. These clusters will be used to build a network model for the mean-field distortion, which will further supervise the clustering process in the second layer. A sketch for this process is shown in figure 5.8.

From the mean-field consideration, the slowly-varying mean-flow field is the ideal candidate for the detection of several invariant sets. A fifth order Butterworth low-pass filter with cutoff

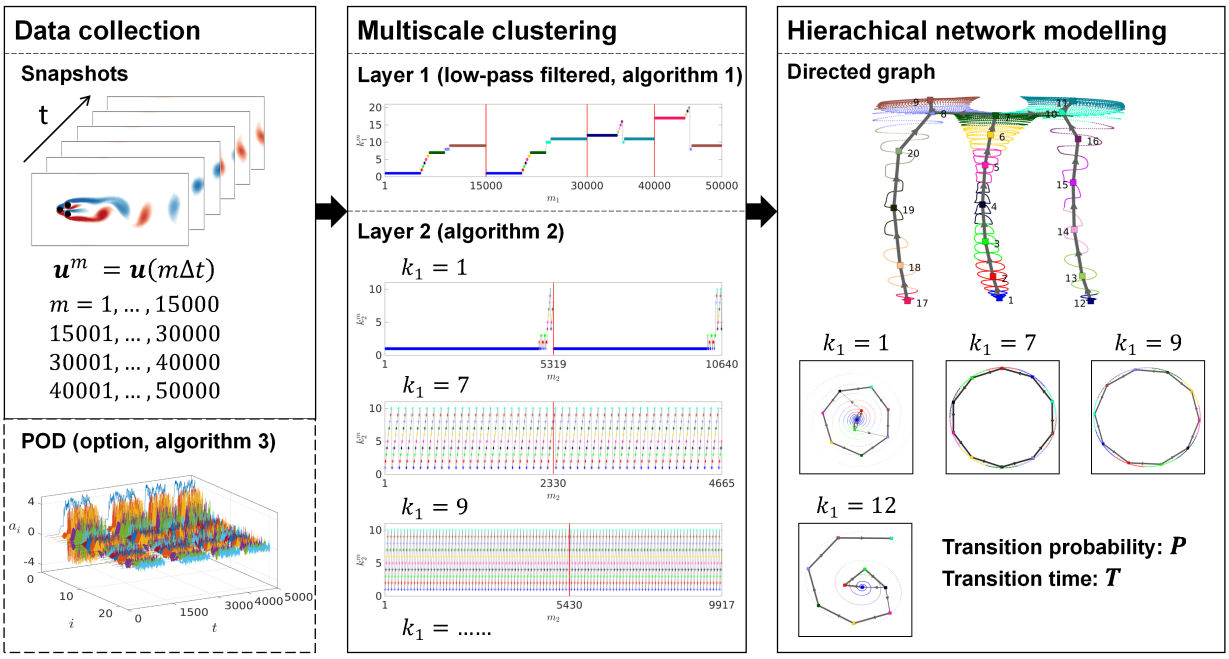


Figure 5.8: Sketch for HiCNM applied to the fluidic pinball at $Re = 80$. See text for the details.

frequency $0.2f_c$ is applied to eliminate the coherent $\tilde{\boldsymbol{u}}$ and incoherent $\boldsymbol{u}'$ components of $\boldsymbol{u}$ in the triple decomposition of Eq.(5.6). The clustering algorithm applied to the first layer is described in algorithm 1. The critical idea of the algorithm is to map the original data to the bounded low-

Algorithm 1 Clustering algorithm with slowly varying mean flow

Input: $\boldsymbol{u}^m$: snapshots; f_c: frequency of coherent part;
K_1: number of clusters

Output: $\chi_{k_1}^m$: characteristic function; k_1^m: cluster indexes of snapshots $\boldsymbol{u}^m$;
$\boldsymbol{c}_{k_1}$: optimal centroids

1: compute the low-pass filtered $\boldsymbol{u}^m$ with cutoff frequency $0.2f_c$, named $\boldsymbol{u}_{\mathrm{LP}}^m$;
2: apply k-means++ algorithm with K_1 clusters to $\boldsymbol{u}_{\mathrm{LP}}^m$, and save the characteristic function $\chi_{k_1}^m$ and the cluster indexes k_1^m;
3: compute and save the centroids in original data space:

$$\boldsymbol{c}_{k_1} = \sum_{m=1}^{M} \chi_{k_1}^m \boldsymbol{u}^m / \sum_{m=1}^{M} \chi_{k_1}^m.$$

frequency space, and then calculate the characteristic function in the low-frequency space. The resulting characteristic function $\chi_{k_1}^m$ is applied to the original data to achieve the clustering of the slowly-varying mean-flow field.

The divisive clustering algorithm of the clusters in the parent layer is described in algorithm 2, under the supervision of the characteristic function $\chi_{k_1}^m$ obtained from the parent layer. The subdivision of the clusters in the parent layer leads to a more detailed network model of the local structures. According to the spectral content of the dynamics, multiple layers are introduced to extract the coherent dynamics. The naming method introduced in § 5.3.2 can clearly locate all clusters in the hierarchy. When dealing with chaotic flow regime, there is no clear frequency boundary. We stop modelling the incoherent components $\boldsymbol{u}'$ with a simple network model which contains all the chaotic dynamics. Snapshots of velocity field can be highly compressed by a lossless POD to accelerate the clustering algorithm, as detailed in appendix 5.C.

Algorithm 2 Divisive clustering algorithm in the child layer under supervision

Input: $\boldsymbol{u}^m$: snapshots;
$\chi^m_{k_1}$: characteristic function from the parent layer
Output: for each cluster $\mathcal{C}_{k_1}$ in parent layer
$\chi^m_{k_1,k_2}$: the characteristic function; k^m_2: cluster indexes of snapshots $\boldsymbol{u}^m$;
$\boldsymbol{c}_{k_1,k_2}$: optimal centroids;

1: **for** $k_1 \leftarrow 1$ to K_1 **do**
2: locate the snapshots $\boldsymbol{u}^m$ in the cluster $\mathcal{C}_{k_1}$ by the characteristic function $\chi^m_{k_1}$, and record the snapshot index m_1 of the resulting n_{k_1} snapshots;
3: extract the snapshots $\boldsymbol{u}^{m_1}$ and renumber them sequentially with $m_2 = 1, \ldots, M_2$;
4: apply k-means++ algorithm with $K_2 = 10$ clusters (as default) to the renumbered snapshots $\boldsymbol{u}^{m_2}$ and save the characteristic function $\chi^{m_2}_{k_1,k_2}$ and the centroids $\boldsymbol{c}_{k_1,k_2}$:
$\boldsymbol{c}_{k_1,k_2} = \sum_{m_2=1}^{M_2} \chi^{m_2}_{k_1,k_2} \boldsymbol{u}^{m_2} / \sum_{m_2=1}^{M_2} \chi^{m_2}_{k_1,k_2}$.
5: **end for**

5.4.2 Hierarchical network model at $Re = 80$

At $Re = 80 > Re_2$, the system has already undergone a supercritical Hopf bifurcation and two coincidental supercritical pitchfork bifurcations on the steady solution and the symmetric limit cycle. As a result, three unstable steady solutions, one unstable (symmetric) limit cycle and two stable (asymmetric) mirror-conjugated limit cycles exist in the state space and organize the dynamics. Thus, there are six invariant sets, the two stable limit cycles being the attractors of the flow state.

The HiCNM in the first layer is based on the clustering results of the low-pass filtered data set (§ 5.4.2). The local dynamics for some typical regimes is further presented with the subclusters in the second layer (§ 5.4.2).

Hierarchical network model in Layer 1

$K_1 = 20$ clusters are used to cluster the snapshots from the low-pass filtered data set which removes the coherent structures with frequency $f_c = 0.1074$. To visualize the cluster topology, we apply the classical multidimensional scaling (MDS) to represent the high-dimensional centroids in a two-dimensional subspace $[\gamma_1, \gamma_2]^{\mathrm{T}}$, while the distances between the centroids are preserved (Kaiser *et al.*, 2014). As shown in figure 5.9, the six exact solutions of the Navier-Stokes equations that organize the state space are well identified. The vorticity field of the resulting centroids can be understood as the slowly varying mean-flow field $< \boldsymbol{u} >_T$ along the transient dynamics, with $T \gg 2\pi/\omega_c$. The three steady solutions belong respectively to clusters $\mathcal{C}_1$ (symmetric steady solution $\boldsymbol{u}_s$), $\mathcal{C}_{12}$ (asymmetric steady solution $\boldsymbol{u}_s^-$) and $\mathcal{C}_{17}$ (asymmetric steady solution $\boldsymbol{u}_s^+$). The three limit cycles are caught by the time-averaged flow in clusters $\mathcal{C}_7$ (symmetric mean-flow field, centroid $\bar{\boldsymbol{u}}^0$), $\mathcal{C}_9$ (asymmetric upward mean-flow field, centroid $\bar{\boldsymbol{u}}^+$) and $\mathcal{C}_{11}$ (asymmetric downward mean-flow field, centroid $\bar{\boldsymbol{u}}^-$). A network of four transient trajectories connects these clusters as follows:

$Trajectory\,1:\ \mathcal{C}_1\ (\boldsymbol{u}_s) \to \ldots \to \mathcal{C}_7\ (\bar{\boldsymbol{u}}^0)\ \to \mathcal{C}_8\ \to \mathcal{C}_9\ (\bar{\boldsymbol{u}}^+)$;
$Trajectory\,2:\ \mathcal{C}_1\ (\boldsymbol{u}_s) \to \ldots \to \mathcal{C}_7\ (\bar{\boldsymbol{u}}^0)\ \to \mathcal{C}_{10} \to \mathcal{C}_{11}\ (\bar{\boldsymbol{u}}^-)$;
$Trajectory\,3:\ \mathcal{C}_{12}\ (\boldsymbol{u}_s^-) \to \ldots \to \mathcal{C}_{16} \to \mathcal{C}_{10} \to \mathcal{C}_{11}\ (\bar{\boldsymbol{u}}^-)$;
$Trajectory\,4:\ \mathcal{C}_{17}\ (\boldsymbol{u}_s^+) \to \ldots \to \mathcal{C}_{20} \to \mathcal{C}_8\ \to \mathcal{C}_9\ (\bar{\boldsymbol{u}}^+)$.

We notice that most of the transitions between clusters are with 100% probability, except for the bifurcating cluster $\mathcal{C}_7$ with half probability to transition to clusters $\mathcal{C}_8$ or $\mathcal{C}_{10}$. The transition matrix in figure 5.10(b) illustrates the probability of all the transitions. The probability is 1 for the surely

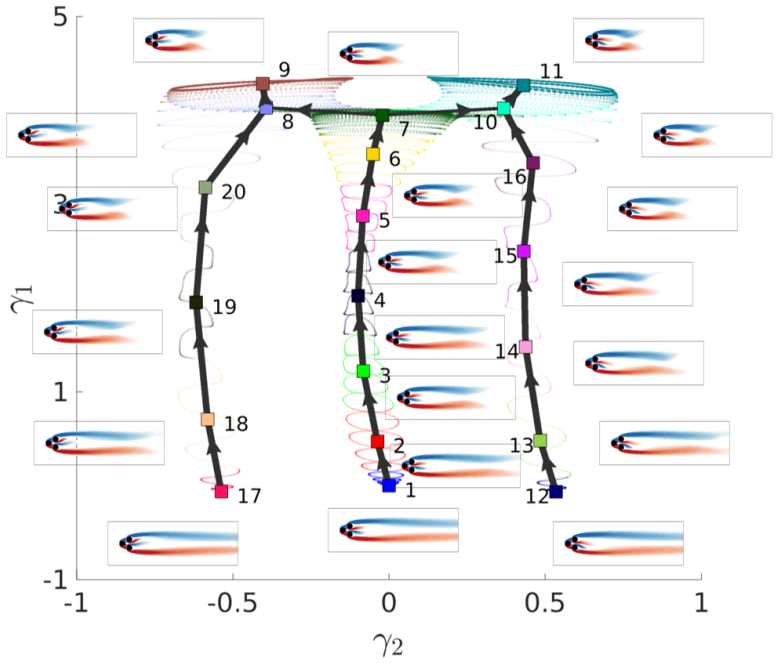


Figure 5.9: Graph of transitions between clusters in layer 1 at $Re = 80$. Cluster centroids are marked with the colored squares, with their vorticity fields in color with $[-1.5, 1.5]$. The snapshots belonging to them are marked as small dots with the same colors. The transitions between clusters are shown with arrows, where the line width presents the probability of transition.

directed transitions. By contrast, $P_{78} = 0.5$ and $P_{710} = 0.5$ for the bifurcating cluster. After entering into $\mathcal{C}_8$ and $\mathcal{C}_{10}$, the flow will surely enters the two clusters $\mathcal{C}_9$ and $\mathcal{C}_{11}$ respectively. The two clusters $\mathcal{C}_9$ and $\mathcal{C}_{11}$ catch the permanent regimes, from which the dynamics cannot escape, imposing all the terms in the ninth and eleventh columns $P_{j\,9} = P_{j\,11} = 0,\ \forall j$.

The time ordering of the cluster transitions is illustrated in figure 5.10(a). The red vertical line separates the four trajectories. All the trajectories are irreversible transition from one steady solution to one of the two stable periodic solutions. In figure 5.10(c), the filled black circles emphasize the transitions starting from $\mathcal{C}_1$, $\mathcal{C}_{12}$, $\mathcal{C}_{17}$ and $\mathcal{C}_7$. The two attracting clusters $\mathcal{C}_9$ and $\mathcal{C}_{11}$ have no transition to any other clusters. Hence, all the terms in the ninth and eleventh columns $T_{j\,9}, T_{j\,11}$ are 0, $\forall j$. The residence time associated with clusters $\mathcal{C}_9$ and $\mathcal{C}_{11}$ is infinite.

According to the above discussion, the network model in this layer has successfully identified the six invariant sets of the dynamics, four being unstable, the two others being the attractors of the system.

Hierarchical network model in layer 2

We apply algorithm 2 based on the clustering result for the layer $\mathcal{L}_1$ in § 5.4.2. In the second layer $\mathcal{L}_2$, we will isolate and analyze the clusters $\mathcal{C}_{k_1},\ k_1 = 9, 7, 1$ associated with three invariant sets of the dynamics.

The permanent regime in cluster $\mathcal{C}_9$

Cluster $\mathcal{C}_9$ is associated with one of the two asymmetric limit cycles. The k-means++ algorithm is directly applied to the intra-cluster snapshots $\boldsymbol{u}^m \in \mathcal{C}_9$. The limit cycle in figure 5.11(a) has been divided into $K_2 = 10$ subclusters according to algorithm 2. The centroids are distributed on the limit cycle at equal distances. The arrows between the centroids form a closed loop, which results from the periodic nature of the oscillating dynamics. The resulting centroids are phase-averaged flow fields along the complete period of the vortex shedding. The inner jet-flow of the centroids in

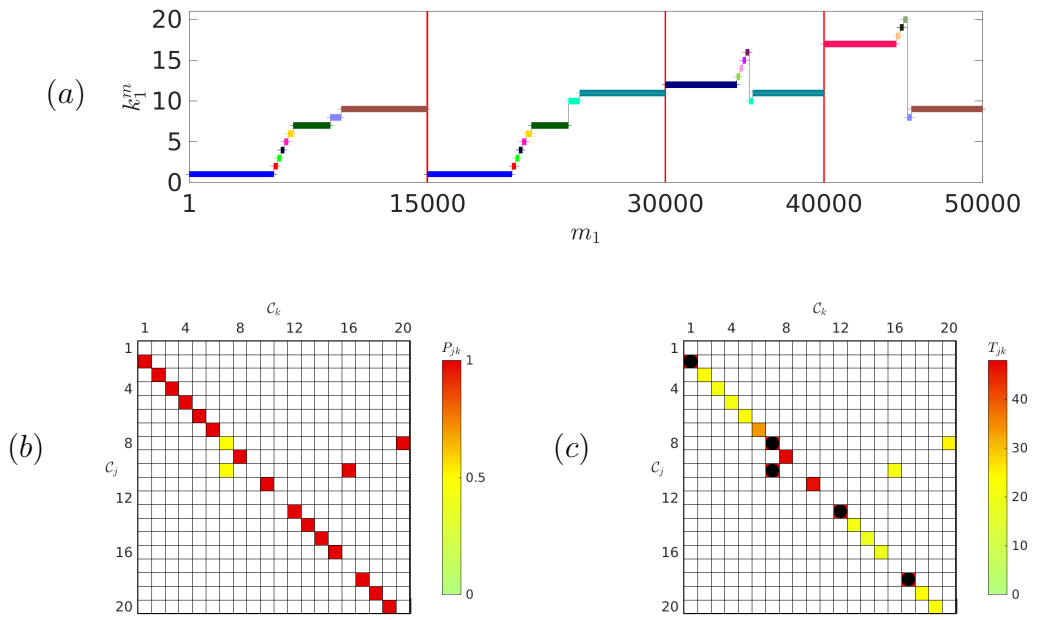


Figure 5.10: Cluster-based analysis at $Re = 80$ in layer 1: (a) transition illustrated with cluster label, (b) transition matrix, (c) residence time matrix. Since subclusters are not considered, the snapshot index in layer 1 m_1 is identical to the original index m. The colorbar indicates the values of the terms. Residence time larger than 100 is marked with a solid black circle and excluded from the colorbar. An extremely long residence time in a cluster indicates data density, and such a cluster is generally associated with an invariant set.

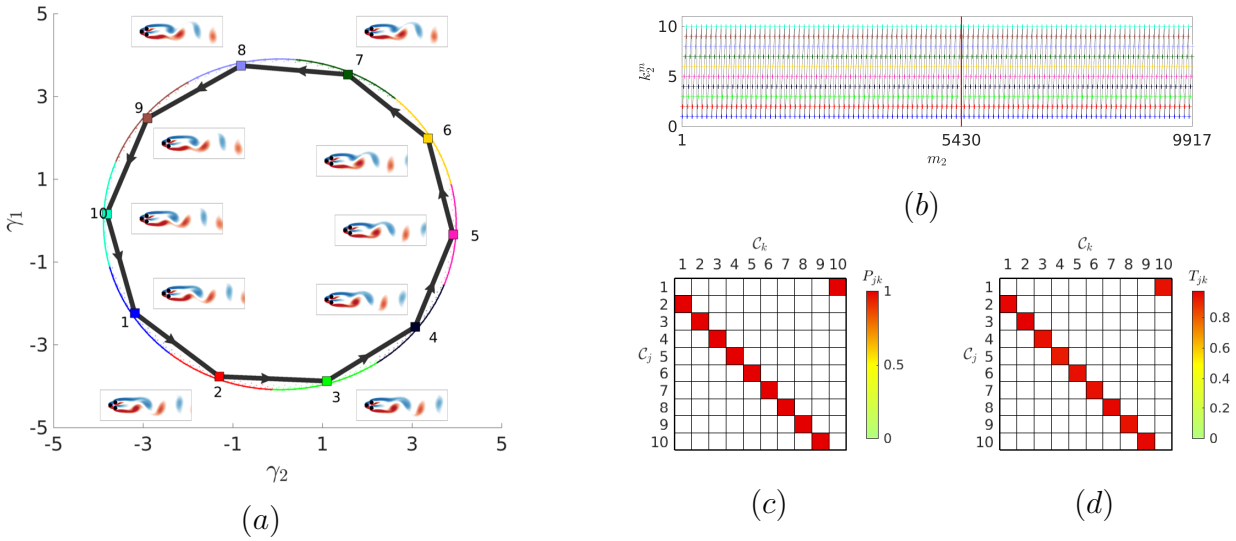


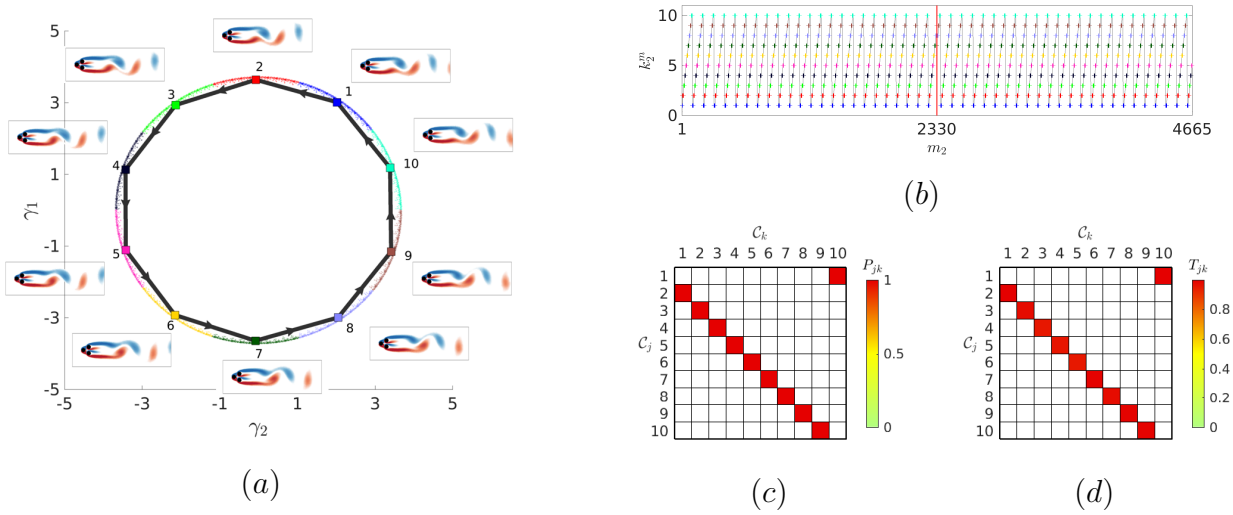
Figure 5.11: Cluster-based analysis in layer 2 at $Re = 80$ for $\mathcal{C}_9$: (a) graph of non-trivial transitions between clusters, as in figure 5.9, (b) transition illustrated with cluster label, (c) transition matrix, (d) residence time matrix, as in figure 5.10. Two trajectories pass through $\mathcal{C}_9$ in the parent layer, one with $m_2 = 1, \ldots, 5430$ and another with $m_2 = 5431, \ldots, 9917$.

figure 5.11(a) are all deflected upwards, as expected for the attractor that belongs to cluster $\mathcal{C}_9$.

Figure 5.11(b) associates the cluster labels to the dynamics. Two different transient trajectories reach $\mathcal{C}_9$ with entering snapshots $m = 9571$ and 45514, one issued from $\mathcal{C}_8$, the other from $\mathcal{C}_{16}$, according to the network model in layer $\mathcal{L}_1$. According to the entering time, the original snapshot index $m = m_2 + 9571 - 1 = 9571, \ldots, 15000$ and $m = m_2 + 45514 - 5431 = 45514, \ldots, 50000$. The flow periodically travels along the subclusters $\mathcal{C}_{9,k_2}$, $k_2 = 1, \ldots, 10$ as $\mathcal{C}_{9,1} \to \ldots \to \mathcal{C}_{9,10} \to \mathcal{C}_{9,1}$. The limit cycle has a clear and stable transition matrix, as each cluster only has one possible destination, see figure 5.11(c). The residence times in each clusters are uniform as shown in figure 5.11(d). The sum of all residence times is 9.51, which is close to the real time period 9.50 of the vortex shedding in the permanent regime computed from the DNS.

The bifurcating state in cluster $\mathcal{C}_7$

Cluster $\mathcal{C}_7$ is associated with the symmetric limit cycle. $K_2 = 10$ clusters are used to classify the snapshots in this cluster.



(a) (b) (c) (d)

Figure 5.12: Cluster-based analysis in layer 2 at $Re = 80$ for $\mathcal{C}_7$: (a) graph of non-trivial transitions between clusters, as in figure 5.9, (b) transition illustrated with cluster label, (c) transition matrix, (d) residence time matrix, as in figure 5.10. Two trajectories pass through $\mathcal{C}_7$ in the parent layer, one with $m_2 = 1, \ldots, 2330$ and another with $m_2 = 2331, \ldots, 4665$.

The resulting limit cycle of figure 5.12(a) differs from the limit cycle of figure 5.11 by its centroids. As expected with the symmetric limit cycle, the inner-jet is not deflected in cluster $\mathcal{C}_7$, while it is deflected in cluster $\mathcal{C}_9$. In figure 5.12(b), two different transient trajectories pass through $\mathcal{C}_7$ with entering snapshots $m = 6541$ and 21552, the original snapshot index is $m = 6541, \ldots, 8870$ and $m = 21552, \ldots, 23886$. The limit cycle has a stable transition matrix, as shown in figure 5.12(c). The sum of the residence times of figure 5.12(d) is 9.79, again very close to the period 9.80 of the symmetric transient vortex shedding computed from the DNS. The bifurcating dynamics can not be detected with these subclusters but can be captured in the parent layer.

The destabilizing regime in cluster $\mathcal{C}_1$

Cluster $\mathcal{C}_1$ is associated with two mirror-conjugated trajectories spiraling out of the symmetric steady solution. $K_2 = 10$ subclusters are used in the child layer $\mathcal{L}_2$. In the $[\gamma_1, \gamma_2]^{\mathrm{T}}$ representation of figure 5.13(a), the centroids are distributed along two diverging trajectories spiraling out of the fixed point $[0, 0]^{\mathrm{T}}$.

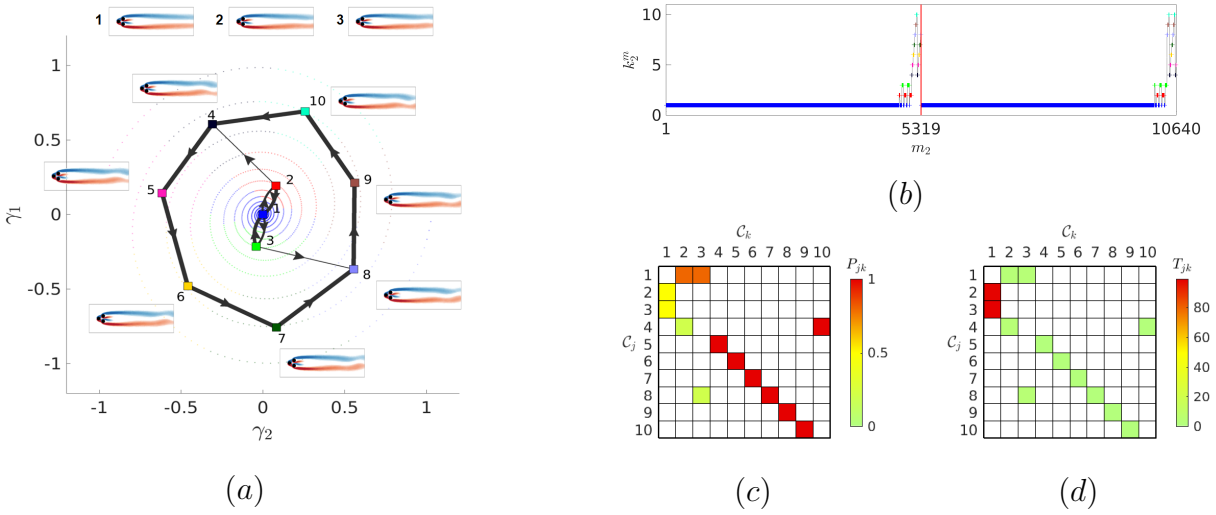


Figure 5.13: Cluster-based analysis in layer 2 at $Re = 80$ for $\mathcal{C}_1$: (a) graph of non-trivial transitions between clusters, as in figure 5.9, (b) transition illustrated with cluster label, (c) transition matrix, (d) residence time matrix, as in figure 5.10. Two trajectories pass through $\mathcal{C}_1$ in the parent layer, one with $m_2 = 1, \ldots, 5319$ and another with $m_2 = 5320, \ldots, 10640$.

The first three subclusters $\mathcal{C}_{1,1}$, $\mathcal{C}_{1,2}$ and $\mathcal{C}_{1,3}$ belong to the inner zone of the spirals, where the distribution of snapshots is dense in the $[\gamma_1, \gamma_2]^{\mathrm{T}}$ proximity map of figure 5.13. As a result, three nonphysical closed-loop cycles are formed between these three clusters, namely $\mathcal{C}_{1,1} \to \mathcal{C}_{1,2} \to \mathcal{C}_{1,1}$, $\mathcal{C}_{1,1} \to \mathcal{C}_{1,3} \to \mathcal{C}_{1,1}$, and $\mathcal{C}_{1,3} \to \mathcal{C}_{1,1} \to \mathcal{C}_{1,2} \to \mathcal{C}_{1,1} \to \mathcal{C}_{1,3}$. The remaining clusters belong to the outer arms of the spirals, with a relatively sparse distribution of the snapshots. The transitions between them form a closed-loop trajectory $\mathcal{C}_{1,4} \to \ldots \to \mathcal{C}_{1,10} \to \mathcal{C}_{1,4}$. The transitions $\mathcal{C}_{1,2} \to \mathcal{C}_{1,4}$ and $\mathcal{C}_{1,3} \to \mathcal{C}_{1,8}$ correspond to the departing dynamics out of the inner zone, due to the growth of the instability. The varying density of distribution comes from the exponential growth of the instability. The flow perturbations are small in the beginning of the instability while the flow distortions evolve faster in the later stages, leading to a multiscale problem in the transient and post-transient flow dynamics. In this case, the later stages with larger distortions are obviously easier to divide into different clusters.

From figure 5.13(b), two transient trajectories pass through $\mathcal{C}_1$ with entering snapshots $m = 1$ and 15001. The original snapshot index is $m = 1, \ldots, 5319$ and $m = 15001, \ldots, 20321$, corresponding to the initial stage of destabilization from the symmetric steady solution. In the transition matrix of figure 5.13(c), the inner and outer portions of the spiral are also apparent. The inner zone is the oscillating dynamics between $\mathcal{C}_{1,1}$ and $\mathcal{C}_{1,2}$, or $\mathcal{C}_{1,1}$ and $\mathcal{C}_{1,3}$. The outer zone has a more obvious periodic dynamics through the remaining clusters from $\mathcal{C}_{1,4}$ to $\mathcal{C}_{1,10}$. In figure 5.13(d), the residence time in each cluster is very short compared to the residence time in $\mathcal{C}_{1,1}$, which the vicinity of the steady solution belong to.

Dynamics reconstruction of the hierarchical network model at $Re = 80$

Figure 5.14 shows the autocorrelation function of the DNS and the HiCNM in different layers. As the autocorrelation function has been normalized by $R(0)$, the unit one presents the level of kinetic energy of the whole transition. For the transient and post-transient dynamics, the autocorrelation function vanishes with increasingtime shift, as shown in figure 5.14(b). The autocorrelation function of the DNS identifies the dominant frequency. In layer $\mathcal{L}_1$, no oscillation can be identified, due to the centroids by averaging the snapshots within clusters in the state space. The RMSE of the autocorrelation function is $R^1_{\mathrm{rms}} = 17.46$. In layer $\mathcal{L}_2$, the autocorrelation function of the model matches perfectly over the entire range, with $R^2_{\mathrm{rms}} = 1.18$, which quantifies the accuracy of the cluster-beasd model.

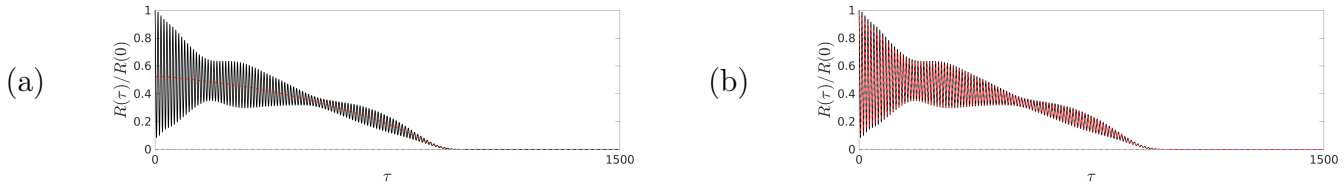


Figure 5.14: Autocorrelation function for $\tau \in [0, 1500)$ from DNS (black solid line) and the hierarchical network model (red dashed line) in the two layers: (a) $\mathcal{L}_1$ and (b) $\mathcal{L}_2$, at $Re = 80$

5.4.3 Advantages of HiCNM as compared to CNM

In this sub-section, we compare the results of the HiCNM in its second layer to the standard CNM with the same number of clusters. The hierarchical structure can systematically present the global trend and local dynamics, which improve the graphic interpretation of transient and post-transient, multi-frequency, multi-attractor behaviours.

Standard CNM at $Re = 80$

We show a standard network model treating the transient dynamics at $Re = 80$ with $K = 200$ clusters. The directed graph is not shown here as too many clusters overlapped in the two-dimensional subspace $[\gamma_1, \gamma_2]^{\mathrm{T}}$, losing the interpretation of the dynamics. In figure 5.15(a), the six exact solutions are well classified. The three steady solutions are divided into $\mathcal{C}_1$, $\mathcal{C}_{192}$ and $\mathcal{C}_{196}$ separately. The three limit cycles are identified with three blocks of oscillating labels, from $\mathcal{C}_{69}$ to $\mathcal{C}_{99}$, from $\mathcal{C}_{100}$ to $\mathcal{C}_{142}$, and from $\mathcal{C}_{148}$ to $\mathcal{C}_{191}$. However, the transient trajectories starting with the asymmetric steady solutions are misidentified, as both travelling through the same clusters from $\mathcal{C}_2$ to $\mathcal{C}_{68}$ and even the block of the symmetric limit cycle. It indicates that the clustering algorithm failed to distinguish the symmetry-breaking in the transient dynamics starting with the different steady solutions. Too many clusters also make the transition matrix hard to read as illustrated in figure 5.15(c), and also for the residence time matrix not shown here. Ignoring the transitions with low probability in the transition matrix, the transitions in the above-mentioned three blocks are almost definite, indicating the correct identification of the cycles. However, the transient dynamics from $\mathcal{C}_2$ to $\mathcal{C}_{68}$ is random. We can hardly find any relevant feature for the building of vortex shedding or symmetric breaking. The cluster distribution directly affects the analysis of the dynamics.

In summary, a large number of clusters tends to increase the resolution of the identified network model, but will misidentify the dynamics with the random transitions between clusters during the transient state. The standard network model fails to describe the transient dynamics between different invariant sets. This kind of problem comes from the poor distribution of clusters and can be solved by the hierarchical clustering strategy, as shown in the following sub-section.

HiCNM at $Re = 80$ in Layer 2

Based on the hierarchical network model at $Re = 80$ recorded in § 5.4.2, we build a network model ensembling all the clusters in the layer $\mathcal{L}_2$, which also contains $K = 200$ clusters. For ease of illustration, the cluster indexes of two layers $\mathcal{C}_{k_1,k_2}$ are denoted with a single index $\mathcal{C}_k$, with $k = 1, \ldots, 200$.

From figure 5.15(b), we found that the cluster distribution is very uniform during the transient states and the post-transient states. The six exact solutions are well classified, and much fewer clusters are used for the limit cycles. The three steady solutions being divided into $\mathcal{C}_1$, $\mathcal{C}_{111}$ and $\mathcal{C}_{161}$ separately. The three limit cycles are identified with three blocks of oscillating labels, from $\mathcal{C}_{61}$ to $\mathcal{C}_{70}$, from $\mathcal{C}_{81}$ to $\mathcal{C}_{90}$, and from $\mathcal{C}_{101}$ to $\mathcal{C}_{110}$. Even the transient states between the limit cycles can also be identified, with two blocks from $\mathcal{C}_{71}$ to $\mathcal{C}_{80}$, and from $\mathcal{C}_{91}$ to $\mathcal{C}_{100}$. The transient trajectories

starting with three different steady solutions are also well separated. The transition matrix in figure 5.15(d) are almost full of definite transitions, which shows clearer transient dynamics than the figure 5.15(c). The global matrix keeps the local dynamics for each cluster $\mathcal{C}_{k_1}$ in the first layer $\mathcal{L}_1$, as shown in each small block of 10 clusters, and ensembles them together. The transitions from block to block have a much lower probability, comparing with the cycling transitions within the block, which makes them hard to see in the figure. However, this kind of transition between the blocks should be also definite, as shown in § 5.4.2 for the hierarchical model in the first layer. Analogously, the multiscale problem of the dynamics in different layers also exists in the global residence time matrix, which makes the small scale terms unseeable. Therefore, we suggest analyzing the slow-varying mean flow and the local dynamics separately in different layers, as in § 5.4.2, to avoid the influence of different scales.

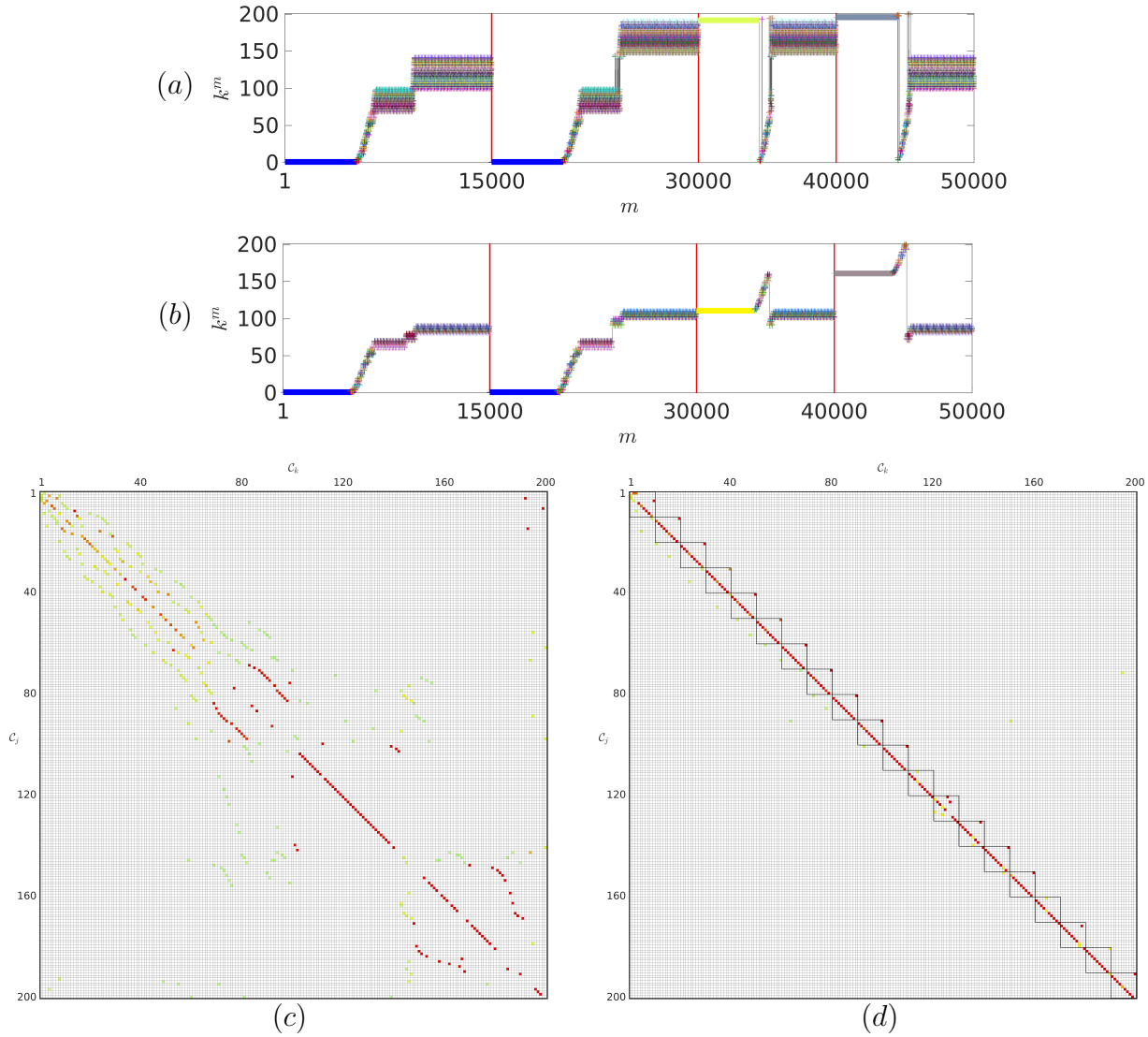


Figure 5.15: Cluster-based analysis at $Re = 80$ for the fluctuating flow: transition illustrated with cluster label and transition matrix of the standard network model (a, c) and the hierarchical network model in the second layer (b, d), as in figure 5.10.

5.4.4 Hierarchical network model for the quasi-periodic dynamics at Re=105

At $Re = 105$, the flow dynamics is quasi-periodic (Deng *et al.*, 2020). The inner jet oscillations are modulated at a non-commensurate low frequency. The flow dynamics considered in the first cluster

layer is low-pass filtered (§ 5.4.4). The basic limit cycle associated with the dominant frequency ω_c of the vortex shedding is clustered and analysed in layer 2 (§ 5.4.4). The low-frequency modulations of the vortex shedding are further described in layer 3 (§ 5.4.4).

Hierarchical network model in Layer 1

The frequency of the coherent component $\tilde{\boldsymbol{u}}$ is $f_c = 0.1172$. The $K_1 = 11$ clusters are used on the low-pass filtered data set, following algorithm 1. The non-trivial transitions are shown in the two-dimensional subspace $[\gamma_1, \gamma_2]^{\mathrm{T}}$ of figure 5.16.

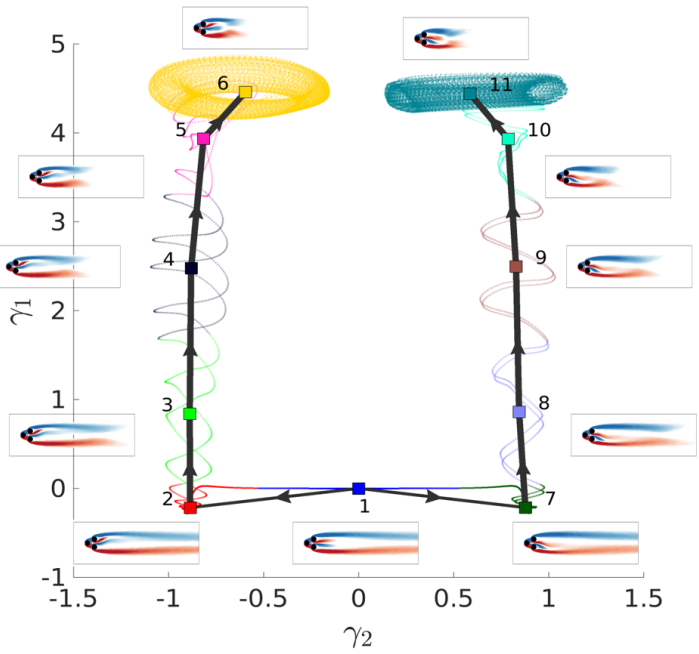


Figure 5.16: Graph of transitions between clusters in layer 1 at $Re = 105$, displayed as in figure 5.9.

The two attractors belong to clusters $\mathcal{C}_6$ and $\mathcal{C}_{11}$. The symmetric and asymmetric steady solutions belong respectively to clusters $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{C}_7$. The transient dynamics observed at $Re = 105$ are different from those identified in figure 5.9 for $Re = 80$. Figure 5.16 shows four trajectories, initiated from mirror-conjugated initial conditions close to the symmetric and asymmetric steady solutions:

$$
\begin{aligned}
&\textit{Trajectory}\,1: \quad \mathcal{C}_1\ (\boldsymbol{u}_s) \to \mathcal{C}_2(\boldsymbol{u}_s^+) \to \ldots \to \mathcal{C}_6(\bar{\boldsymbol{u}}^+);\\
&\textit{Trajectory}\,2: \quad \mathcal{C}_1\ (\boldsymbol{u}_s) \to \mathcal{C}_7(\boldsymbol{u}_s^-) \to \ldots \to \mathcal{C}_{11}(\bar{\boldsymbol{u}}^-);\\
&\textit{Trajectory}\,3: \quad \mathcal{C}_2(\boldsymbol{u}_s^+) \to \ldots \to \mathcal{C}_6(\bar{\boldsymbol{u}}^+);\\
&\textit{Trajectory}\,4: \quad \mathcal{C}_7(\boldsymbol{u}_s^-) \to \ldots \to \mathcal{C}_{11}(\bar{\boldsymbol{u}}^-).
\end{aligned}
$$

The state trajectories start from the symmetric steady solution to one of the two asymmetric steady solutions, then converge to the corresponding attracting torus. The two trajectories on the left side have a significant phase delay, while the two on the right side are almost in the same phase. This phase difference is a random function that depends on the initial condition. The two tori also look different because the feature vectors associated with $[\gamma_1, \gamma_2]^{\mathrm{T}}$ are asymmetrical from the multidimensional scaling.

In figure 5.17(a), the evolution of the cluster label of the snapshots illustrates four irreversible transient dynamics. The transition matrix in figure 5.17(b) exhibits two red diagonals associated with trajectories to the final state, and two yellow elements associated with the initial state starting close to the symmetric steady solutions in cluster $\mathcal{C}_1$. The transition from $\mathcal{C}_1$ to clusters $\mathcal{C}_2$ and $\mathcal{C}_7$ each have $1/2$ probability. The two attracting clusters $\mathcal{C}_6$ and $\mathcal{C}_{11}$, associated with the attractors

of the system, have no transition to any other clusters, as all the terms in the sixth and eleventh columns $P_{j\,6} = P_{j\,11} = 0,\ \forall j$. In the residence time matrix of figure 5.17(c), the filled black circles indicates three typical clusters $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{C}_7$, which correspond to the vicinity of the three steady solutions. All the terms in the sixth and eleventh columns $T_{j\,6}, T_{j\,11}$ are 0 $\forall j$, which means that the residence time is infinite, as expected for attractors.

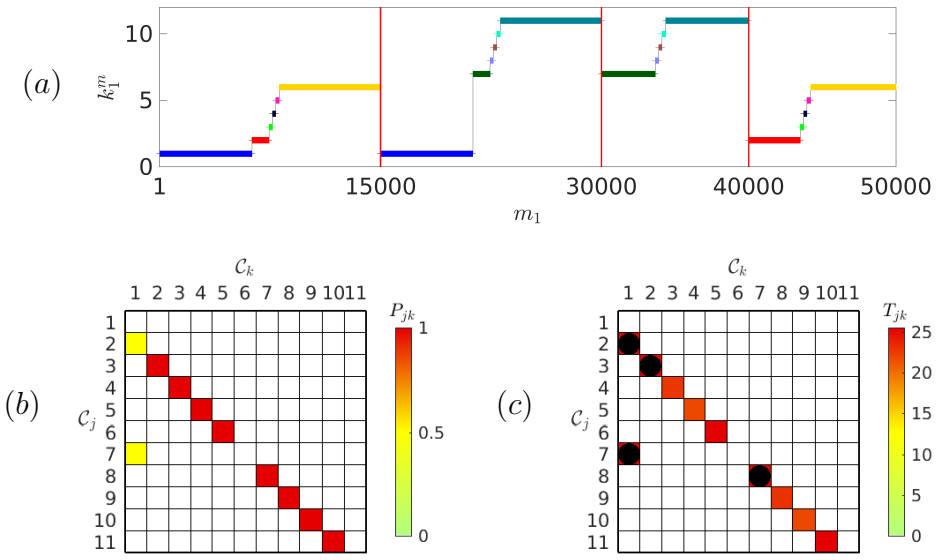


Figure 5.17: Cluster-based analysis at $Re = 105$ in layer 1: (a) transition illustrated with cluster label, (b) transition matrix, (c) residence time matrix, displayed as in figure 5.10.

Hierarchical network model in Layer 2

We focus on the permanent regime in the cluster $\mathcal{C}_6$, and apply the sub-division clustering algorithm 2, which results in $K_2 = 10$ subclusters in layer $\mathcal{L}_2$.

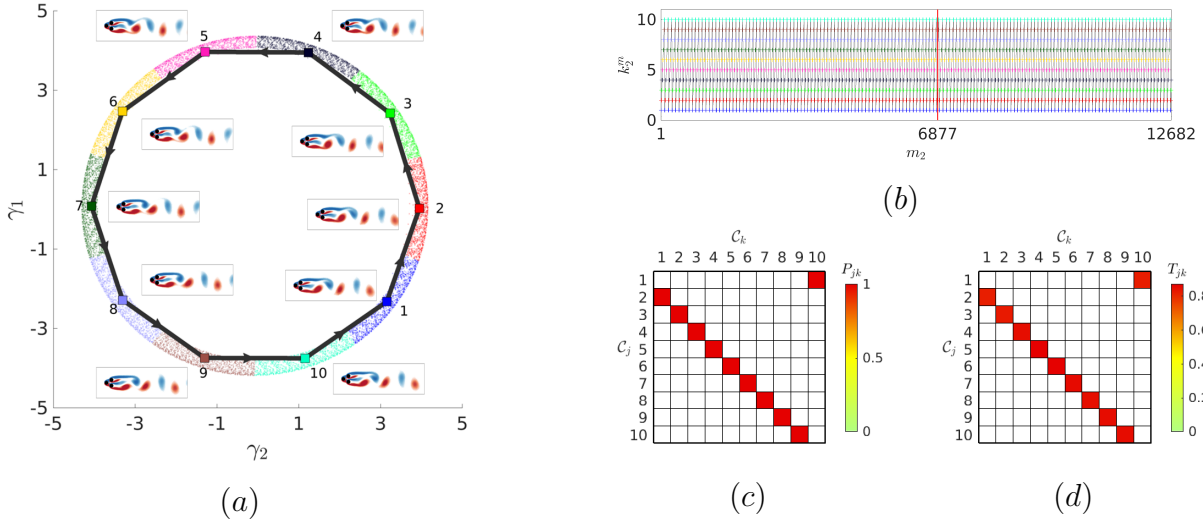


Figure 5.18: Cluster-based analysis in layer 2 at $Re = 105$ for $\mathcal{C}_6$: (a) graph of non-trivial transitions between clusters, as in figure 5.9, (b) transition illustrated with cluster label, (c) transition matrix, (d) residence time matrix, as in figure 5.10. Two trajectories pass through $\mathcal{C}_6$ in the parent layer, one with $m_2 = 1, \ldots, 6877$ and another with $m_2 = 6878, \ldots, 12682$.

As illustrated in figure 5.18(a), a closed orbit between the 10 clusters is found. Figure 5.18(b) shows two transient trajectories pass through $\mathcal{C}_6$, the original snapshot index is $m = 8124, \ldots, 15000$

and $m = 44196, \ldots, 50000$, involving the asymptotic regime of an attractor. The clusters in this closed orbit have a clear transition rule, as is evidenced by the transition matrix of figure 5.18(c). The sum of the residence times is 8.45 for the cycle of $\mathcal{C}_3 \to \mathcal{C}_4 \to \ldots \to \mathcal{C}_{10} \to \mathcal{C}_3$ from the residence time matrix in figure 5.18(d).

Hierarchical network model in Layer 3

Based on the detected cycle found in layer $\mathcal{L}_2$, the entering snapshots of one subcluster can be used to sample the quasi-periodic regime. The snapshots recurrently enter into each cluster of layer $\mathcal{L}_2$. The entering states, defined by T^m_{o,k_1,k_2} in Eq.(5.20), are the entry in the clusters. The entering snapshots of a given cluster can be considered as hits in a "Poincaré section" Guckenheimer & Holmes (2013).

The clustering algorithm in the layer $\mathcal{L}_3$ is applied to the snapshots of all the Poincaré sections. We still apply the sub-division clustering algorithm 2, with T^m_{o,k_1,k_2} as characteristic function.

We consider the entry in cluster $\mathcal{C}_{6,10}$ for illustration, and build a network model in the third layer $\mathcal{L}_3$. In the following, the cluster symbol in the first two layers $\mathcal{C}_{k_1=6,k_2=10,k_3}$ is omitted.

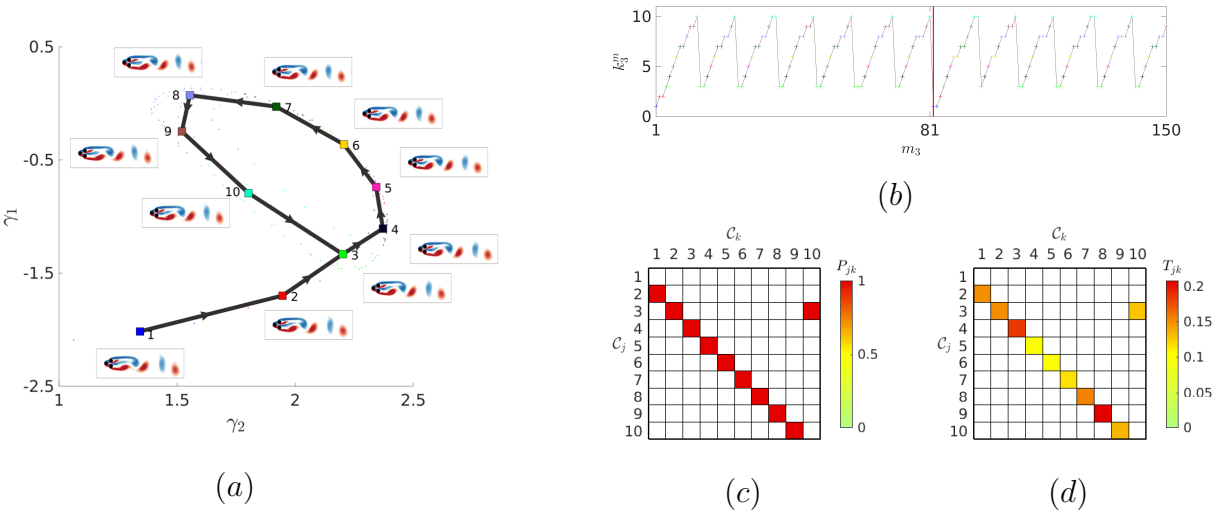


Figure 5.19: Cluster-based analysis in layer 3 at $Re = 105$ for $\mathcal{C}_{6,10}$: (a) graph of non-trivial transitions between clusters, as in figure 5.9, (b) transition illustrated with cluster label, (c) transition matrix, (d) residence time matrix, as in figure 5.10. Two trajectories are detected in the "Poincaré section" of $\mathcal{C}_{6,10}$, one with $m_3 = 1, \ldots, 81$ and another with $m_3 = 82, \ldots, 150$.

As shown by the graph of non-trivial transitions shown in figure 5.19(a), there exists a cycle $\mathcal{C}_3 \to \mathcal{C}_4 \to \ldots \to \mathcal{C}_{10} \to \mathcal{C}_3$. The periodically changing cluster label in figure 5.19(b) indicates the existence of a recurrent dynamics. The original cluster index corresponds to two sets of discrete snapshots in the Poincaré section, with interval approximate to the periodic detected in the $\mathcal{L}_2$. From the transition matrix of figure 5.19(c), the two clusters $\mathcal{C}_1$ and $\mathcal{C}_2$ are not part of the cycle, but they form the transient part of the dynamics, before entering the cycle. This indicates that the low frequency modulations only start after the high-frequency oscillations have started in the second layer. In other words, the low frequency does not exist during the building process of the vortex shedding, but appears after the vortex shedding has developed to a certain degree.

From the residence time matrix of figure 5.19(d), the number of snapshots that belong to the cycle in the Poincaré section, determined by averaging over multiple trajectories, is 11.41. The clustering analysis in the second layer $\mathcal{L}_2$ indicates that the trajectories periodically hit the Poincaré section with 8.45, by summing up the elements in the blocks from $\mathcal{C}_3$ to $\mathcal{C}_{10}$ in figure 5.19(d). Hence, the resulting period of the cyclic process, considering both frequencies, is around 96.41, which is very close to the real period of 97.10 determined from the DNS.

Dynamics reconstruction of the hierarchical network model at $Re = 105$

Figure 5.20 shows the autocorrelation function of the DNS and the HiCNM in the three layers. The autocorrelation function of the DNS identifies the two dominant frequencies of the dynamics.

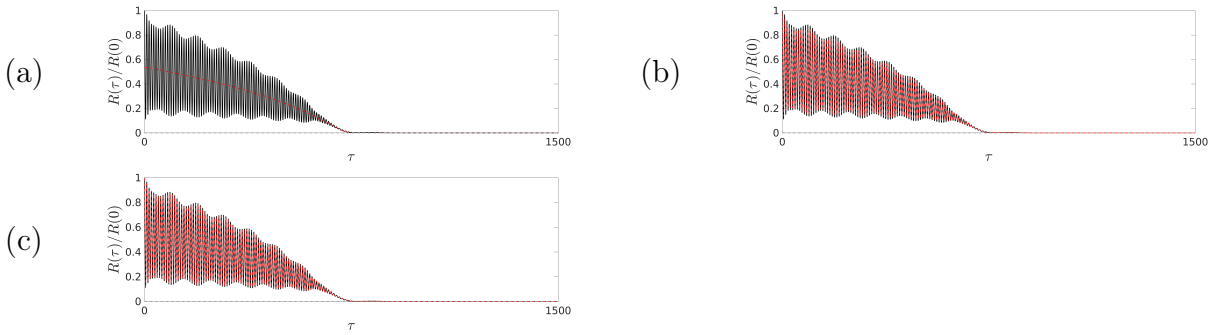


Figure 5.20: Autocorrelation function for $\tau \in [0, 1500)$ from DNS (black solid line) and the hierarchical network model (red dashed line) in the two layers: (a) $\mathcal{L}_1$, (b) $\mathcal{L}_2$, and (c) $\mathcal{L}_3$, at $Re = 105$

In layer $\mathcal{L}_1$, no oscillation can be identified, and the RMSE of the autocorrelation function is $R^1_{\text{rms}} = 22.20$. In layer $\mathcal{L}_2$, the autocorrelation function of the model matches well with the high-frequency oscillations. The low-frequency oscillations can be also found, but the amplitude does not fit well. The error is $R^2_{\text{rms}} = 1.52$, which is good enough for the accuracy. In layer $\mathcal{L}_3$, the amplitude of the low-frequency oscillations can be better reproduced. The error is further reduced to $R^3_{\text{rms}} = 0.77$ with higher accuracy.

5.4.5 Hierarchical network model at Re=130

At $Re = 130$, the asymptotic dynamics is chaotic. We apply the clustering algorithm 1 first to the low-pass filtered data set (§ 5.4.5), before considering some typical flow regimes in the subclusters (§ 5.4.5).

Hierarchical network model in layer 1

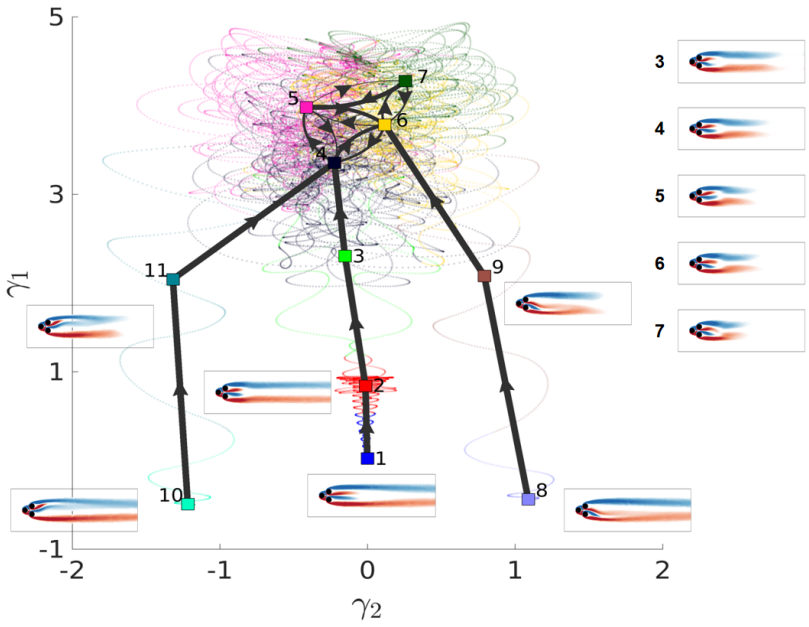


Figure 5.21: Graph of transitions between clusters in layer 1 at $Re = 130$, displayed as in figure 5.9.

In the first layer, the filtered data set is used to analyze the mean-field dynamics with $K_1 = 11$ clusters. The detected frequency associated with the coherent component $\tilde{\boldsymbol{u}}$ is $f_c = 0.1225$. The clustering algorithm 1 is applied to the data set, and the non-trivial transitions are shown in the two-dimensional subspace $[\gamma_1, \gamma_2]^{\mathrm{T}}$ of figure 5.21. Four trajectories are found, each issued from one of the three steady solutions. The symmetric steady solution and the two asymmetric steady solutions respectively belong to clusters $\mathcal{C}_1$, $\mathcal{C}_8$, and $\mathcal{C}_{10}$. These clusters evolve through $\mathcal{C}_1 \to \mathcal{C}_2 \to \mathcal{C}_3$, $\mathcal{C}_8 \to \mathcal{C}_9$ and $\mathcal{C}_{10} \to \mathcal{C}_{11}$ before entering into the same chaotic cloud, consisting of the remaining clusters.

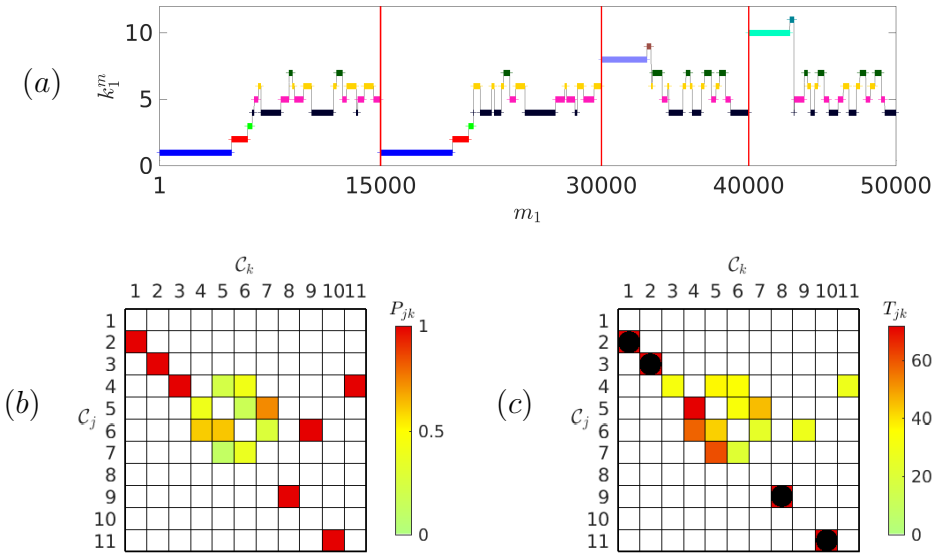


Figure 5.22: Cluster-based analysis at $Re = 130$ in layer 1: (a) transition illustrated with cluster label, (b) transition matrix, (c) residence time matrix, displayed as in figure 5.10.

There is no obvious periodic block of oscillating dynamics in the transient dynamics, as illustrated by the cluster labels of figure 5.22(a). However, the initial destabilizing process is characterized by a very long residence time in the clusters to which the steady solutions belong. From the transition matrix of figure 5.22(b), three transitions leading to the chaotic region with 100% probability, as P_{34}, P_{96} and $P_{11\,4}$. Each cluster in the chaotic cloud has at least two possible destinations, with nearly equal probability. The hidden transition dynamics for this chaotic regime will be analized in the next layer. The residence time matrix of figure 5.22(c) shows four black filled circles associated with clusters $\mathcal{C}_1$, $\mathcal{C}_8$, $\mathcal{C}_{10}$ and $\mathcal{C}_2$. The steady solutions belong to the first three, while the symmetric limit cycle belongs to $\mathcal{C}_2$, as it will become clear in the next sections.

Hierarchical network model in layer 2

In the second layer $\mathcal{L}_2$, we focus on the clusters associated with four typical states detected in layer $\mathcal{L}_1$: the destabilizing state from the symmetric steady solution in $\mathcal{C}_1$, the destabilizing state from the (upward) asymmetric steady solution in $\mathcal{C}_{10}$, the transient state with long residence time before chaos in $\mathcal{C}_2$, and the chaotic state in the group of clusters $\mathcal{C}_4, \ldots, \mathcal{C}_7$.

The destabilizing state in the cluster $\mathcal{C}_1$

Cluster $\mathcal{C}_1$ gathers snapshots in the initial stage of the instability starting from the symmetric steady solution. In the second layer $\mathcal{L}_2$, the snapshots of $\mathcal{C}_1$ are dispatched into $K_2 = 10$ subclusters $\mathcal{C}_{1,k_2}$, with $k_2 = 1, \ldots, K_2$. Figure 5.23 shows the transient trajectories with the centroids in the $[\gamma_1, \gamma_2]^{\mathrm{T}}$ plane.

Similar to figure 5.13(a), the snapshots of figure 5.23(a) form two diverging trajectories spiraling out of the center $[0, 0]$. A loop is formed between the subclusters $\mathcal{C}_{1,1}$ and $\mathcal{C}_{1,2}$ in the inner zone.

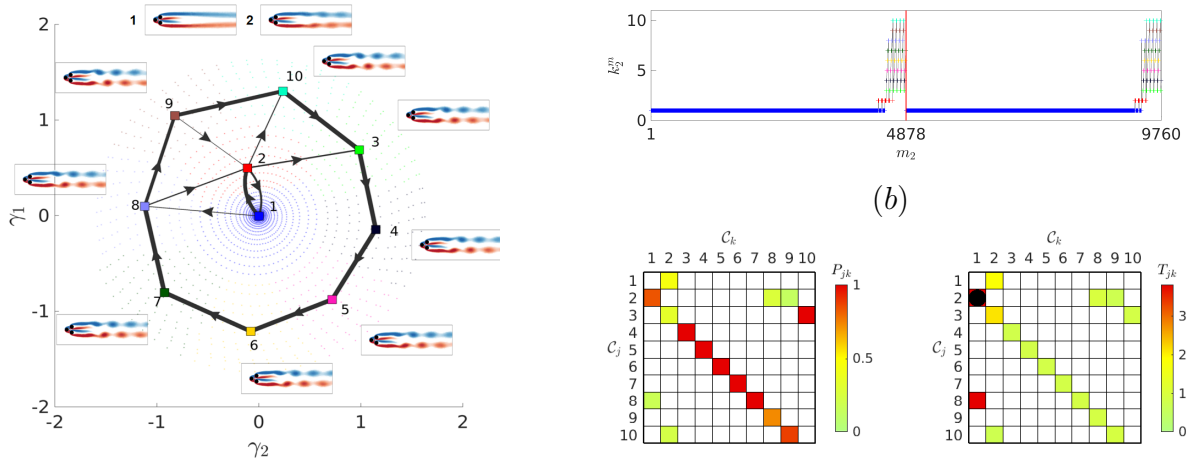


Figure 5.23: Cluster-based analysis in layer 2 at $Re = 130$ for $\mathcal{C}_1$: (a) graph of non-trivial transitions between clusters, as in figure 5.9, (b) transition illustrated with cluster label, (c) transition matrix, (d) residence time matrix, as in figure 5.10. Two trajectories pass through $\mathcal{C}_1$ in the parent layer, one with $m_2 = 1, \ldots, 4878$ and another with $m_2 = 4879, \ldots, 9760$.

In the outer zone, a cycle appears with the periodic trajectory $\mathcal{C}_{1,3} \to \ldots \to \mathcal{C}_{1,10} \to \mathcal{C}_{1,3}$. Figure 5.23(b) shows two trajectories leaving $\mathcal{C}_{1,1}$. The original snapshot index is $m = 1, \ldots, 4878$ and $m = 15001, \ldots, 19882$, corresponding to the initial stage of the instability.

The transition matrix in figure 5.23(c) corroborates this periodic cycle. The black filled circle in figure 5.23(d) marks out the transition with a long-residence time, due to the unstable center that belong to cluster $\mathcal{C}_{1,1}$.

The transitions $\mathcal{C}_{1,1} \to \mathcal{C}_{1,8}$, $\mathcal{C}_{1,2} \to \mathcal{C}_{1,3}$ and $\mathcal{C}_{1,2} \to \mathcal{C}_{1,10}$ correspond to the departing dynamics out of the inner zone, due to the development of the instability. We also note the returning transitions $\mathcal{C}_{1,8} \to \mathcal{C}_{1,2}$ and $\mathcal{C}_{1,9} \to \mathcal{C}_{1,2}$. However, the latter do not mean that the flow actually returns back to the destabilizing center, as both trajectories are spiralling out of the center. This confusing result comes from the clustering process. The edge between the inner and outer zones is not well defined, due to the varying density distribution of snapshots along the arms of the spirals. Cluster $\mathcal{C}_{1,2}$ overlaps the outer zone, to the difference of cluster $\mathcal{C}_{1,1}$, which fully belong to the inner zone. If we ignore the loop in the inner zone and merge $\mathcal{C}_{1,1}$ and $\mathcal{C}_{1,2}$, it shows a dynamical evolution from one cluster to the cycle of a group of clusters.

The destabilizing state in cluster $\mathcal{C}_{10}$

Cluster $\mathcal{C}_{10}$ contains the trajectory spiraling out from the upward-deflected asymmetric steady solution, as shown in figure 5.24(a). The snapshots are dispatched into $K_2 = 10$ of subclusters $\mathcal{C}_{10,k_2}$, with $k_2 = 1, \ldots, K_2$. Together with figure 5.24(b), it shows a one-way transition, departing from the unstable center with a sparse spiral. The original snapshot index is $m = 40001, \ldots, 42782$ for the initial stage of the destabilisation from one asymmetric steady solution.

The graph of figure 5.24(a) is different from the graphs of figure 5.13(a) and figure 5.23(a), the latter being associated with the symmetric steady solution. The flow destabilisation from the asymmetric steady solution develops faster than from the symmetric steady solution, as illustrated by the linear growth rates $\sigma_{sym} = 0.032$ and $\sigma_{asym} = 0.106$ of the respective pairs of unstable eigenmodes. As a result, the distribution of snapshots is sparser in figure 5.24(a) than in figure 5.23(a). As indicated by figure 5.24(b), 93.2% of the 2782 snapshots of cluster $\mathcal{C}_{10}$ belongs to subcluster $\mathcal{C}_{10,1}$. The flow quickly travels through all the remaining clusters in only 188 snapshots.

The centroids of figure 5.23(a) and figure 5.24(a) have two main differences: (i) The inner jet is symmetric in the centroids of figure 5.23(a) while it is deflected upwards in the controids

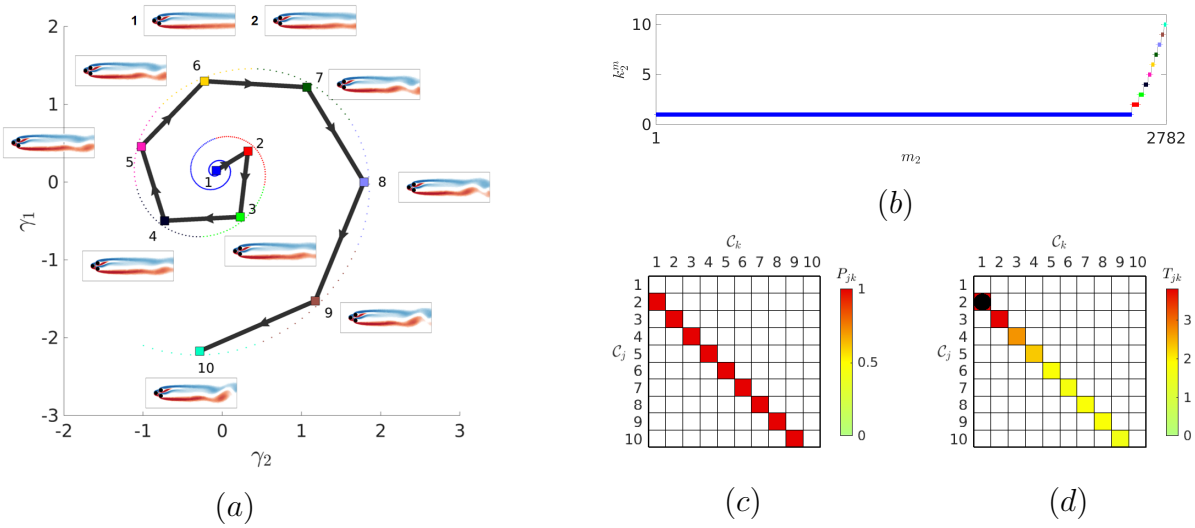


Figure 5.24: Cluster-based analysis in layer 2 at $Re = 130$ for $\mathcal{C}_{10}$: (a) graph of non-trivial transitions between clusters, as in figure 5.9, (b) transition illustrated with cluster label, (c) transition matrix, (d) residence time matrix, as in figure 5.10. A sole trajectorie pass through $\mathcal{C}_{10}$ in the parent layer with $m_2 = 1, \ldots, 2782$.

of figure 5.24(a); (ii) the von Kármán street of vorticies of figure 5.23(a) exhibits positive and negative vortices well apart from each other in the y-axis. In figure 5.24(a), the positive and negative vortices are of a larger strength and adjacent to the x-axis, together with a longer shear layer.

Transient regime before chaos in cluster $\mathcal{C}_2$

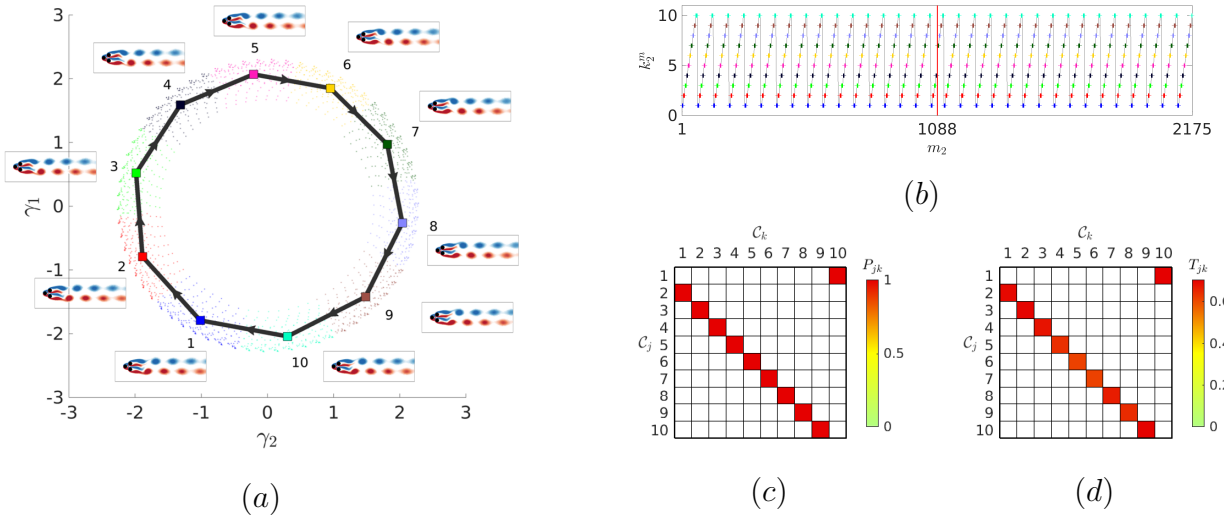


Figure 5.25: Cluster-based analysis in layer 2 at $Re = 130$ for $\mathcal{C}_2$: (a) graph of non-trivial transitions between clusters, as in figure 5.9, (b) transition illustrated with cluster label, (c) transition matrix, (d) residence time matrix, as in figure 5.10. Two trajectories pass through $\mathcal{C}_2$ in the parent layer, one with $m_2 = 1, \ldots, 1088$ and another with $m_2 = 1089, \ldots, 2175$.

Cluster $\mathcal{C}_2$ contains two transient trajectories which connect the symmetric steady solution in $\mathcal{C}_1$ to the chaotic cloud, with the original snapshot index $m = 4879, \ldots, 5966$ and $19883, \ldots, 20969$.

Figure 5.25(b) shows that the flow periodically travels through the ten subclusters in the child layer $\mathcal{L}_2$ of cluster $\mathcal{C}_2$. As for the centroids of figure 5.25, the alley of vortices do not cross the x-axis. The closed-transitions $\mathcal{C}_1 \to \ldots \to \mathcal{C}_{10} \to \mathcal{C}_1$ is further evidenced in the transition matrix

of figure 5.25(c). The residence times of figure 5.25(d) are rather uniform and the averaged period is 6.62.

The chaotic state in the group of clusters $\mathcal{C}_4, \ldots, \mathcal{C}_7$

When dealing with the chaotic dynamics, considering each cluster $\mathcal{C}_4, \ldots, \mathcal{C}_7$ separately is doable. However, instead of building a network model for each cluster separately, we can build an overall model for these four clusters in the chaotic regime. Therefore, for analyzing the chaotic dynamics, we consider all the clusters that belong to the chaotic cloud in layer $\mathcal{L}_1$ as a whole.

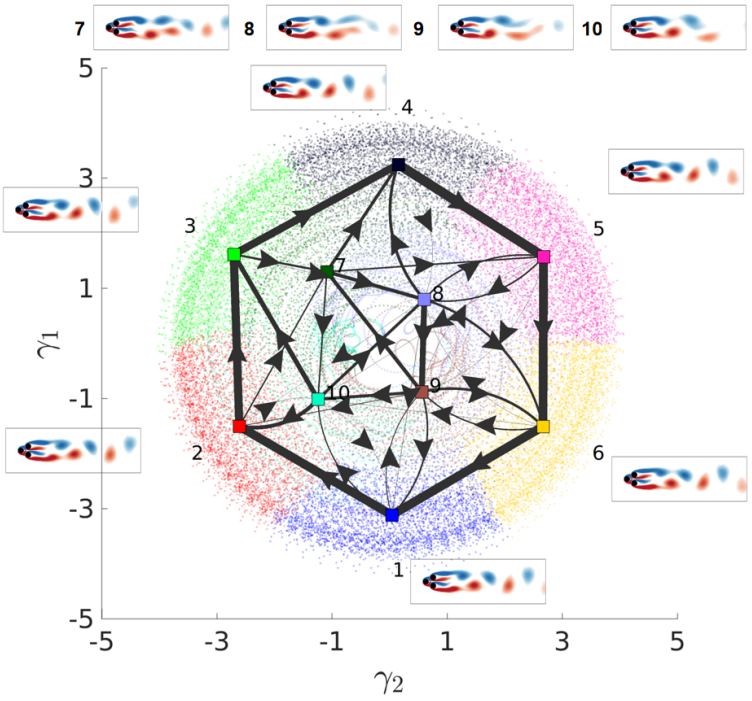


Figure 5.26: Graph of transitions between clusters in layer 2 at $Re = 130$ for the chaotic clusters of $\mathcal{L}_1$, displayed as in figure 5.9.

From figure 5.22(a), all four trajectories will reach the same chaotic attractor described by the chaotic clusters, with the original snapshot index $m = 6281, \ldots, 15000, 21301, \ldots, 30000, 33365, \ldots, 40000$ and $43068, \ldots, 50000$. In figure 5.26, the closed orbit of the clusters $\mathcal{C}_1 \to \ldots \to \mathcal{C}_6 \to \mathcal{C}_1$ is formed, with a relatively high probability of transition between the successive clusters. The flow field of the centroids of the first six clusters form a complete cycle of vortex shedding. This is interesting, as it reminds the periodic and quasi-periodic dynamics respectively observed at $Re = 80$ and $Re = 105$. The periodic block of the transition matrix of figure 5.27(b), from $\mathcal{C}_1$ to $\mathcal{C}_6$, corroborates the existence of the periodic dynamics. The centroids of the clusters in the cycle present a similar structure of coherence, as can be seen in figure 5.26. There are also other possible transitions from clusters on this orbit to the remaining clusters with much smaller probability. The residence times shown in figure 5.27(c) for each cluster on the orbit are uniform, and the averaged period along the complete cycle is 7.78 by summing up the elements in the block from $\mathcal{C}_1$ to $\mathcal{C}_6$.

The remaining clusters $\mathcal{C}_7, \ldots, \mathcal{C}_{10}$ have multiple destinations with quasi-random possibilities. Even though the probabilities of these random transitions are small, they contribute to the chaotic dynamics of the flow field, with recurrent transitions $\mathcal{C}_1 \to \mathcal{C}_{10}$, $\mathcal{C}_9 \to \mathcal{C}_1$, and so on. The flow fields of the associated centroids are shown in figure 5.26. Their structure looks like distortions of the vortex shedding cycle formed by the first six clusters. The network model indicates that the fully chaotic state still contains a main cycle of clusters associated with a periodic vortex shedding, together with the random jumping to the clusters associated with a stochastic disorder

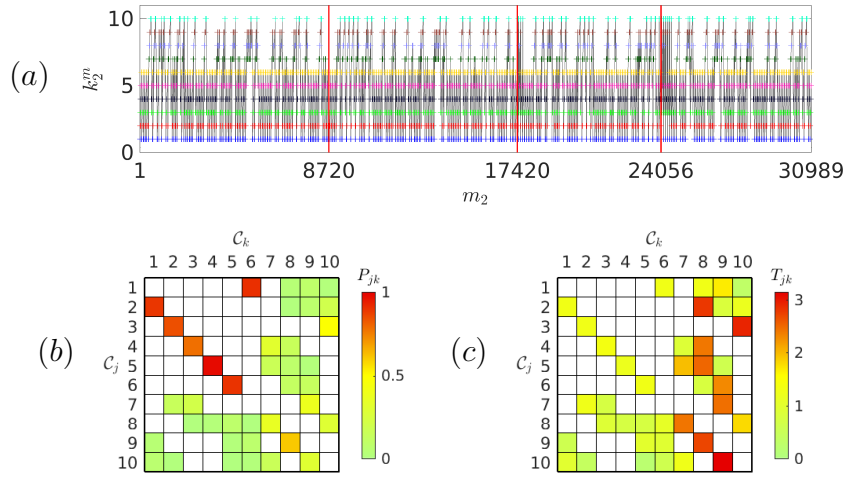


Figure 5.27: Cluster-based analysis in layer 2 at $Re = 130$ for the chaotic clusters of $\mathcal{L}_1$: (a) transition illustrated with cluster label, (b) transition matrix, (c) residence time matrix, as in figure 5.10. All four trajectories reach the chaotic clusters,with $m_2 = 1, \ldots, 8720, 8721, \ldots, 17420, 17421, \ldots, 24056$ and $24057, \ldots, 30989$.

in the wake. In this case, the transition matrix can be used to build a stochastic model, as shown in appendix 5.D.

Dynamics reconstruction of the hierarchical network model at $Re = 130$

Figure 5.28 shows the autocorrelation function of the DNS and the HiCNM in the two layers. We stop the hierarchical modelling in layer $\mathcal{L}_2$, as both the transient and chaotic dynamics can be fairly reproduced with a limited number of clusters. The autocorrelation function of the DNS

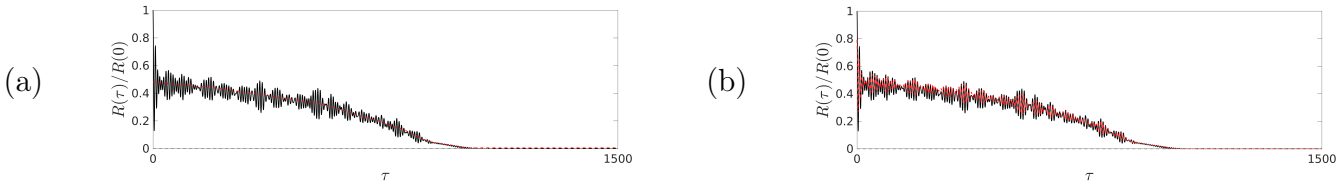


Figure 5.28: Autocorrelation function for $\tau \in [0, 1500)$ from DNS (black solid line) and the the hierarchical network model (red dashed line) in the two layers: (a) $\mathcal{L}_1$ and (b) $\mathcal{L}_2$, at $Re = 130$.

shows chaotic oscillations with a dominant frequency. In layer $\mathcal{L}_1$, no oscillation can be identified, and the RMSE of the autocorrelation function is $R^1_{\text{rms}} = 6.94$. In layer $\mathcal{L}_2$, $R^2_{\text{rms}} = 3.42$. The autocorrelation function of the model matches well with the dominant frequency of the oscillations, but the $\hat{R}_{\mathcal{L}_2}(\tau)$ value can hardly match.

5.5 Conclusion

We have proposed a data-driven modelling methodology, which consists of hierarchical clustering and network modelling on top of the cluster-based reduced-order model (CROM). The hierarchical structure is physically consistent with the weakly nonlinear model derived from the mean-field consideration. The flow field is decomposed into a hierarchy of components, namely the slowly varying mean-flow field, the dominant vortex shedding and the secondary components. The resulting hierarchical cluster-based reduced-order model (HiCROM) can automate the modelling process based on the representative states and systematically trace the flow dynamics on multiple scales, involving multiple frequencies and multiple attractors. The cluster-based hierarchical network

model (HiCNM) presented in this work is a HiCROM using the directed network to describe the non-trivial transitions between clusters. Based on the classical CNM, we derived the HiCNM for the transient and post-transient dynamics of the two-dimensional incompressible "fluidic pinball", characterized by multiple invariant sets and dynamics, for different Reynolds numbers.

The considered data set consists of snapshots of the velocity field computed from the direct numerical simulation starting with different initial conditions, which refer to four trajectories: two mirror-conjugated trajectories starting in the vicinity of the symmetric steady solution and the two others starting from the two asymmetric steady solutions. At the considered Reynolds numbers, all the steady solutions are unstable. In this sense, the data set contains the transient and post-transient dynamics involving all the invariant sets of the system. The hierarchical modelling is based on the hierarchical clustering under the mean-field consideration. In the first layer, the first k-means++ clustering algorithm is applied to the low-pass filtered data and partitions the snapshots into different clusters, as in algorithm 1. The flow field fluctuations are responsible for the nonlinear mean-field distortions through the Reynolds stress. As a result, the snapshots in different clusters exhibit different states of slowly changing mean-flow field. The network model in the first layer focuses on the global transitions between different invariant sets. The clustering result in the first layer will guide the clustering process in the second layer. The second clustering process in algorithm 2 partitions the snapshots in the same cluster again into sub-clusters, according to the original data. Based on the sub-clusters, a new network model can be built with a better interpolation of the local dynamics. The clustering process is similar to the divisive hierarchical clustering, which can continue until each snapshot is a cluster. However, the number of layers depends on the number of characteristic scales in the system, such as the number of coherent frequencies or the fast and slow terms. Hence, two or three layers will be enough to extract the transient and post-transient dynamics out of multiple invariant sets and multiple frequencies in our case.

At $Re = 80$, the six invariant sets were well identified in the first layer of the HiCNM, including the dynamics around three unstable steady solutions ($\mathcal{C}_1$, $\mathcal{C}_{12}$, $\mathcal{C}_{17}$), one unstable symmetric limit cycle ($\mathcal{C}_7$) and two stable asymmetric limit cycles ($\mathcal{C}_9$, $\mathcal{C}_{11}$). The transient dynamics between the multiple invariant sets, and the temporal development of the degrees of freedom associated with the static symmetry breaking, are identified by the model in the first layer. We further presented the model in the second layer involving the three exact solutions: the destabilization of the symmetric steady solution in cluster $\mathcal{C}_1$, the dynamics around the symmetric limit cycle in cluster $\mathcal{C}_7$, and the permanent regime on the asymmetric limit cycle in cluster $\mathcal{C}_9$. Compared to a CNM with the same number of clusters, the HiCNM preserves the advantage of automatable modelling, optimizes the cluster distribution, and makes it human-interpretable.

For the quasi-periodic flow regime at $Re = 105$, the first two layers are identical to the case at $Re = 80$. The HiCNM in the first layer identified the different invariant sets, and the model in the second layer described the local dynamics on the invariant sets. We further introduced the third layer to characterize the new coherent structures at low frequency. The sub-division clustering in algorithm 2 was applied on the entering snapshots of the cluster $\mathcal{C}_{6,10}$. The low frequency was successfully identified, while the centroids identified the tiny changes of the oscillating jet in the near wake.

At $Re = 130$, three unstable steady solutions and one chaotic attracting set have been caught in the first layer. The chaotic zone was divided into several clusters. In the second layer, we focused on the local structures around the invariant sets. We determined the dynamics of the initial transients from the unstable steady solutions in $\mathcal{C}_1$ and $\mathcal{C}_{10}$. Besides, to preserve the continuity of the data, the second clustering process was applied to the group of clusters in the chaotic regime $\mathcal{C}_{k_1}$, $k_1 = 4, \ldots, 7$. An unstable cycle was identified for the chaotic regime, characterized by random transitions from and to the chaotic clusters with low probability.

Compared to other reduced-order modelling strategies, HiCROM inherits the excellent recognition performance of classical CROM, and provides a universal modelling strategy for identifying

transient and post-transient dynamics in a self-supervised manner. Multiple transient dynamics can be considered at the same time, which gives a global view of the trajectories between the different invariant sets. Thus, it provides a better understanding of the complex flow dynamics for the multiscale, multi-frequencies and multi-attractors problem. To summarize, the HiCNM applied in this work has the following advantages comparing to classical CNM:

(i) A more robust clustering result with the hierarchical modelling under mean-field consideration, and a better distribution of the clusters.
(ii) Better ability to identify the topology on the multi-attractor and multiscale problem. HiCNM identifies transient trajectories between different invariant sets, and locally constructs new CNMs for the different invariant sets if necessary.
(iii) No need to find a good compromise between the resolution (the number of clusters) and the network complexity. All clustering algorithms use 10-20 clusters, and the number of clusters depends on the accurate representation needs of the scale involved in the layer. The number of layers depends on the number of characteristic scales in the system. The clusters of different scales or characteristics can be systematically distributed in multiple layers with a clear hierarchy.
(iv) The subdivision is flexible according to the actual needs. We can freely choose one or more object clusters that need to be divided in the next layer to preserve the time continuity and the local characteristics of the data.

The price is the need for a physical intuition guiding the hierarchical clustering and modelling by adjusting the design parameters. As exemplified in this work, the HiCNM at $Re = 80$ is fully automated. After reaching the second layer, the HiCNM can well identify the mean flow and the coherent structure with a single frequency. The deeper layers cannot identify other meaningful dynamics. At $Re = 105$, the third layer can identify the secondary frequency with the entering sanpshots in one cluster in the second layer. At $Re = 130$, we merge the asymptotic chaotic clusters in the first layer together, and build a model for them in the second layer to maintain their dynamic continuity. This sacrifices automation but results in a better dynamic representation. We foresee other data-driven methods for these decisions in future work to promote the automation of the hierarchical network model.

In summary, HiCNM provides a flexible and automatable cluster-based modelling framework for complex flow dynamics, and shows its ability and applicability to identify transient and post-transient, multi-frequency, multi-attractor dynamics. Since the Reynolds decomposition under the mean-field consideration is common for fluid flows, the hierarchical strategy should be extendable to other flows. Especially for cases that require high-resolution analysis with a large number of clusters, HiCNM greatly simplifies the complexity of the analysis and improves its interpretation. For high turbulence without frequency boundaries, such as isotropic turbulence, it is reasonable to use a network model first to check whether there is a grouping relationship between clusters, before deciding whether a HiCNM needs to be constructed. HiCNM is promising for a variety of potential applications, such as topology identification of the state space for complex dynamics, recognition and analysis of the temporal evolution of degrees of freedom associated with different types of instabilities (Hopf, pitchfork, etc.), and feature extraction of the dynamical structure when different spatial/temporal scales are involved. An alternative direction of this work is the HiCNM-based control, with the aim to find an optimized control strategy from various control laws with multiple scales in different layers.

Acknowledgements

Nan Deng appreciates the support of the China Scholarships Council (No.201808070123) during his Ph.D. thesis in the ENSTA Paris of Institut Polytechnique de Paris, and numerical supports from the laboratories LISN (CNRS-UPR 3251) and IMSIA (UMR EDF-ENSTA-CNRS-CEA 9219).

This work is supported by a public grant overseen by the French National Research Agency (ANR) under grant 'FlowCon' (ANR-17-ASTR-0022), the National Natural Science Foundation of China (NSFC) under grant 12172109, and the Polish Ministry of Science and Higher Education (MNiSW) under the Grant No.: 0612/SBAD/3567.

We appreciate valuable discussions with Guy Cornejo-Maceda, Hao Li, and Franҫois Lusseyran and the HIT fluidic pinball team: Bingxi Huang, Wenpeng Li, Yiqing (Anne) Li, Qixin (Kiki) Lin, Ruixuan (Rick) Shen, Xin Wang and Shangyan (Shane) Xie. Last but not least, we thank the referees for their thoughtful suggestions, which have significantly improved the manuscript.

Declaration of Interests. The authors report no conflict of interest.

5.A Blockage effect in the fluidic pinball

We run the DNS with an enlarged computational domain with the blockage ratio $B = 0.025$, which is bounded by a rectangular box of size $[-25D, +75D] \times [-50D, +50D]$, as shown in figure 5.29. The unstructured grid has 14 831 triangles and 29 961 vertices as compared to the grid in § 5.2.1.

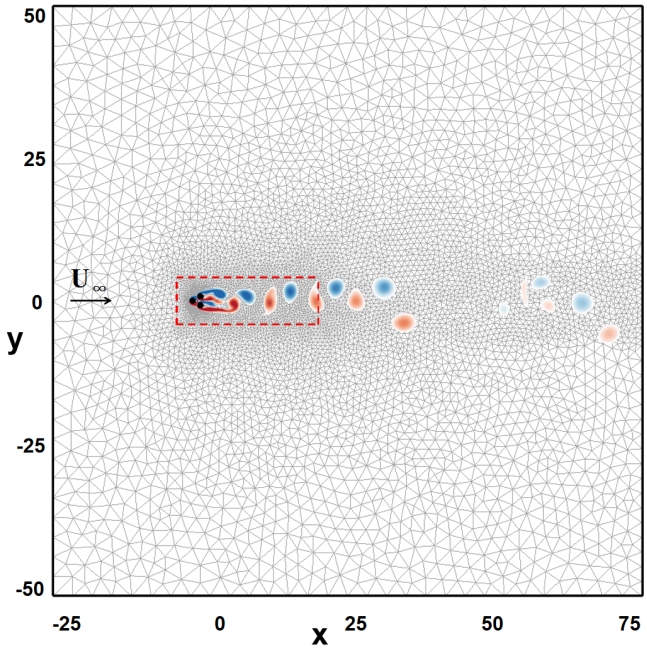


Figure 5.29: Configuration of the fluidic pinball and dimensions of the simulated domain of the blockage ratio $B = 0.025$. A typical field of vorticity at $Re = 150$ is represented in color with $[-1.5, 1.5]$. The upstream velocity is denoted U_∞. An observation zone of size $[-4D, +20D] \times [-4D, +4D]$ is marked out with a red dashed box.

The vortices in the near wake $0 < x < 20D$ are concentrated in $|y| < 4D$. Linear stability analysis of the symmetric steady solution indicates that the critical value of the primary Hopf bifurcation does not change $Re_1 = 18$, but the next bifurcations are found for larger Reynolds numbers. The pitchfork bifurcation of the symmetric steady solution is changed to $Re_2 = 81$, but the transient and post transient dynamics for different flow regimes remain qualitatively the same, as shown in figure 5.30.

Enlarging the computational domain reduces the blockage effect, but the blockage is practically difficult to suppress or even reduce. In this work, we are interested in the richness of the dynamics to evaluate our method. The blockage is not critical as we have similar numerical results for the transient and post-transient dynamics. How the location of lateral boundaries, as well as upstream and downstream boundaries, affect the bounded flow will be discussed in our future work.

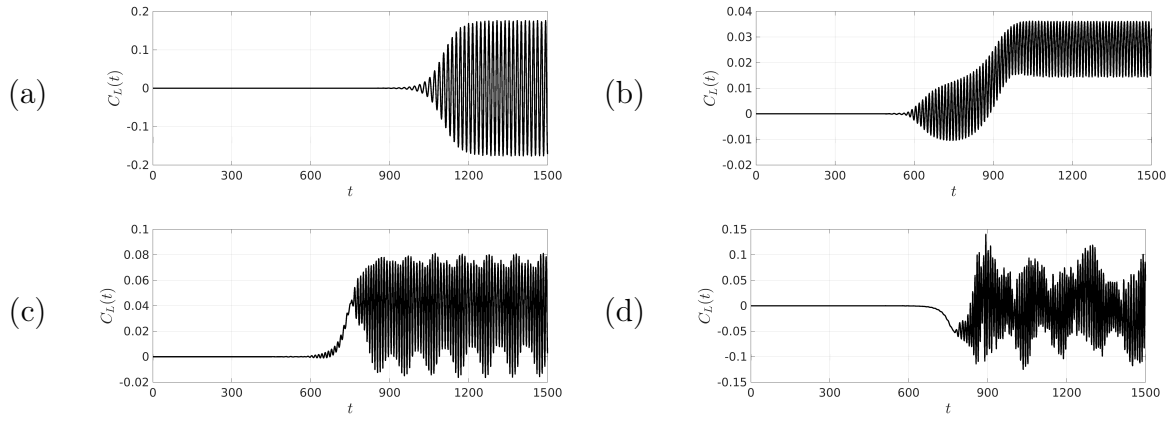


Figure 5.30: Transient and post-transient dynamics of the fluidic pinball with the enlarged grid, illustrated with the time evolution of the lift coefficients C_L starting with the symmetric steady solution at different Reynolds numbers: $Re = 30$ (a) , 90 (b) , 120 (c), 150 (d).

5.B Vertical transitions in the hierarchical network model

In figure 5.7, the snapshots entering into and leaving from $\mathcal{C}_{k_{l-1}}$ are marked out for each trajectory by the characteristic function (5.20), and can be used to describe the vertical transitions.

When a cluster in the parent layer $\mathcal{C}_{k_{l-1}}$ is activated, it also activates a horizontal transition through all its subclusters in the child layer. The ports of entry and exit for the subclusters are indicated by the entering and exiting snapshots. The entering snapshots belong to the first activated sub-cluster in each trajectory. The horizontal transition through subclusters ends with the exiting snapshots, and is forced to return to the parent layer. At the next time step after the exiting snapshots, $\mathcal{C}_{k_{l-1}}$ and all its subclusters will deactivate.

With an additional condition from the subclusters of the child layer, the characteristic function (5.20) can be defined as

$$\chi^m_{o,k_{l-1}\to k_l} := \begin{cases} 1, & \text{if } \boldsymbol{u}^{m-1} \notin \mathcal{C}_{k_{l-1}} \ \ \& \ \ \boldsymbol{u}^m \in \mathcal{C}_{k_{l-1},k_l}, \\ 0, & \text{otherwise.} \end{cases} \tag{5.25a}$$

$$\chi^m_{e,k_{l-1}\to k_l} := \begin{cases} 1, & \text{if } \boldsymbol{u}^{m+1} \notin \mathcal{C}_{k_{l-1}} \ \ \& \ \ \boldsymbol{u}^m \in \mathcal{C}_{k_{l-1},k_l}, \\ 0, & \text{otherwise.} \end{cases} \tag{5.25b}$$

The number of entering snapshots n_o and of exiting snapshots n_e in each subclusters $\mathcal{C}_{k_1,\cdots,k_l}$ read

$$n_{o,\,k_l} = \sum_{m=1}^{M} \chi^m_{o,k_{l-1}\to k_l}, \quad n_{e,\,k_l} = \sum_{m=1}^{M} \chi^m_{e,k_{l-1}\to k_l}, \tag{5.26}$$

where only the final subscript of the subcluster index in the current layer is indicated, and as well in the following.

For the cluster $\mathcal{C}_{k_{l-1}}$, the probability of vertical transition into and out of the child layer Q_{o,k_l} and Q_{e,k_l}, are defined as

$$Q_{o,\,k_l} = \frac{n_{o,\,k_l}}{n_o}, \ \ Q_{e,\,k_l} = \frac{n_{e,\,k_l}}{n_e}. \tag{5.27}$$

Note that $\sum_{k_l=1}^{k_l} Q_{o,\,k_l} = 1$ and $\sum_{k_l=1}^{k_l} Q_{e,\,k_l} = 1$.

An example of the hierarchical network model

An example of HiCNM is illustrated in figure 5.31. The model is constructed from the full data set containing one or several individual trajectories of discrete state snapshots, each starting from

different initial conditions and converging toward possibly different attracting sets. From the top to the bottom of this tree structure, only one cluster in each parent layer has been chosen and divided into subclusters in the child layer. For the clusters in each layer, we only indicate the sub-cluster number in the current layer.

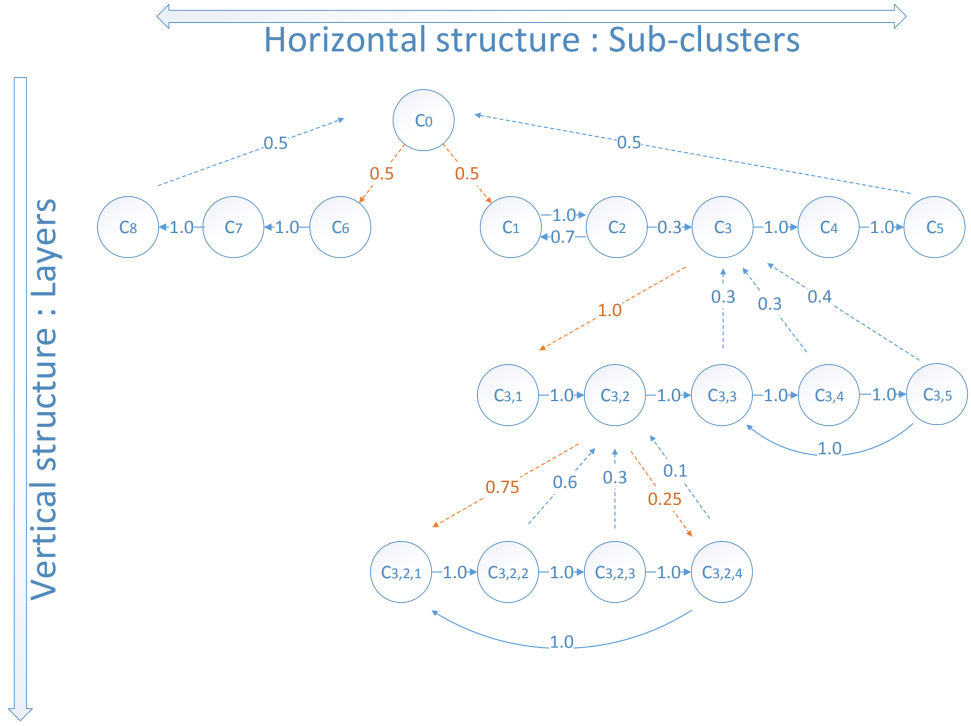


Figure 5.31: An illustration of an hierarchical network model. The vertical structure describes the transitions between the parent-child layers with dashed arrows: the red dashed arrows present the possible inlet from the parent layer to the child layer, and the blue dashed arrows denote the outlet back to the parent layer. The horizontal structure presents the transition between the subclusters with arrows. The numbers indicate the possibility of each transition.

The full data set is treated as one cluster $\mathcal{C}_0$ on the top, which ensembles all the snapshots. The clustering algorithm in the first layer has divided the snapshots in this cluster into eight subclusters. Two entering subclusters $\mathcal{C}_1$ and $\mathcal{C}_6$ are sketched with the same probability, as could be found when the full data set contains two pathways with equal probability in the state space of the system. The trajectories starting from the two entering subclusters have no intersecting cluster, and return to $\mathcal{C}_0$ from their respective existing subclusters, $\mathcal{C}_5$ and $\mathcal{C}_8$, with the same probability. The trajectory in layer $\mathcal{L}_1$, $\mathcal{C}_6 \to \mathcal{C}_7 \to \mathcal{C}_8$ is a simple one-way transition. The trajectory $\mathcal{C}_1 \to \cdots \to \mathcal{C}_5$ is more complex, due to a possible return dynamics from $\mathcal{C}_2$ to $\mathcal{C}_1$. The clustering algorithm works on the snapshots in cluster $\mathcal{C}_3$ and has divided them into five subclusters in the second layer $\mathcal{L}_2$. For the vertical transition, only one entering sub-cluster has been found but with three exiting subclusters, which means that the snapshots leaving cluster $\mathcal{C}_3$ belong to one of these three subclusters. The dynamics between the subclusters is a simple one-way transition, but with a cycle $\mathcal{C}_{3,3} \to \mathcal{C}_{3,4} \to \mathcal{C}_{3,5} \to \mathcal{C}_{3,3}$, which indicates a periodic dynamics. Next the clustering algorithm is applied to the snapshots in cluster $\mathcal{C}_{3,2}$, resulting in four subclusters in the third layer $\mathcal{L}_3$. Among the subclusters, there exist two entering subclusters and three exiting subclusters. We notice that $\mathcal{C}_{3,2,4}$ can work as either an entering sub-cluster or an exiting sub-cluster. The reason is that the entering snapshots and the exiting snapshots of $\mathcal{C}_{3,2}$ belong to the same sub-cluster in the child layer. A periodic dynamics exists between the subclusters $\mathcal{C}_{3,2,1} \to \cdots \to \mathcal{C}_{3,2,4} \to \mathcal{C}_{3,2,1}$.

A full dendrogram, as in figure 5.6, is also available if all the clusters in the parent layer are divided. However, in actual practice, it is not necessary to divide every cluster in the state space. The clustering algorithm in the first layer divides different invariant sets and the transient

states into different clusters. A classic CNM is used to describe the transient dynamics between the invariant sets. In the second layer, the refined dynamics on the invariant sets exhibits new interesting features. The snapshots are often concentrated close to the stable/unstable invariant sets. Within the same invariant set, the snapshots have a relatively homogeneous distribution according to certain rules. The local dynamics are relatively simple and easy to extract because the clustering result depends entirely on the distribution in the state space. The transient states from an unstable set to a stable set mix the dynamical behavior of different invariant sets. Hence, the clusters close to invariant sets need to be divided again, namely close to steady solution, to metastable solution, or to the stable solution. Usually, these clusters have some characteristics, like a large number of snapshots in the cluster or multiple possible transitions to other clusters.

5.C Clustering with POD

The computational cost can be significantly reduced with POD, as a lossless POD can highly compress the flow field data to accelerate the clustering algorithm. The clustering algorithm can be applied to the compressed data instead of the high-dimensional velocity fields. In this work, the snapshots of velocity field are pre-processed by a proper orthogonal decomposition (POD), where $\boldsymbol{u}_s(\boldsymbol{x})$ is the symmetric steady solution at the Reynolds number under consideration. Compared to the classical POD method, the symmetric steady solution $\boldsymbol{u}_s(\boldsymbol{x})$ has been used instead of the ensemble-averaged mean flow $\bar{\boldsymbol{u}}(\boldsymbol{x})$, because our analysis deals with multiple invariant sets and the mean flow is not a Navier-Stokes solution, which has no dynamical relevance. The fluctuating flow field can be decomposed on the basis of the POD modes $\boldsymbol{u}_i(\boldsymbol{x})$,

$$\boldsymbol{u}^m(\boldsymbol{x}) - \boldsymbol{u}_s(\boldsymbol{x}) \approx \sum_{i=1}^{N} a_i^m \, \boldsymbol{u}_i(\boldsymbol{x}), \tag{5.28}$$

where the a_i^m are the mode amplitudes. A complete basis for the modal decomposition is given when $N = M$ (Berkooz *et al.*, 1993). For our cluster-based analysis, the number of modes could be reduced to $N = 400$ without loss of relevant information. The pre-processing algorithm is detailed in algorithm 3. The computational cost for the cluster analysis can be significantly reduced in the

Algorithm 3 Pre-processing the velocity field by POD

Input: $\boldsymbol{u}^m$: snapshots of velocity field; $\boldsymbol{u}_s$: symmetric steady solution
Output: $\boldsymbol{u}_i$: leading POD modes; a_i^m: mode amplitudes

1: compute POD modes $\boldsymbol{u}_i$, $i = 1, \ldots, N$, for the data base $\{\boldsymbol{u}^m(\boldsymbol{x})\}$, $m = 1, \ldots, M$, with choosing $\boldsymbol{u}_s$ as the base-flow;
2: compute the mode amplitudes $a_i^m = (\boldsymbol{u}^m - \boldsymbol{u}_s, \boldsymbol{u}_i)_\Omega$;
3: save the leading N POD modes and the corresponding mode amplitudes.

POD subspace (Kaiser *et al.*, 2014; Li *et al.*, 2021), thus enabling an accurate compressed sensing of the original datasets. The centroids of the velocity field based on the POD mode amplitudes now read:

$$\boldsymbol{c}_k = \boldsymbol{u}_s + \frac{1}{n_k} \sum_{m=1}^{M} \chi_k^m \sum_{i=1}^{N} a_i^m \boldsymbol{u}_i. \tag{5.29}$$

We note that this POD process is just an option for data compression to speed up the clustering process, which can approximate the data distribution in a POD subspace with high accuracy. We can even apply the cluster-based approach to a feature-based subspace of the flow data, as Nair *et al.* (2019) who applied CROM to a 3D phase space of the drag and lift forces.

In contrast to the non-linear reconstruction of flows using a POD basis, the cluster-based approach neither decomposes the flow field nor extracts the dominant structures with the most

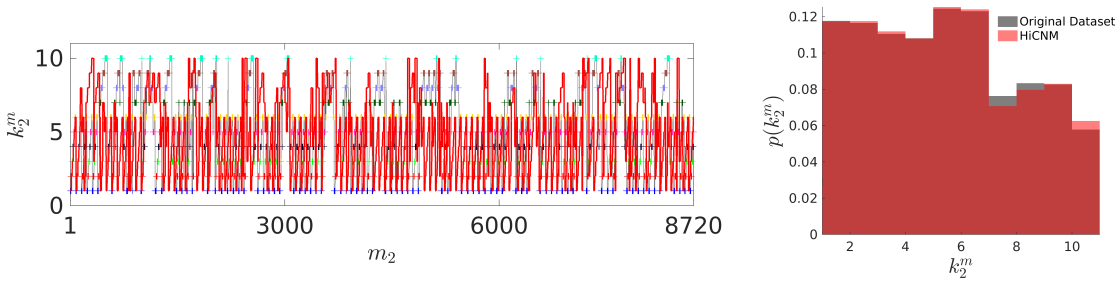
fluctuating energies. Instead, it gathers similar snapshots and represents them with a linear combination of snapshots within the cluster. This combination is always in the original data space and there is no projection comparing to a POD model. In addition, the modelling process of the cluster-based method is automated based on the data topology in the state space, which is even suitable for multiple invariant sets. The differences between the cluster-based methods and the POD-based model have been discussed in Kaiser *et al.* (2014); Li *et al.* (2021). As an extension, the HiCNM strategy provides a multiscale solution for describing complex flow with transient dynamics.

5.D Stochastic model for asymptotic regime

In § 5.4, we have applied HiCNMs for the identification and analysis of complex dynamics with multiple scales and multiple invariant sets. The centroids and cluster index provide a concise representation of the original data-set, and we can reconstruct the flow based on the time evolution of the cluster index. From the cluster analysis, we use a transition matrix to statistically record the possible motions between clusters, which can be used to predict the evolution of the cluster index. Hence, the dynamics reconstruction can also come from the stochastic model based on the transition matrix.

For a single trajectory, the transient dynamics is fully predictable and converges to the asymptotic dynamics. A stochastic model for the transient dynamics is not suggested because the clustering result may suffer from the multiscale problem, and some random walks will be mistakenly introduced into the transition matrix, like the destabilizing stage from the steady solution. In contrast, if there is no multiscale problem, all the transition probabilities should be 1 due to the predictable transient dynamics. Hence, there is no need to build the dynamics from a stochastic model and to discuss the probability distribution.

The asymptotic limit cycle at $Re = 80$ and torus at $Re = 105$ are also fully predictable as there is no random walk. Only for the chaotic dynamics in the asymptotic regime at $Re = 130$, can the transition matrix be used to construct a stochastic model. According to the current state in one of the clusters, it will choose the next destination according to the probability of transitions. Based on the local dynamics for the chaotic state in the group of clusters $\mathcal{C}_{k_1}$, $k_1 = 4, \ldots, 7$, the reconstructed dynamics and the probability distribution provided by the stochastic model are shown in figure 5.32. The main cycle $\mathcal{C}_1 \to \cdots \to \mathcal{C}_6 \to \mathcal{C}_1$ and the random walks to $\mathcal{C}_{k_2}$,



(a) (b)

Figure 5.32: Time evolution of the cluster index: the reconstructed dynamics (red curve) with the same initial cluster of the original dynamics (gray curve with colored markers) for $m_2 = 1, \ldots, 8720$ (a); and probability distribution of all m_2 in figure 5.27 by a stochastic model using the transition matrix (b), at $Re = 130$.

$k_2 = 7, \ldots, 10$ have been fairly reproduced, with the appropriate probability distribution.

Chapter 6

Conclusion

This final chapter proposes an overall conclusion and some perspectives for future work.

In this thesis, we presented a challenging benchmark example for automatable reduced-order modelling using first principles and machine learning. We chose the incompressible two-dimensional wake flow around the fluidic pinball because it has many typical characteristics of wake flows. Direct numerical simulations (DNS) at different values of the Reynolds number were performed, yielding different flow regimes: periodic (symmetric and asymmetric), quasi-periodic and chaotic flow regimes. For each considered value of the Reynolds number, the simulations were initiated in the vicinity of an unstable fixed point. The transient and post-transient dynamics of the unsteady flow showed various transient behaviours, involving the steady solutions, periodic solutions, quasi-periodic and chaotic attractors. Reduced-order modelling for the complex dynamics that results from the interaction between the dominant structures associated with different kinds of instabilities is undoubtedly a challenging task.

In order to obtain a sparse human-interpretable model of the flow dynamics, we started with the theoretical and numerical analyses of the governing equations. Linear stability analysis of the steady solutions, Floquet analysis of the periodic solutions, and the nonlinear analysis of the asymptotic dynamics provided a comprehensive understanding of the underlying mechanisms at play in the flow. This approach resulted in three different mean-field models: (i) the least-order mean-field model of Chapter 2, based on the elementary degrees of freedom associated with the Hopf and pitchfork bifurcations; (ii) the Galerkin force model supported by the mean-field model of Chapter 4, bridging the Galerkin model in the state space to the unsteady force dynamics; (iii) the hierarchical cluster-based reduced-order model of Chapter 5, generalizing to more complex dynamics which involve multi-scale, multi-frequency, and multi-attractor properties, as in the case of quasi-periodic or chaotic regimes. The mean-field models proposed in this work led to a better understanding of the coupling between the fluctuating and mean-flow fields. It also improved the performance and applicability of data-driven models. In the following, we summarize the contributions of this work and propose some perspectives for future work from the viewpoint of both flow analysis and mean-field modelling.

6.1 Flow analysis with transient and post-transient dynamics

For multi-structure configurations, like the fluidic pinball with three cylinders, various flow mechanisms may exist. Due to the complex interactions of the structures, the flow behaviour is sensitive to the spatial arrangement and the Reynolds number, which is far more complex than a single-structure configuration. This study mainly focuses on the unforced fluidic pinball and study the flow characterized by a single control parameter, the Reynolds number. The wake flow still presents various flow patterns at different values of the Reynolds number, including: (i) three steady flows with different states of the base-bleeding jet; (ii) three periodic von Kármán vortex street with different states of the base-bleeding jet; (iii) two quasi-periodic von Kármán vortex street with the

base-bleeding jet deflected up or down oscillating at a low frequency; (iv) a chaotic von Kármán vortex street with the base-bleeding jet switching up and down randomly. These flow patterns represent exact Navier-Stokes solutions. They are invariant sets and attractors in the phase space, and result from the Bénard-Von Kármán instability, the symmetry-breaking instability, and the complex interactions between them. From the standpoint of bifurcation theory, the system undergoes a supercritical Hopf bifurcation, a supercritical pitchfork bifurcation, a Neimar-Säcker bifurcation and a transition to chaos, as the Reynolds number is increased. These are common bifurcations in fluid mechanics. The flow analysis and modelling presented in this thesis can therefore apply to other flow configurations exhibiting similar instabilities and dynamics. Hence, the fluidic pinball can be seen as a great benchmark configuration for the analysis of complex flow dynamics and modelling strategies.

Transient and post-transient dynamics discussed in this work involve unsteady dynamics initiated close to unstable steady solutions. The coupling, competition and interaction of the main features associated with the above instabilities lead to very different transient behaviour. The temporal evolution of flow dynamics highly depends on the initial condition. A trajectory starting from an unstable steady solution can first be attracted by an unstable solution before being captured by an attractor. Multi-scale dynamics (in time and space) are concerned, including the linear dynamics in the neighbourhood of the steady solution, the weakly nonlinear dynamics during the transition and the nonlinear dynamics in the asymptotic regime on the attractor.

In order to get a comprehensive overview of the underlying mechanisms in this configuration, we conducted the linear stability analysis of the steady solutions and the nonlinear analysis of the transient and post-transient dynamics. As discussed in Chapter 2, the linear stability analysis of the steady solutions at different Reynolds numbers revealed two Hopf bifurcations and a pitchfork bifurcation. The corresponding unstable eigenmodes are reflection-antisymmetric. The theoretical analysis of the nonlinear saturation provides deep insights into the coupling between the unstable eigenmodes and the mean-field deformation, which provides a theoretical foundation for our mean-field modelling.

A closer view of the two Hopf bifurcations shows that their eigenvalues have different growth rates but similar frequencies. The related two pairs of complex-conjugated eigenvectors share similar spatial structures. The first Hopf bifurcation occurs supercritically at $Re = 18$. The second Hopf bifurcation occurs at $Re = 64$ with a smaller growth rate. A competing relation exists between these two Hopf bifurcations, with the size of their growth rates swapping at a higher Reynolds number. Interestingly, a stable quasi-periodic regime was observed at $Re = 104$. The bifurcation of periodic solution occurs almost coincidentally with the crossover of the growth rate of the two Hopf bifurcations of the base flow, as shown in figure 1.9. As already known, the linear stability analysis only involves the dynamics at the onset of the instability, with tiny perturbations in the neighbourhood of the steady solution. In contrast, the stable quasi-periodic regime is an attractor in the asymptotic regime. A better understanding of the quasi-periodic regime could arise from an in-depth study of the coupling between these two Hopf instabilities.

The Bénard-Von Kármán instability discussed above changes the wake globally, resulting in an oscillatory dynamics. The symmetry-breaking instability at the onset of the pitchfork bifurcation, occurring at $Re = 68$, have only a local impact on the flow, since it only affects the base-bleeding jet. However, the symmetry breaking of the jet changes the wake flow from a global viewpoint. In Chapter 3, we discussed the coupling and competition between the two instabilities. Generically, a local bifurcation only affects the stability of one equilibrium and cannot affect the stability of other equilibria. However, two coinciding pitchfork local bifurcations of the steady and periodic Navier-Stokes solutions were found in the fluidic pinball. The symmetry breaking of both the steady flow and the von Kármán vortex street occurs at nearly the same critical Reynolds number. The linear stability analysis of the steady flow and the Floquet stability analysis of the von Kármán street provide similar eigenmodes with the dominant structure concentrated in the near-wake. This indicates that the symmetry-breaking instability of the base-bleeding jet is spatially nearly

independent to the Bénard-Von Kármán instability in the downstream wake. The uncoupling of these two different instabilities leads to a coincident bifurcating mechanism, which is observed and discussed for the first time in fluid mechanics.

In this work, we mainly discussed the first two instabilities and the coupling relation between them. With an increasing Reynolds number, the effect of the von Kármán vortex street on the base-bleeding jet becomes stronger. The strong coupling mechanisms in the quasi-periodic and chaotic regimes remain, however, to be elucidated. The bifurcations reported in this work are supercritical. How will the model be modified in the sub-critical case? Are there any changes to the elementary degrees of freedom associated with a subcritical Hopf bifurcation or a subcritical pitchfork bifurcation? We suspect that additional shift modes will have to be introduced, but their structure remains to be seen.

Sensitivity analysis could be conducted in order to determine the optimal external excitation which would give rise to energy amplification in the wake flow. Such a study could provide new insights into the most sensitive regions associated with different instabilities. An alternative direction is to analyse the fluidic pinball under different control laws. The combinations of three rotating cylinders offer many possible control laws. The analytical works presented in this thesis could benefit the three elementary control laws of boat tailing, base-bleeding and Coanda forcing, which would provide a parameter-dependent dynamics analysis of this system. More generally, the gap ratio between the cylinders can also be a control parameter for various flow instabilities.

6.2 Mean-field modelling with first principles and machine learning

Numerous machine learning techniques enable dimensionality reduction and data-driven modelling. A mathematical model is precise to explain the first principles and works as the governing equations of the considered system. However, due to the high-dimensionality, non-linearity, and spatial and temporal multi-scale behaviour, we still need a human-interpretable model for a complete understanding of the transient dynamics and instabilities in the considered flow.

The most prominent finding in this thesis is that the mean-field consideration can provide a theoretical foundation for data-driven modelling, forcing the dynamical model onto the mean-field manifold. The mean-field models presented in this work can be mainly divided into two kinds of models: projection-based and cluster-based models. Both were applied to the complex dynamics of the fluidic pinball with multiple trajectories issued from the vicinity of three steady solutions, at the different Reynolds numbers.

For the projection-based modelling strategies, we first derived the least-order mean-field Galerkin model of Chapter 2, for the first two successive bifurcations, introducing the mean-field assumption into the standard Galerkin framework. The mean-field considerations enable us to use as few modes as possible to model the dynamics, and these modes are the optimal choice for the underlying instabilities. The original system is projected onto a low-dimensional subspace with an optimal basis, which guarantees the elementary components of the original dynamical system. The resulting mean-field Galerkin system inherits the quadratic nonlinearity of the governing Navier-Stokes equations, and present sparsity through a symmetry-based decomposition. The challenging issues for the standard Galerkin model, like instability growth, transient behaviour, and robustness of the identified model, can be cured by data-driven techniques. The SINDy (Sparse Identification of Nonlinear Dynamics) algorithm (Brunton *et al.*, 2016*a*) provides an established elegant framework for a quadratic Galerkin regression. We further proposed the Galerkin force model of Chapter 4, supported by the least-order mean-field model. Based on a general Galerkin decomposition of the unsteady incompressible flow, the aerodynamic force on the immersion body could be derived as a constant-linear-quadratic function of the mode amplitudes by projecting the force formulae onto a linearly independent low-dimensional subspace. In relation to the least-order mean-field model,

the force modelling process could be further simplified by considering the symmetry properties of the modes and using sparse calibration. The above two mean-field modelling strategies combine data-driven modelling under physical constraints. The resulting models are sparse and provide insights into the complex dynamics that can result from the interaction between the elementary degrees of freedom associated with the two bifurcations.

There are several directions for future works. The first direction is the automated physics-based modelling, like the mean-field Galerkin model and the Galerkin force model presented above. The physical or mathematical models can be projected onto an optimal subspace, then using machine learning techniques to identify the model. The dimensionality reduction is essential for this methodology, and the selected low-dimensional manifold directly decide the performance of the model. In our case, the modal selection is from dynamical consideration based on the analytical works. It is promising to automate this process with physical constraints, like mean-field, symmetry and frequency considerations.

The alternative choice is to automate manifold identification. The traditional way is mode-based with choosing the elementary modes for different instabilities, as exemplified in our works. The considered manifold basis are global modes with typical spatial structures. Another strategy is the feature-based manifold modelling (Loiseau *et al.*, 2018*a*), liberating from the spatial modes and using the data from sensors or the forces information as the manifold basis.

For quasi-periodic and chaotic dynamics, the flow may have undergone many different instabilities with increasing Reynolds number. It is extremely difficult to carry out the analysis of the constitutive elements of the dynamics and their interactions. The data-driven cluster-based modelling strategy better serves this goal. In Chapter 5, a self-supervised hierarchical cluster-based reduced-order modelling (HiCROM) methodology is proposed for bifurcations and topology identification of complex dynamics. The centroids after clustering are still in the original data space. The clustering process works as a kinematic compression of input data instead of projecting the data into an approximate subspace. The classical cluster-based network only considers the asymptotic dynamics on attractors. The hierarchical modelling strategy enables the network model for more complex transient and post-transient dynamics. The hierarchical clustering process is self-supervised with Reynolds decomposition under the mean-field consideration.

This methodology is a universal modelling strategy for complex dynamics of multiple scales and frequencies. The HiCNMs are successfully applied to the transient and post-transient dynamics of four trajectories starting with four different initial conditions at three Reynolds numbers: for the case involving six exact solutions at $Re = 80$, for the case involving three steady solutions and two quasi-periodic torus attractor at $Re = 105$, and for the case involving three steady solutions and a chaotic attractor at $Re = 130$, respectively.

HiCROM provides a flexible and general framework of cluster-based modelling for complex transient and post-transient dynamics for various fluid flows. The test for three cases with different coupling mechanisms shows its ability and applicability to extract transient and post-transient, multi-frequency, multi-attractor behaviours. A variety of potential applications is promissing, such as the topology identification of the state space with complex dynamics, the feature recognition of different instabilities, and the structure extraction with different spatial/time scales. We can also build a hierarchical projection-based model with local POD modes (Amsallem *et al.*, 2012) under the mean-field consideration. The cluster-based control strategy (Nair *et al.*, 2019) is also a promising direction to achieve an optimization from various control laws with multiple scales in hierarchical structure. Last but not least, the low-order statistical modelling of intermittence dynamics based on clusters also shows a good promise with considering stochasticity (Sapsis, 2021).

We have focused on the projection-based mean-field modelling and the cluter-based mean-field modelling in this thesis. The mean-field model sacrifices certain liberty to enhance a universal and inherent property from Reynolds stress closure. The data-driven mean-field modelling would be a fruitful area for further work.

Appendix A

Introduction et résumé en Français

Contents

Dans cette thèse, nous ouvrons la voie à la réduction de modèle automatisée pour la dynamique des fluides utilisant les premiers principes et des techniques d'apprentissage automatique. L'objectif est d'établir un problème de référence pour les caractéristiques dynamiques les plus importantes des écoulements de sillage. La configuration de référence choisie est l'écoulement de sillage incompressible bidimensionnel autour du pinball fluidique, un ensemble de trois cylindres parallèles dont les axes sont situés aux sommets d'un triangle équilatéral pointant vers l'amont.

Les structures à cylindres multiples se retrouvent généralement dans des applications d'ingénierie. L'organisation spatiale et la taille des cylindres ont une influence significative sur l'écoulement de sillage. Pour notre pinball fluidique, malgré la simplicité géométrique de la configuration, l'écoulement peut présenter une riche dynamique d'écoulement due aux interactions de sillage derrière les cylindres. Par exemple, à un faible nombre de Reynolds, les trois cylindres fonctionnent comme un unique corps non profilé et nous pouvons trouver une allée tourbillonnaire de

von Kármán en aval. À un nombre de Reynolds plus élevé, un écoulement de jet apparaît au centre des deux cylindres arrières. Avec l'augmentation du nombre de Reynolds, la longueur du jet augmente, et son instabilité est également affectée. Une rupture de symétrie se produit avec une déviation du jet vers le haut ou vers le bas. L'allée tourbillonnaire de von Kármán et le jet sont tous deux des exemples typiques de la dynamique des fluides et ont été bien étudiés pour leur transition des conditions laminaires aux conditions turbulentes. Dans cette étude, le lacher tourbillonnaire périodique et la rupture de symétrie résultent d'une bifurcation de Hopf et d'une bifurcation fourche. Les bifurcations fourches prennent naissance localement dans le sillage proche, tandis que l'allée tourbillonnaire de von Kármán prend naissance dans le sillage des trois cylindres pris comme un seul obstacle. Les relations de couplage et l'interaction entre ces deux instabilités dépendent fortement de leur distribution spatiale.

L'objectif de ce projet de recherche est de comprendre la dynamique transitoire et post-transitoire et l'interaction de ces deux instabilités fondamentales. L'analyse des dynamiques linéaires et non linéaires est basée sur l'analyse de stabilité linéaire des équations de Navier-Sokes et la saturation non linéaire dans les régimes asymptotiques. Sur la base de la compréhension physique de la transition, nous introduirons trois types de stratégies d'établissement d'un modèle d'ordre réduit, qui améliorent les performances de la méthode Galerkin, de la modélisation orientée vers les forces et de la modélisation de réseau basée sur les clusterspour les écoulements complexes. Les travaux analytiques peuvent garantir aux modèles résultants la capacité à capturer la dynamique transitoire et l'interprétation humaine, mais ces modèles nécessitent un effort considérable pour comprendre les mécanismes sous-jacents. Cependant, seule une partie des connaissances a priori est nécessaire pour qu'un modèle de type boîte grise puisse équilibrer interprétation et coûts des travaux analytiques. La connaissance des équations de Navier-Stokes, de la géométrie et des considérations sur le champ moyen amélioreront largement les performances du modèle. En raison de la simplicité géométrique de la configuration et les bifurcations génériques discutées dans ce travail, le processus d'analyse et de modélisation peuvent être appliqué à tout autre écoulement ayant une géométrie similaire.

Le chapitre introductif est structuré comme suit. § A.1 discute de l'objectif de la modélisation d'ordre réduit et décrit les différentes catégories de modélisation. § A.2 présente la problématique de cette thèse. Les dynamiques transitoires et post-transitoires avec attracteurs multiples sont brièvement étudiées. Dans la partie § A.3, les travaux analytiques sont présentés, y compris l'analyse de stabilité linéaire des solutions de Navier-Stokes, les instabilités et les bifurcations. Dans la partie § A.4, nous discutons de la saturation non linéaire pour la dynamique post-transitoire, et de la déformation du champ moyen due à l'effet du tenseur de Reynolds. § A.5 présente les trois stratégies de modélisation.

A.1 Modélisation d'ordre réduit

La complexité des écoulements fluides provient de sa haute dimensionnalité, des non-linéarités et de son comportement spatial et temporel multi-échelle. Lorsque l'on traite des problèmes d'écoulement de fluides, une quantité massive d'informations détaillées sur l'écoulement est générée à mesure que la résolution augmente. De toute évidence, il est extrêmement difficile de comprendre tous ces détails à différentes échelles, et l'optimisation de contrôle est impossible à appliquer. Malgré la dimensionnalité presque infinie d'un écoulement naturel, sa dynamique peut généralement être capturée par projection dans un espace de faible dimension. L'examen et l'analyse de la dynamique dans cet espace à faible dimension permettent de construire un modèle approximatif pour le système d'écoulement complexe. L'objectif de la modélisation d'ordre réduit est d'obtenir des modèles simplifiés et interprétables de l'écoulement des fluides, qui peuvent permettre de mieux comprendre les mécanismes sous-jacents. De tels modèles à faible dimension sont essentiels pour la prédiction de la dynamique de l'écoulement et la conception de contrôles efficaces.

Le processus de d'établissement d'un modèle d'ordre réduit d'un système d'écoulement peut

être divisé en deux étapes. La première étape consiste généralement par une décomposition modale pour obtenir une réduction de la dimensionnalité. Plusieurs modèles, présentant un comportement typique dans l'espace ou le temps, sont souvent suffisants pour capturer les mécanismes essentiels de l'écoulement. De nombreuses techniques ont été développées pour extraire ces modèles primaires, comme la décomposition orthogonale propre (POD), la décomposition en modes dynamiques (DMD). Sur la base des modes extraits, nous avons toujours besoin d'un modèle non linéaire pour décrire la dynamique dans l'espace à faible dimension. Contrairement au développement rapide de la décomposition modale ou de la réduction de la dimensionnalité, les modèles non linéaires dépendent de notre compréhension des systèmes non linéaires, et davantage de théories et de techniques doivent être développées. Sur la base des méthodes de décomposition, les approches générales appliquées sont la projection de Galerkin et la théorie de Koopman.

A.1.1 Pourquoi un modèle d'ordre réduit?

En raison des non-linéarités, des dimensions infinies et du problème multi-échelle dans l'espace et le temps, un modèle d'ordre réduit (ROM) du système d'écoulement sera utile pour comprendre les mécanismes physiques, améliorer l'efficacité des calculs et concevoir des lois de contrôle. Dans de nombreux cas, l'écoulement évolue généralement sur un attracteur de faible dimension, qui peut être caractérisé par un nombre fini de structures élémentaires. Ces structures typiques, présentant certaines échelles spatiales et temporelles, offrent la possibilité d'utiliser des modèles d'ordre réduit pour révéler les mécanismes sous-jacents cachés par la haute dimensionnalité.

Les modèles d'ordre réduit (ROM) présentent de nombreux avantages :

1. **Simplicité:** Une bonne extraction des caractéristiques dominantes de l'écoulement peut promettre la simplicité de la ROM résultante. De nouvelles techniques de réduction de dimension se développent rapidement, et il existe encore de nombreuses possibilités de trouver le meilleur ensemble de modes.
2. **Compréhension:** La représentation de faible dimension de la dynamique non linéaire originale nous aide à filtrer le bruit et à révéler le mécanisme sous-jacent. Les dynamiques linéaires et non linéaires peuvent être interprétées en termes de plusieurs degrés de liberté. Les techniques d'apprentissage automatique offrent de nouvelles méthodes et possibilités pour comprendre les dynamiques complexes directement à partir des données.
3. **Analyse:** L'interaction entre les degrés de liberté élémentaires permet de comprendre en profondeur la dynamique non linéaire. Les travaux analytiques permettent de découvrir les mécanismes fondamentaux et nous aident à comprendre les transitions non linéaires.
4. **Control:** Un modèle d'ordre réduit peut comprendre la dynamique dominante qui permet la mise en place efficace d'un contrôle linéaire et du contrôle optimal.
5. **Prédiction:** Une bonne approximation de la dynamique non linéaire dans des sous-espaces de faible dimension assure une prédiction dynamique rapide des transitions imminentes et permet une réaction par anticipation.
6. **Accélération du calcul:** Un modèle de substitution peut explorer la dynamique dans l'espace à faible dimension. Le modèle de substitution est dynamiquement équivalent au système original et peut être relié aux données à haute dimension pour accélérer la simulation.

A.1.2 White box modelling vs black-box modelling

Dans cette sous-section, nous classons brièvement les différents types de modélisation. Nous suivons Wiener (1948) et distinguons les modèles en tant que modèles boîte blanche (*white-box models* : WBM), modèles boîte grise (*gray-box models* : GBM) et modèles boîte noire (*black-box models* : BBM), comme indiqué dans A.1

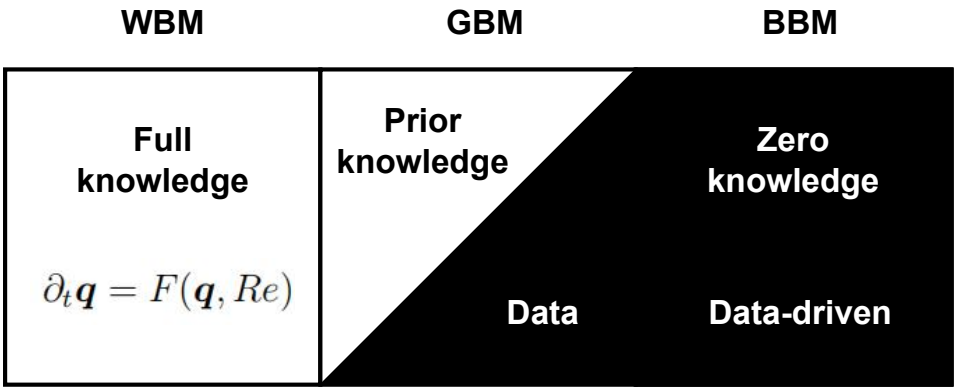


Figure A.1: Illustration du concept des trois types de stratégies de modélisation. Les modèles boîte blanche (WBM) sont des modèles purement théoriques, qui possèdent une connaissance complète du système d'écoulement. Les modèles boîte grise (GBM) combinent une base théorique partielle avec des données pour compléter le modèle. Les modèles boîte noire (BBM), quant à eux, sont des modèles statistiques sans aucune connaissance préalable du système.

Modélisation boîte blanche

Les équations de Navier-Stokes fournissent un modèle mathématique en mécanique des fluides, qui peut être considéré comme un modèle boîte blanche. Ce processus de modélisation est également connu sous le nom de modélisation en boîte claire, car il permet de connaître parfaitement les équations qui régissent le système d'écoulement. Des informations complètes sur l'état peuvent être dérivées de ce modèle, mais elles nécessitent des calculs lourds. Comme nous l'avons mentionné au début de ce chapitre, de nombreux problèmes ne sont toujours pas résolus avec ce modèle de type boîte blanche. Une forme faible de ce modèle a été fournie pour résoudre ces problèmes, comme RANS, LES, et d'autres modèles de turbulence. Il s'agit toujours de modèles à boîte blanche, car toutes les hypothèses proviennent de la dérivation théorique. Ces modèles sont coûteux à simuler numériquement et donc difficiles à utiliser pour la conception de contrôle, l'optimisation et les applications d'ingénierie.

En résumé, le but de la modélisation en boîte blanche est de trouver le "bon" modèle mathématiquement et physiquement. Par conséquent, le processus de modélisation est analytique, ce qui met l'accent sur la cause et l'effet, la forme et les principes. Il requiert des connaissances humaines sur les équations de Navier-Stokes, l'instabilité, les non-linéarités, les transferts d'énergie, etc. Les exigences élevées en matière d'expérience humaine et le coût élevé de la simulation rendent difficile son application dans les domaines de l'industrie et de l'ingénierie.

Modélisation boîte noire

Un modèle boîte noire est issu d'un processus de modélisation purement axé sur les données, mettant l'accent sur l'identification, la capacité de prédiction et la vitesse de modélisation. L'objectif est de trouver un modèle "précis" pour les caractéristiques, la dynamique et les comportements d'entrée-sortie du système. Une bonne précision rend le modèle plus utile dans les applications industrielles et le diagnostic médical.

La précision est généralement obtenue au prix d'un sacrifice de l'interprétabilité, comme c'est le cas pour la modélisation des réseaux neuronaux. Les relations fonctionnelles internes dans les couches multiples possèdent une liberté et une complexité accrues, ce qui les rend difficiles à expliquer d'un point de vue physique ou mathématique. Cependant, ce réseau neuronal complexe peut construire un modèle précis des comportements d'entrée-sortie, ce qui le rend plus utile dans les applications industrielles et le diagnostic médical.

De manière analogue, selon la théorie de Koopman (Schmid, 2010), la dynamique non linéaire

dans un espace de dimension finie peut être représentée par une dynamique linéaire dans un espace de dimension infinie. La décomposition en modes dynamiques (DMD) fournit une représentation linéaire des données d'un système non linéaire par un opérateur de Koopman de dimension infinie. Les modèles résultants ont une bonne capacité à capturer les caractéristiques cohérentes, mais ne peuvent pas être utilisés pour expliquer la dynamique transitoire non linéaire.

Différente des approches susmentionnées, la modélisation d'ordre réduit basée sur les clusters (CROM) offre une nouvelle façon de décrire la dynamique transitoire. Les données sont divisées en clusters en fonction de leur similarité. La dynamique transitoire est décrite par les transitions entre les clusters. Kaiser *et al.* (2014) utilise un modèle de Markov basé sur les clusters (CMM) pour décrire la dynamique transitoire avec des probabilités. Nair *et al.* (2019) a appliqué le CMM au contrôle d'écoulement par rétroaction non linéaire avec une extension de la chaîne de Markov qui met l'accent sur les transitions non triviales entre les clusters. Le graphe orienté (Newman, 2018) a été introduit avec les clusters, où les clusters sont considérés comme des nœuds et les transitions entre clusters comme des arêtes . Fernex *et al.* (2021) et Li *et al.* (2021) ont ensuite proposé le modèle de réseau basé sur les clusters (CNM) pour les données résolues dans le temps en introduisant l'interpolation locale entre les clusters avec les temps de transition déterminés.

Modélisation boîte grise

Par rapport au WBM basé sur des travaux analytiques et au BBM basé sur des données, GBM peut être considérée comme une approche hybride pour construire un modèle d'ordre réduit basé sur une structure théorique et des techniques basées sur les données. Par exemple, les modèles POD-Galerkin approchent l'état complet par projection Galerkin des équations de Navier-Stokes sur la base orthogonale d'un espace d'état de faible dimension. Le système Galerkin d'équations différentielles ordinaires quadratiques qui en résulte décrit la non-linéarité des équations de Navier-Stokes. Les modes de base proviennent d'un processus POD piloté par les données, ou même des modes propres des équations de Navier-Stokes linéarisées.

Dans cette thèse, nous nous concentrons sur la modélisation de la boîte grise en utilisant les données de l'écoulement, les premiers principes et les techniques d'apprentissage automatique. Les trois méthodes de modélisation suivantes ont été appliquées avec succès au pinball fluidique, qui sera examiné plus en détail dans la section 1.5.

A.2 Difficultés liées aux attracteurs multiples

Dans cette thèse, nous considérons la dynamique transitoire et post-transitoire d'un écoulement traversant un groupe de trois cylindres circulaires de même rayon. L'analyse de la dynamique de l'écoulement peut être grossièrement divisée en deux parties. La première se concentre sur la dynamique linéaire déstabilisant les solutions stationnaires ou les solutions périodiques, en appliquant l'analyse de stabilité linéaire de ces solutions exactes de Navier-Stokes. L'autre se concentre sur la dynamique post-transitoire évoluant sur un attracteur, en étudiant la saturation non linéaire dans le régime asymptotique, où la dynamique transitoire s'est calmée. En raison des bifurcations successives qui se produisent dans le système d'écoulement, le pinball fluidique non forcé présente un comportement multi-attracteur. Il existe trois solutions stationnaires et trois solutions périodiques, ainsi que de multiples trajectoires issues du voisinage de trois solutions stationnaires.

En général, la modélisation d'ordre réduit considère la dynamique post-transitoire évoluant sur un seul attracteur. La modélisation de la transition entre le point fixe et le régime asymptotique nécessite plus d'efforts pour saisir la distorsion du champ moyen, la dynamique locale et la saturation non linéaire pendant la partie transitoire. La construction d'un modèle universel d'ordre réduit pour des trajectoires transitoires multiples avec des attracteurs multiples est un défi majeur. Il est encore plus difficile de construire un modèle universel d'ordre réduit pour des trajectoires transitoires multiples avec plusieurs ensembles invariants. La dynamique d'écoulement considérée

dans cette thèse implique plusieurs régimes transitoires et post-transitoires, trois points fixes, trois cycles limites, deux attracteurs quasi-périodiques et un attracteur chaotique, à trois nombres de Reynolds différents.

A.2.1 Une configuration de référence - le pinball fluidique

L'écoulement d'un fluide autour d'un cylindre circulaire est l'un des plus célèbres problèmes fondamentaux de la mécanique des fluides et a été bien étudié au fil des années. Par rapport au cylindre unique, les structures à cylindres multiples sont plus fréquemment utilisées dans l'industrie, par exemple dans les échangeurs de chaleur et les conduites sous-marines. Les cylindres multiples offrent plus de possibilités pour différents types d'instabilités, le sillage interagissant les uns avec les autres, et une dynamique d'écoulement plus complexe.

Dans ce travail, nous nous intéressons à l'écoulement sur un groupe de trois cylindres parallèles, qui a été étudié expérimentalement incluant notamment le transfert de chaleur, les interactions fluide-structure et les interactions de fréquences multiples au cours des dernières décennies (Price & Paidoussis, 1984; Sayers, 1987; Lam & Cheung, 1988; Tatsuno *et al.*, 1998; Bansal & Yarusevych, 2017). Cinq modèles d'écoulement différents ont été identifiés en faisant varier les rapports d'espacement et les nombres de Reynolds par la simulation numérique (Bao *et al.*, 2010; Zheng *et al.*, 2016; Gao *et al.*, 2019; Chen *et al.*, 2020).

Pour le pinball fluidique, on utilise trois cylindres fixes de diamètre D dont les axes sont situés aux sommets d'un triangle équilatéral de côté $3D/2$ dans le plan (x, y) et qui sont orientés perpendiculairement à ce plan. Les cylindres sont espacés de R, le rayon des cylindres et le triangle formé par les trois cylindres pointe en amont. Ils sont placés dans un écoulement uniforme incompressible visqueux à la vitesse U_∞, comme le montre la figure 1.3. Les cylindres peuvent tourner à différentes vitesses, ce qui permet de modifier les trajectoires du fluide entrant, tout comme des pales manipulent la balle d'un flipper conventionnel. Cette configuration est un excellent bac à sable pour tester différentes stratégies de contrôle d'écoulement et a été utilisée pour l'évaluation de contrôleurs à plusieurs entrées et plusieurs sorties (Ishar *et al.*, 2019; Cornejo Maceda *et al.*, 2021). Dans cette étude, les trois cylindres restent statiques car nous nous intéressons à la dynamique naturelle de l'écoulement lorsque le nombre de Reynolds augmente. Le pinball fluidique non forcé a déjà montré une dynamique étonnamment riche.

L'écoulement est étudié dans le système de coordonnées cartésiennes délimité dans un domaine rectangulaire $[-6, 20] \times [-6, 6]$. L'origine du système de coordonnées cartésiennes est placée au milieu des deux cylindres arrière. Le domaine de calcul est discrétisé sur une grille non structurée. La configuration du pinball utilise une grille comportant 4 225 triangles et 8 633 sommets. Les conditions aux limites comprennent une condition de non-glissement sur les cylindres et une vitesse unitaire dans le champ lointain (1.4). Les conditions aux limites du champ lointain sont exercées sur les frontières d'entrée, supérieure et inférieure, tandis que la frontière de sortie est supposée être une frontière sans contrainte, transparente pour les structures du fluide sortant.

Le champ d'écoulement instantané est calculé par des simulations numériques directes bidimensionnelles (DNS). Pour tester la dépendance de la solution vis-à-vis de la grille, nous avons comparé le résultat des simulations avec une grille raffinée dans le Chapitre Chapter:JFM1. Les deux simulations prouvent l'indépendance de la grille et donnent des résultats dynamiquement cohérents.

A.2.2 Dynamique transitoire et post-transitoire

Dans cette sous-section, les dynamiques transitoires et post-transitoires avec attracteurs multiples seront décrites avec l'évolution temporelle des forces de traînée et de portance sur les cylindres. Le champ de vorticité des états d'écoulement explorés pendant la dynamique transitoire sera montré. Enfin, nous présenterons l'espace de phase 3D avec la dynamique des forces.

Dans la figure 1.5 sont représentées les dynamiques transitoires des coefficients de portance C_L et de traînée C_D depuis la solution symétrique stable instable $\mathbf{u}_s$ jusqu'au régime asymptotique, pour différentes valeurs du nombre de Reynolds Re. Théoriquement, le système ne devrait jamais quitter l'état initial car il s'agit d'un point fixe des équations. En pratique, les erreurs de discrétisation numérique rendent la solution imparfaite. Malgré un transitoire extrêmement long (plusieurs centaines de temps de convection), la trajectoire du système finit par s'échapper de $\mathbf{u}_s$ et atteint asymptotiquement l'attracteur. La dynamique transitoire représentée sur la figure 1.5 révèle les caractéristiques suivantes :

- Pour $Re = 30$, le système reste au voisinage de la solution stationnaire instable $\mathbf{u}_s$ pendant un temps extrêmement long. Une transition évidente est observée sur l'intervalle de temps $t \in [800, 1000]$. C_D augmente rapidement, et C_L commence à osciller autour d'une moyenne évanouissante avec une amplitude croissante. Le système d'écoulement entre finalement dans un régime asymptotique. Au voisinage de l'attracteur, la fréquence d'oscillation de C_L est la même que la fréquence de délestage des tourbillons f_{VS}, tandis que C_D oscille à une fréquence deux fois plus élevée.
- Pour $Re = 80$, pendant la dynamique transitoire, on peut voir que C_L commence à osciller autour d'une valeur moyenne évanouissante avant d'atteindre sa valeur moyenne asymptotique non nulle. Dès lors, la symétrie de l'écoulement moyen n'est brisée que dans un deuxième temps, puisque des centaines d'unités de temps convectives se sont écoulées après que le sillage ait commencé à osciller. Cela indique que la déstabilisation due à la bifurcation de Hopf se produit avant la bifurcation en fourche. L'état transitoire, observé sur la plage de temps $t \in [700, 800]$, a un coefficient de traînée plus faible que celui du régime final. La traînée la plus faible est associée à l'état initial $\mathbf{u}_s$.
- Pour des valeurs du nombre de Reynolds de 100 à 120, qui couvrent les régimes périodique ($Re = 100$), quasi-périodique ($Re = 105$) et chaotique ($Re \geq 115$), C_L brise d'abord la symétrie avant de commencer à osciller. Cela signifie que la déstabilisation due à la bifurcation de Hopf se produit après la bifurcation en fourche. Il est intéressant de noter que la traînée de l'état d'équilibre asymétrique observée autour de $t \approx 700$ pour $Re = 100$ à 120, est minimale, tandis que la traînée associée à l'état final est maximale.
- A partir de $Re = 80$ & 100, l'amplitude d'oscillation de la force de traînée est très faible, et la fréquence est identique à la fréquence de lâcher du vortex f_{VS}. En revanche, la force de portance oscille à la fréquence de lâcher des tourbillons avec une amplitude beaucoup plus grande. Pour $Re = 105$ & 110, la traînée oscille avec une plus grande amplitude à la fréquence de la modulation du jet, qui est inférieure d'un ordre de grandeur à la fréquence de lâcher des tourbillons, $f_{\mathrm{JET}} \approx f_{\mathrm{VS}}/12$.
- Pour $Re = 120$ & 130, avant d'entrer dans le régime entièrement chaotique, la partie initiale de la dynamique transitoire montre une dynamique transitoire similaire à celle du cas à faible nombre de Reynolds. A $Re = 120$, le coefficient de portance C_L atteint d'abord la valeur d'une solution asymétrique stable, et il commence à osciller autour d'une valeur moyenne évanouissante à $Re = 130$. Cela indique qu'une seule instabilité de ces deux bifurcations se produit pendant la déstabilisation initiale.

Les états d'écoulement explorés pendant la dynamique transitoire mentionnée ci-dessus sont illustrés par le champ de vorticité dans la figure 1.4.

Les dynamiques transitoires et post-transitoires dépendent des conditions initiales. En suivant le DNS appliqué dans la figure 1.5, nous choisissons les solutions stationnaires symétriques/asymétriques comme condition initiale, et le système présente une dynamique riche. Ces trajectoires sont tracées dans un espace d'intégration du coefficient de traînée $C_D(t)$, du coefficient de portance $C_L(t)$ et du coefficient de portance retardé $C_L(t-\tau)$, comme le montre la figure 1.6. La dynamique transitoire et post-transitoire est différente : au stade initial, elle présente une dynamique linéaire

au voisinage de la solution stationnaire correspondante et une dynamique non linéaire dans le régime asymptotique d'un attracteur. Dans tous les cas, l'écoulement s'éloigne des solutions stationnaires instables et se rapproche finalement du régime asymptotique. À $Re = 30$, il n'existe que l'instabilité de von Kármán et une seule trajectoire d'un point fixe au cycle limite. Après la rupture symétrique, il existe de multiples trajectoires entre les trois solutions stationnaires et les deux attracteurs. L'écoulement héritera de la symétrie des solutions stationnaires avant d'être piégé par un régime chaotique, comme $Re \geq 120$. En particulier pour les trajectoires de la solution stationnaire symétrique, elles s'approcheront d'abord d'un état transitoire, par exemple, un cycle limite instable, comme $Re = 80$, ou les solutions stationnaires asymétriques instables, comme $Re = 100, 105, 120$, avant d'entrer dans l'état permanent.

En résumé, le pinball fluidique sans forçage fournit une dynamique d'écoulement riche. Nous nous intéressons à la construction d'un modèle de champ moyen à différents nombres de Reynolds, en évaluant la dynamique transitoire et post-transitoire pour trois cas : avec six ensembles invariants, avec l'état quasi-périodique, et avec l'état chaotique.

A.3 Analyse de la stabilité linéaire

Dans cette section, nous nous intéressons à la dynamique de l'écoulement de petites perturbations au voisinage de la solution stationnaire. Sous cette considération, l'évolution temporelle d'une petite perturbation peut être considérée comme une dynamique linéaire dans un temps transitoire fini.

A.3.1 Stabilité linéaire d'un écoulement de base.

Dans un écoulement de fluide, une analyse de stabilité linéaire se concentre sur la dynamique de perturbation linéaire à partir d'un écoulement de base. Cet écoulement de base est une solution exacte des équations gouvernantes, et peut être une solution stationnaire ou une solution périodique. L'analyse de stabilité peut être globale ou locale, selon que l'on étudie ou non la perturbation sur l'ensemble du domaine. Dans ce travail, l'analyse de stabilité linéaire globale de l'écoulement de base révèle l'instabilité sous-jacente du système d'écoulement.

Analyse de stabilité linéaire de la solution stationnaire

Autour de l'équilibre $\boldsymbol{U}_s(\boldsymbol{x})$, on considère une perturbation infinitésimale $\boldsymbol{u}'(\boldsymbol{x}, t)$:

$$\boldsymbol{u}(\boldsymbol{x}, t) = \boldsymbol{U}_s(\boldsymbol{x}) + \boldsymbol{u}'(\boldsymbol{x}, t). \tag{A.1}$$

En définissant l'opérateur de Navier-Stokes linéarisé autour de la solution stationnaire $\boldsymbol{U}_s$ comme $\mathcal{L}_{\boldsymbol{U}_s}$ et le vecteur d'état de perturbation $\boldsymbol{q}' = (\boldsymbol{u}', p')$, le système linéarisé peut être écrit comme :

$$\partial_t \boldsymbol{q}' = \mathcal{L}_{\boldsymbol{U}_s} \boldsymbol{q}'. \tag{A.2}$$

L'analyse modale de la stabilité conduit à un problème de valeur propre $(\sigma + i\omega)\,\hat{\boldsymbol{q}} = \mathcal{L}_{\boldsymbol{U}_s}\hat{\boldsymbol{q}}$, avec une perturbation de la forme $\boldsymbol{q}'(\boldsymbol{x}, t) = \hat{\boldsymbol{q}}(\boldsymbol{x})e^{(\sigma+i\omega)t}$, ce qui donne

$$\boldsymbol{u}'(\boldsymbol{x}, t) = \hat{\boldsymbol{u}}(\boldsymbol{x})e^{(\sigma+i\omega)t}, \quad p'(\boldsymbol{x}, t) = \hat{p}(\boldsymbol{x})e^{(\sigma+i\omega)t}. \tag{A.3}$$

La partie réelle des valeurs propres σ, correspondant au taux de croissance de la perturbation, elle décide de la stabilité de l'écoulement de base considéré $\boldsymbol{U}_s$. Si toutes les valeurs propres σ sont négatives, l'écoulement de base est stable, ce qui indique que toutes les perturbations se réduiront à zéro, sinon l'écoulement de base est instable. Comme l'écoulement de base dépend du nombre de Reynolds, les valeurs propres changeront avec le changement du nombre de Reynolds. En général, un système d'écoulement stable tend à devenir instable lorsque le nombre de Reynolds augmente,

ce qui se traduit par une valeur propre passant par zéro. La partie imaginaire ω détermine la fréquence initiale, qui est non nulle pour une bifurcation de Hopf.

Analyse de Floquet pour un écoulement périodique

Pour un écoulement périodique temporel, le problème de stabilité utilisant la théorie de Floquet fonctionne avec un écoulement de base T-périodique $\boldsymbol{Q}_p(\boldsymbol{x},t) = (\boldsymbol{U}_p(\boldsymbol{x},t), P_p(\boldsymbol{x},t))$. Le système linéarisé est le même que Eq. (A.2) mais l'opérateur linéaire $\mathcal{L}_{\boldsymbol{U}_p(t)}$ est T-périodique. Les solutions du problème propre conduisent aux modes de Floquet T-périodiques $\hat{\boldsymbol{q}}(\boldsymbol{x},t)$ et aux exposants de Floquet correspondants $\sigma + i\omega$. Les multiplicateurs de Floquet peuvent être écrits comme $\lambda_F = e^{(\sigma+i\omega)T}$.

A.3.2 Instabilités et bifurcations

L'analyse de stabilité linéaire des solutions stationnaires symétriques $\boldsymbol{U}_s$ à différents nombres de Reynolds a été réalisée sur un sous-espace de Krylov de dimension 9 avec 100 itérations. Deux paires de valeurs propres conjuguées complexes, et une valeur propre de fréquence nulle peuvent être trouvées avec une partie réelle positive lorsque le nombre de Reynolds augmente. La première paire de valeurs propres conjuguées avec partie réelle positive apparaît pour la première fois lorsque le nombre de Reynolds passe de 18 à 19, voir figure 1.8(gauche). La deuxième paire de valeurs propres conjuguées croise l'axe $\sigma = 0$ lorsque le nombre de Reynolds passe de 64 à 65, voir la figure 1.8(milieu). La valeur propre réelle devient positive entre $Re = 68$ et 69, voir figure 1.8(droite). Ceci confirme qu'une bifurcation de Hopf se produit sur la solution stationnaire symétrique à $Re_1 \approx 18$, une autre bifurcation de Hopf se produit à $Re_2 \approx 64$ et une bifurcation en fourche à $Re_3 \approx 68$.

La partie réelle des valeurs propres indique le taux de croissance des modes propres correspondants pour la dynamique linéaire au voisinage de la solution stationnaire donnée $\boldsymbol{U}_s$. La partie imaginaire non nulle correspond quant à elle à la fréquence initiale. Le taux de croissance positif indique les modes propres instables, déstabilisant le champ d'écoulement de $\boldsymbol{U}_s$. Nous traçons le taux de croissance des modes propres correspondants dans la figure 1.9 en fonction du nombre de Reynolds. Nous remarquons que la première paire complexe-conjuguée a le plus grand taux de croissance à $Re \leq 90$, et la valeur propre réelle a le plus grand taux de croissance à $Re \geq 95$. Pour les deux paires complexes-conjuguées, la première paire a un taux de croissance plus important à $Re \leq 105$. Nous rappelons que ces taux de croissance indiquent seulement la dynamique linéaire au voisinage de la solution stationnaire $\boldsymbol{U}_s$. Avec le développement de l'instabilité, la dynamique linéaire ne s'applique plus. Le taux de croissance sera réduit à zéro après la saturation, et la fréquence peut également être modifiée par l'interaction non linéaire. Nous discuterons de la dynamique non linéaire dans les régimes asymptotiques dans la section suivante.

A.4 Analyse faiblement non linéaire

La dynamique linéaire ne concerne qu'un voisinage restreint de la solution stationnaire, car la perturbation $\boldsymbol{u}'$ est minime au début de l'instabilité. Avec le développement de la perturbation, le terme non linéaire $\boldsymbol{u}' \otimes \boldsymbol{u}'$ ne peut être ignoré, et va commencer à modifier l'écoulement moyen $\overline{\boldsymbol{U}}$ par l'effet du tenseur de Reynolds $\overline{(\boldsymbol{u}' \cdot \nabla)\boldsymbol{u}'}$. Au fur et à mesure que la distorsion du champ moyen par rapport à la solution stationnaire augmente, la dynamique linéaire originale ne convient plus, et l'itération non linéaire va charger la dynamique transitoire jusqu'à saturation.

La fIgure 1.17 montre la dynamique transitoire et post-transitoire à partir de la solution stationnaire d'un écoulement d'un unique cylindre. Cette instabilité oscillatoire correspond à une simple bifurcation de Hopf supercritique, où une paire de modes oscillatoires conjugués suffit à décrire les composantes fluctuantes. La dynamique non linéaire de ce système peut être projetée

dans un plongement spatial tridimensionnel avec les évolutions temporelles de deux modes oscillatoires et d'un mode shift (i.e. le mode de déformation du champ moyen) décrivant la déformation du champ moyen. Pendant l'évolution du point fixe au cycle limite, les modes oscillatoires sont constamment déformés et se rapprochent du cylindre. La longueur de la bulle de recirculation diminue continuellement de la solution stationnaire à l'écoulement moyen saturé. Cette déformation sur l'écoulement moyen peut être vue comme le résultat des contraintes de Reynolds générées par le champ fluctuant (Barkley, 2006).

A.4.1 Brisure de symétrie de l'écoulement de base

Dans cette sous-section, nous examinons la saturation non linéaire de la bifurcation fourche de l'écoulement de base (la solution stationnaire de symétrie $\boldsymbol{U}_s(\boldsymbol{x})$, qui conduit à trois solutions stationnaires.

Pour $Re > Re_{\mathrm{PF}}$, il existe trois solutions stationnaires satisfaisant les équations de Navier-Stokes stables (1.5). En considérant une décomposition pour une bifurcation fourche:

$$\boldsymbol{U}(\boldsymbol{x},t) = \boldsymbol{U}_s(\boldsymbol{x}) + \boldsymbol{u}_4(\boldsymbol{x},t) + \boldsymbol{u}_5(\boldsymbol{x},t), \tag{A.4}$$

avec les perturbations $\boldsymbol{u}_4(\boldsymbol{x},t)$ et un mode shift $\boldsymbol{u}_5(\boldsymbol{x},t)$ pour la dynamique à variation lente l'écoulement moyen. Le mode $\boldsymbol{u}_4$ est antisymétrique, et les autres sont symétriques par réflexion.

Introduisons Eq. (A.4) dans les équations de Navier-Stokes et séparerons la partie antisymétrique et la partie symétrique. La partie symétrique décrit la distorsion du champ moyen $\boldsymbol{u}_5$. En introduisant $\langle \boldsymbol{U} \rangle \equiv \boldsymbol{U}_s + \boldsymbol{u}_5$, on a:

$$(\boldsymbol{u}_4 \cdot \nabla)\boldsymbol{u}_4 = -(\boldsymbol{U}_s \cdot \nabla)\boldsymbol{u}_5 - (\boldsymbol{u}_5 \cdot \nabla)\boldsymbol{U}_s + \nu \triangle u_5 - \nabla p_5. \tag{A.5}$$

Le terme non linéaire$(\boldsymbol{u}_4 \cdot \nabla)\boldsymbol{u}_4$ correspond au tenseur de Reynolds de la partie antisymétrique des perturbations.

À $\partial_t \boldsymbol{u}_4 = 0$, le système atteint l'une des solutions stationnaires asymétriques ($\boldsymbol{u}_4 \neq 0$). La saturation non linéaire peut être présentée avec :

$$(\boldsymbol{u}_5 \cdot \nabla)\boldsymbol{u}_4 + (\boldsymbol{u}_4 \cdot \nabla)\boldsymbol{u}_5 = -(\boldsymbol{U}_s \cdot \nabla)\boldsymbol{u}_4 - (\boldsymbol{u}_4 \cdot \nabla)\boldsymbol{U}_s + \nu \triangle \boldsymbol{u}_4 - \nabla p_4. \tag{A.6}$$

Le côté droit de Eq. (A.6) peut être écrit comme $\mathcal{L}_{\boldsymbol{U}_s}\boldsymbol{u}_4$, ce qui contribue à l'instabilité linéaire de la solution stationnaire symétrique. Dans ce cas, la croissance de $\boldsymbol{u}_4$ a été saturée par l'interaction non linéaire du côté gauche.

A.4.2 Analyse de stabilité de l'écoulement moyen

L'écoulement moyen moyenné dans le temps d'un écoulement périodique n'est pas une solution exacte des équations de Navier-Stokes, à proprement parler. Il s'agit uniquement d'une solution statistique qui ne peut pas révéler la stabilité du système physique. Pour l'écoulement périodique résultant d'une bifurcation de Hopf, comme dans la figure 1.17, l'écoulement moyen diffère de la solution stationnaire car déformé par le tenseur de Reynolds du fait des perturbations durant la phase transitoire et post-transitoire. La distorsion du champ moyen est un résultat direct du processus de saturation non linéaire. L'analyse de stabilité linéaire de la solution stationnaire ne peut prédire l'instabilité qu'au début de la déstabilisation. Comme la dynamique est déjà saturée, l'écoulement moyen est marginalement stable, le taux de croissance étant presque nul. La fréquence après la saturation non linéaire est également différente de l'analyse de stabilité linéaire de la solution stationnaire. Cependant, cette fréquence non linéaire peut être correctement capturée par l'analyse de stabilité de l'écoulement moyen. Cela correspond à une propriété dite RZIF : la partie réelle est nulle, et la partie imaginaire est la fréquence du cycle limite (Turton *et al.*, 2015). Cette propriété est un résultat direct de l'équilibre harmonique (Dušek *et al.*, 1994). A partir

de la même idée, Mantič-Lugo *et al.* (2014) a proposé le modèle auto-consistant pour prédire la saturation non linéaire de $\boldsymbol{U}_s$ à $\overline{\boldsymbol{U}}$ pour une bifurcation de Hopf supercritique de l'écoulement du cylindre seul. Au stade final, nous avons $\sigma = 0$, et nous pouvons obtenir les valeurs exactes de $\overline{\boldsymbol{U}}$, $\boldsymbol{u}_1$ et ω du cycle limite.

A.5 Modélisation en champ moyen

En physique et en statistique, la théorie du champ moyen, également connue sous le nom de théorie du champ autoconsistant, est une simplification de l'étude d'un système stochastique à haute dimension par le calcul de moyennes statistiques. L'étude de l'écoulement moyen peut révéler certaines caractéristiques du système original avec un coût de calcul plus faible et a été appliquée à un large éventail de disciplines scientifiques. En mécanique des fluides, l'écoulement moyen fait généralement référence à une solution moyenne dans le temps de l'écoulement périodique, les composantes fluctuantes disparaissant. Contrairement à la solution stationnaire, l'écoulement moyen peut être obtenu à la fois numériquement et expérimentalement. L'analyse de cette solution a un champ d'application plus large et peut contribuer à une meilleure compréhension de l'interaction non linéaire pendant la transition.

Comme mentionné dans la section précédente § A.4.2, l'analyse de la stabilité de l'écoulement moyen peut prédire avec succès la fréquence non linéaire de l'état saturé. Il s'avère que la propriété RZIF est un résultat direct de la saturation non linéaire de l'instabilité oscillatoire. Les considérations sur le champ moyen conduisent également à la décomposition de Reynolds du champ d'écoulement (A.7) et aux équations de Navier-Stokes moyennées selon Reynolds (1.27). Le modèle autoconsistant proposé dans Mantič-Lugo *et al.* (2014) utilise un seul mode oscillatoire pour prédire la dynamique de saturation non linéaire d'une bifurcation de Hopf supercritique.

A.5.1 Mean-field ansatz

Sous la considération du champ moyen, la triple décomposition du champ d'écoulement est similaire à Reynolds & Hussain (1972)

$$\boldsymbol{u}(\boldsymbol{x},t) = \underbrace{\langle \boldsymbol{u}(\boldsymbol{x},t)\rangle_T}_{\omega \ll \omega_c} + \underbrace{\tilde{\boldsymbol{u}}(\boldsymbol{x},t)}_{\omega \sim \omega_c} + \underbrace{\boldsymbol{u}'(\boldsymbol{x},t)}_{\omega \gg \omega_c}, \tag{A.7}$$

où la fréquence angulaire dominante ω_c est définie comme le pic dominant dans le spectre de Fourier du champ de vitesse. Ici, le champ de vitesse est décomposé en un champ d'écoulement moyen à variation lente $\langle \boldsymbol{u}\rangle_T$, une composante cohérente sur des échelles de temps d'ordre $2\pi/\omega_c$, impliquant des structures cohérentes $\tilde{\boldsymbol{u}}$, et les fluctuations non cohérentes restantes à petite échelle $\boldsymbol{u}'$. Ce type de décomposition se retrouve également dans les modèles Galerkin d'ordre inférieur de Tadmor *et al.* (2011) et dans la modélisation faiblement non linéaire de Rigas *et al.* (2017*a*). La déformation de l'écoulement moyen provient de l'effet du tenseur de Reynolds de la composante cohérente $\overline{(\tilde{\boldsymbol{u}}\cdot\nabla)\tilde{\boldsymbol{u}}}$ et des fluctuations non cohérentes à petite échelle $\overline{(\boldsymbol{u}'\cdot\nabla)\boldsymbol{u}'}$.

A.5.2 Structure de la thèse

Cette thèse vise à faciliter la modélisation automatisée d'ordre réduit pour un modèle parcimonieux interprétable en utilisant les premiers principes et les techniques d'apprentissage automatique. Plus précisément, nous nous concentrons sur la dynamique de sillage complexe que l'on trouve généralement en mécanique des fluides et nous développons différents types de modèles de champ moyen. Les considérations relatives au champ moyen constituent un élément clé pour la construction d'un modèle de champ moyen, en supposant que la déformation du champ moyen variant lentement sous l'effet du tenseur de Reynolds est générée par les fluctuations, par exemple la décomposition

de Reynolds. Les modèles de champ moyen proposés dans cette étude constituent un exemple de référence stimulant pour la modélisation automatisable d'ordre réduit.

Dans ce chapitre d'introduction, nous avons abordé la motivation de la modélisation d'ordre réduit et avons donné un aperçu des principes et des méthodologies pour les stratégies de modélisation en champ moyen basées sur les clusters et sur les projections :

(a) **Modélisation en champ moyen par clusters :** Le clustering et la modélisation sont effectués sur la base de la considération du champ moyen, en incorporant une structure hiérarchique issue d'une hypothèse théorique dans la modélisation pilotée par les données. Le clustering réalise une compression cinématique des données d'entrée, accompagnée d'une moyenne statistique des données de regroupement (centroïdes) dans l'espace de données original. Les données sont étiquetées avec l'indice de cluster, qui regroupe automatiquement les états dans l'espace d'état sans approximation. La dynamique identifiée peut préserver la structure du système original en fonction de la distribution des clusters. Le modèle hiérarchique peut identifier systématiquement la dynamique multi-échelle, y compris les transitions entre différentes solutions, la dynamique de bifurcation vers différents attracteurs, et les structures locales pour l'écoulement moyen déformé dans le temps.

(b) **Modélisation en champ moyen par projection :** Le système original est projeté sur un sous-espace de faible dimension linéairement indépendant, et la dynamique non linéaire sera examinée dans le sous-espace choisi. La base de la représentation à basse dimension est essentielle pour la dynamique approximative, ce qui se reflète sur la stabilité, le comportement transitoire et la robustesse du modèle identifié. L'hypothèse du champ moyen pour le système Galerkin conduit à un système Galerkin contraint, où les degrés de liberté sont optimaux pour les instabilités sous-jacentes du système. La connaissance de l'analyse de stabilité et des états saturés peut aider à déterminer certains coefficients du modèle de champ moyen et à augmenter la robustesse du modèle parcimonieux (en anglais *sparse*) de Galerkin en identifiant les coefficients restants par une régression parcimonieuse.

(c) **Modélisation des forces par projection :** La modélisation de la force est basée sur la projection de l'expression mathématique de la force sur un sous-espace linéaire indépendant de faible dimension. Elle peut être généralisée pour tout autre modèle de Galerkin sans hypothèse de champ moyen. Cependant, la symétrie modale et les relations d'asservissement dans un modèle à champ moyen peuvent introduire une sparsité supplémentaire dans le modèle de force et améliorer l'interprétabilité du modèle identifié à partir de la régression parcimonieuse de Galerkin.

Toutes les stratégies de modélisation mentionnées ci-dessus sont exemplifiées pour la dynamique transitoire et post-transitoire du pinball fluidique non forcé à différents nombres de Reynolds. La thèse est organisée comme suit :

Dans le chapitre 2, le code de simulation numérique du pinball fluidique est présentée. Les comportements d'écoulement à différents nombres de Reynolds sont étudiés numériquement. L'analyse de stabilité linéaire des solutions stationnaires et l'analyse de stabilité de Floquet des solutions périodiques sont effectuées pour la vérification des deux premières bifurcations. Un modèle Galerkin à champ moyen est obtenu pour la bifurcation primaire de Hopf et la bifurcation fourche secondaire. L'identification du système Galerkin est basée sur la régression quadratique parcimonieuse avec des contraintes basées sur la physique. Le modèle d'ordre minimal est seulement à cinq dimensions mais peut reproduire les caractéristiques clés du comportement transitoire et post-transitoire de la dynamique complète.

Dans le chapitre 3, nous discutons des bifurcations locales coïncidentes des solutions stationnaires et périodiques de Navier-Stokes trouvées dans le pinball fluidique. Cette coïncidence n'aurait pas dû être observée dans l'écoulement de sillage, et le mécanisme de ce type de coïncidence non-générique est modélisé et expliqué. L'instabilité de rupture de symétrie du jet est indépendante de l'instabilité de von Kármán ou du lâcher de tourbillonds.

Dans le chapitre 4, nous dérivons un modèle de force aérodynamique associé à un modèle Galerkin pour les deux premières bifurcations du pinball fluidique. L'identification de la formule de traînée et de portance est simplifiée en exploitant la symétrie modale dans le cadre de la considération du champ moyen et d'une calibration parcimonieuse. Sur la base du modèle de champ moyen d'ordre minimal, les modèles de force Galerkin de base sont dérivés pour la bifurcation de Hopf supercritique et pour la bifurcation de la fourche supercritique. Ensuite, les modèles de force pour le cas de couplage de deux bifurcations successives sont étudiés avec des modes élémentaires. Les avantages de cette méthodologie sont discutés en comparaison avec l'approche purement basée sur la projection et le modèle de régression basé sur POD.

Dans le chapitre 5, nous proposons une méthodologie de modélisation d'ordre réduit basée sur des clusters hiérarchiques auto-supervisés pour les bifurcations et l'identification de la topologie des dynamiques complexes. Nous commençons par le processus standard d'un modèle de réseau basé sur les clusters et introduisons la structure hiérarchique pour permettre au modèle de réseau de prendre en charge des dynamiques transitoires plus complexes. Cette méthodologie est une stratégie de modélisation universelle pour identifier les dynamiques transitoires et post-transitoires de manière auto-supervisée. Les HiCNM sont appliqués avec succès aux dynamiques transitoires et post-transitoires de trajectoires multiples à trois nombres de Reynolds : pour le cas impliquant six solutions exactes à $Re = 80$, pour le cas quasi-périodique à $Re = 105$, et pour le cas chaotique à $Re = 130$, respectivement. De plus, il promet d'automatiser l'identification et l'analyse de dynamiques complexes avec plusieurs attracteurs et plusieurs échelles.

Le chapitre 6 conclut les contributions de cette thèse et fournit quelques perspectives pour les travaux futurs.

A.6 Résumé

Dans cette thèse, nous avons présenté un exemple de référence stimulant pour la modélisation automatisable d'ordre réduit en utilisant les premiers principes et l'apprentissage automatique. Nous avons choisi l'écoulement de sillage incompressible bidimensionnel autour du pinball fluidique pour les caractéristiques générales du sillage rapportées dans ce travail. Des simulations numériques directes (DNS) à différentes valeurs du nombre de Reynolds sont effectuées pour chaque régime d'écoulement, en commençant par les solutions stationnaires existantes. La dynamique transitoire et post-transitoire de l'écoulement instable a montré divers comportements transitoires, impliquant les solutions stationnaires, les solutions périodiques, les attracteurs quasi-périodiques et chaotiques. La modélisation d'ordre réduit pour la dynamique complexe issue de l'interaction entre les structures dominantes de différentes instabilités est sans aucun doute une tâche difficile.

Pour obtenir un modèle parcimonieux interprétable, nous avons commencé par les travaux analytiques de cette configuration d'écoulement. L'analyse numérique et théorique des équations gouvernantes est utilisée pour obtenir une compréhension globale des mécanismes sous-jacents, par des outils mathématiques avec l'analyse de stabilité linéaire des solutions stationnaires, l'analyse de stabilité de Floquet des solutions périodiques, et l'analyse non linéaire de la dynamique asymptotique. Ceci a contribué au processus de modélisation et a conduit aux trois différents modèles de champ moyen, y compris le modèle de champ moyen d'ordre minimal dans le chapitre 2, pour les degrés de liberté élémentaires des bifurcations successives ; le modèle de force de Galerkin complémentaire au modèle de champ moyen dans le chapitre 4, reliant le modèle de Galerkin dans l'espace d'état à la dynamique instationnaire des forces; et le modèle d'ordre réduit hiérarchique basé sur les clusters dans le chapitre 5, s'adaptant à des dynamiques plus complexes avec des comportements multi-échelles, multi-fréquences, multi-attracteurs pour ou les régimes quasi-periodiques et chaotiques. Les modèles de champ moyen proposés dans ce travail ont permis de mieux comprendre le couplage entre les fluctuations et l'écoulement moyen, et d'améliorer les performances et l'applicabilité des modèles basés sur les données. Dans ce qui suit, nous résumons les contributions de chaque chapitre.

A.6.1 Résumé du Chapitre 2 – Modèle de champ moyen d'ordre minimal

Nous proposons le premier modèle Galerkin d'ordre minimal d'un écoulement incompressible subissant deux bifurcations supercritiques successives de type Hopf et Pitchfork. Un élément clé est la considération du champ moyen qui exploite la symétrie de l'écoulement moyen et l'asymétrie de la fluctuation. Ces symétries généralisent la théorie du champ moyen, par exemple, aucune hypothèse de taux de croissance lent n'est nécessaire. Le modèle de Galerkin à 5 dimensions qui en résulte décrit avec succès le phénoménogramme du pinball fluidique, un écoulement de sillage bidimensionnel autour d'un groupe de trois cylindres espacés de façon équidistante. Il est démontré que le scénario de transition correspondant subit deux bifurcations supercritiques successives, à savoir une bifurcation de Hopf et une bifurcation fourche sur la voie du chaos. La méthodologie Galerkin à champ moyen généralisé peut être utilisée pour décrire d'autres scénarios de transition.

A.6.2 Résumé du Chapitre 3 – Bifurcations locales coïncidentes

Généralement, une bifurcation locale n'affecte qu'une seule branche de solution. Cependant, des branches très différentes peuvent néanmoins partager certains vecteurs propres et valeurs propres, conduisant à des bifurcations coïncidentes. Pour le pinball fluidique, deux bifurcations fourches supercritiques, de la solution d'équilibre et de la solution périodique, se produisent presque au même nombre de Reynolds. Le mécanisme de ce type de coïncidence non-générique est modélisé et expliqué.

A.6.3 Résumé du Chapitre 4 – Modèle de force de Galerkin

Nous proposons un modèle de force aérodynamique associé à un modèle Galerkin pour le pinball fluidique non forcé, l'écoulement bidimensionnel autour de trois cylindres égaux à une distance radiale l'un de l'autre. Le point de départ est un modèle Galerkin d'un écoulement autour des corps non profilés. Les forces sur ce corps sont dérivées comme des fonctions constantes, linéaires et quadratiques des amplitudes de mode à partir des premiers principes, suivant le travail pionnier de Noca (1997); Noca *et al.* (1999) et Liang & Dong (2014). Le modèle de force est simplifié pour le modèle de champ moyen du pinball fluidique non forcé (Deng *et al.*, 2020) en utilisant les propriétés de symétrie et la calibration parcimonieuse. Le modèle est appliqué avec succès à la dynamique transitoire et post-transitoire dans différents régimes de nombre de Reynolds : l'éjection périodique de vortex après la bifurcation de Hopf et l'éjection asymétrique de vortex après la bifurcation fourche comprenant six solutions différentes de Navier-Stokes. Nous prévoyons de nombreuses applications du modèle de force de Galerkin pour d'autres corps non profilés et des écoulements forcés.

A.6.4 Résumé du Chapitre 5 – Modèle de réseau hiérarchique de clusters

Nous proposons une méthodologie de modélisation d'ordre réduit basée sur des clusters hiérarchiques auto-supervisés pour modéliser et analyser la dynamique complexe résultant d'une séquence de bifurcations pour un écoulement incompressible bidimensionnel du flipper fluidique non forcé. La hiérarchie est guidée par une triple décomposition séparant un écoulement de base variant lentement, un lâcher de tourbillons dominant et des structures d'écoulement secondaires. Toutes ces composantes de l'écoulement sont cinématiquement résolues par une hiérarchie de clusters, en commençant par l'écoulement de base dans la première couche, en résolvant le lâcher de tourbillons dans la deuxième couche et en distillant les structures d'écoulement secondaires dans la troisième couche. La dynamique de transition entre ces clusters est décrite par un graphe orienté, appelé modèle de réseau hiérarchique basé sur les clusters (HiCNM) dans la suite. Trois régimes consécutifs de nombre de Reynolds pour des dynamiques différentes sont considérés : (i)

délestage périodique à $Re = 80$, (ii) shedding quasi-périodique à $Re = 105$, et (iii) mue chaotique à $Re = 130$, impliquant trois points fixes instables, trois cycles limites, deux attracteurs quasi-périodiques et un attracteur chaotique. Le HiCNM permet d'identifier les dynamiques transitoires et post-transitoires entre plusieurs ensembles invariants de manière auto-supervisée. Les tendances globales et les structures locales pendant la transition sont bien résolues par un nombre modéré de clusters hiérarchiques. La modélisation d'ordre réduit proposée fournit une représentation visuelle du comportement transitoire et post-transitoire, multifréquence et multi-attracteur et peut automatiser l'identification et l'analyse de dynamiques complexes avec plusieurs échelles et plusieurs attracteurs.

Bibliography

Albert, R. & Barabási, A.-L. 2002 Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74** (1), 47.

Amsallem, D. & Farhat, C. 2008 Interpolation method for adapting reduced-order models and application to aeroelasticity. *AIAA journal* **46** (7), 1803–1813.

Amsallem, D., Zahr, M. J. & Farhat, C. 2012 Nonlinear model order reduction based on local reduced-order bases. *Int. J. Numer. Methods Eng.* **92** (10), 891–916.

Arthur, D. & Vassilvitskii, S. 2006 k-means++: The advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, p. 1027–1035. USA: Society for Industrial and Applied Mathematics.

Aubry, N., Holmes, P., Lumley, J. L. & Stone, E. 1988 The dynamics of coherent structures in the wall region of a turbulent boundary layer. *J. Fluid Mech.* **192**, 115–173.

Bagheri, S., Brandt, L. & Henningson, D. S. 2009*a* Input-output analysis, model reduction and control of the flat-plate boundary layer. *J. Fluid Mech.* **620** (2), 263–298.

Bagheri, S., Henningson, D. S., Hoepffner, J. & Schmid, P. J. 2009*b* Input-output analysis and control design applied to a linear model of spatially developing flows. *Appl. Mech. Rev.* **62** (2).

Bansal, M. S. & Yarusevych, S. 2017 Experimental study of flow through a cluster of three equally spaced cylinders. *Exp. Thermal Fluid Sci.* **80**, 203–217.

Bao, Y., Zhou, D. & Huang, C. 2010 Numerical simulation of flow over three circular cylinders in equilateral arrangements at low Reynolds number by a second-order characteristic-based split finite element method. *Computers & Fluids* **39** (5), 882–899.

Barabási, A.-L. 2013 Network science. *Philos. Trans. Royal Soc. A* **371** (1987), 20120375.

Barbagallo, A., Sipp, D. & Schmid, P. J. 2009 Closed-loop control of an open cavity flow using reduced-order models. *J. Fluid Mech.* **641**, 1–50.

Barkley, D. 2006 Linear analysis of the cylinder wake mean flow. *EPL (Europhysics Letters)* **75** (5), 750.

Barkley, D. & Henderson, R. 1996*a* Three-dimensional floquet stability analysis of the wake of a circular cylinder. *J. Fluid Mech.* **322**, 215–241.

Barkley, D. & Henderson, R. D. 1996*b* Three-dimensional floquet stability analysis of the wake of a circular cylinder. *J. Fluid Mech.* **322**, 215–241.

Barkley, D., Tuckerman, L. S. & Golubitsky, M. 2000 Bifurcation theory for three-dimensional flow in the wake of a circular cylinder. *Phys. Rev. E* **61** (5), 5247.

Barros, D., Borée, J., Cadot, O., Spohn, A. & Noack, B. R. 2017 Forcing symmetry exchanges and flow reversals in turbulent wakes. *J. Fluid Mech.* **829**, R1.

Beneddine, S., Sipp, D., Arnault, A., Dandois, J. & Lesshafft, L. 2016 Conditions for validity of mean flow stability analysis. *J. Fluid Mech.* **798**, 485–504.

Bengana, Y. & Tuckerman, L. S. 2021 Frequency prediction from exact or self-consistent mean flows. *Phys. Rev. Fluids* **6** (6), 063901.

Bergmann, M., Bruneau, C. H. & Iollo, A. 2009 Enablers for robust POD models. *J. Comput. Phys.* **228** (2), 516–538.

Bergmann, M. & Cordier, L. 2008 Optimal control of the cylinder wake in the laminar regime by trust-region methods and POD reduced-order models. *J. Comput. Phys.* **227** (16), 7813–7840.

Berkooz, G., Holmes, P. & Lumley, J. L. 1993 The proper orthogonal decomposition in the analysis of turbulent flows. *Ann. Rev. Fluid Mech.* **25** (1), 539–575.

Blonigan, P. J., Farazmand, M. & Sapsis, T. P. 2019 Are extreme dissipation events predictable in turbulent fluid flows? *Phys. Rev. Fluids* **4** (4), 044606.

Bollt, E. M. 2001 Combinatorial control of global dynamics in a chaotic differential equation. *Int. J. Bifurc. Chaos* **11** (08), 2145–2162.

Bonnavion, G. & Cadot, O. 2018 Unstable wake dynamics of rectangular flat-backed bluff bodies with inclination and ground proximity. *J. Fluid Mech.* **854**, 196–232.

Borońska, K. & Tuckerman, L. S. 2010 Extreme multiplicity in cylindrical rayleigh-bénard convection. ii. bifurcation diagram and symmetry classification. *Phys. Rev. E* **81** (3), 036321.

Bourgeois, J. A., Noack, B. R. & Martinuzzi, R. J. 2013 Generalised phase average with applications to sensor-based flow estimation of the wall-mounted square cylinder wake. *J. Fluid Mech.* **736**, 316–350.

Brunton, S. & Rowley, C. 2009 Modeling the unsteady aerodynamic forces on small-scale wings. In *47th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition*, p. 1127.

Brunton, S. L. & Noack, B. R. 2015 Closed-loop turbulence control: Progress and challenges. *Appl. Mech. Rev.* **67** (5), 050801:01–48.

Brunton, S. L., Noack, B. R & Koumoutsakos, P. 2020 Machine learning for fluid mechanics. *Annu. Rev. Fluid Mech.* **52**, 477–508.

Brunton, S. L., Proctor, J. L. & Kutz, J. N. 2016*a* Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Natl. Acad. Sci.* **113** (5), 3932–3937.

Brunton, S. L., Proctor, J. L. & Kutz, J. N. 2016*b* Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Natl. Acad. Sci.* **113** (15), 3932–3937.

Buffoni, M., Camarri, S., Iollo, A. & Salvetti, M. V. 2006 Low-dimensional modelling of a confined three-dimensional wake flow. *J. Fluid Mech.* **569**, 141–150.

Burkardt, J., Gunzburger, M. & Lee, H.-C. 2006 POD and CVT-based reduced-order modeling of navier–stokes flows. *Comput. Methods Appl. Mech. Eng.* **196** (1-3), 337–355.

Cadot, O., Evrard, A. & Pastur, L. 2015 Imperfect supercritical bifurcation in a three-dimensional turbulent wake. *Phys. Rev. E* **91** (6), 063005.

Cavalieri, A. V. G., Jordan, P. & Lesshafft, L. 2019 Wave-packet models for jet dynamics and sound radiation. *Appl. Mech. Rev.* **71** (2).

Chen, W., Ji, C., Alam, M. M., Williams, J. & Xu, D. 2020 Numerical simulations of flow past three circular cylinders in equilateral-triangular arrangements. *J. Fluid Mech.* **891**, A14.

Choi, H., Jeon, W.-P. & Kim, J. 2008 Control of flow over a bluff body. *Ann. Rev. Fluid Mech.* **40**, 113–139.

Chomaz, J.-M. 2005 Global instabilities in spatially developing flows: non-normality and nonlinearity. *Annu. Rev. Fluid Mech.* **37**, 357–392.

Cornejo Maceda, G. Y. 2017 Machine learning control applied to wake stabilization. MS2 Internship Report, LIMSI and ENSAM, Paris, France.

Cornejo Maceda, G. Y., Li, Y., Lusseyran, F., Morzyński, M. & Noack, B. R. 2021 Stabilization of the fluidic pinball with gradient-enriched machine learning control. *J. Fluid Mech.* **917**.

Crawford, J. D. & Knobloch, Edgar 1991 Symmetry and symmetry-breaking bifurcations in fluid dynamics. *Annual Review of Fluid Mechanics* **23** (1), 341–387.

Cross, M. C. & Hohenberg, P. C. 1993 Pattern formation outside of equilibrium. *Rev. Mod. Phys.* **65** (3), 851.

Deng, N., Noack, B. R., Morzyński, M. & Pastur, L. R. 2020 Low-order model for successive bifurcations of the fluidic pinball. *J. Fluid Mech.* **884**, A37.

Deng, N., Noack, B. R., Morzyński, M. & Pastur, L. R. 2021*a* Galerkin force model for transient and post-transient dynamics of the fluidic pinball. *J. Fluid Mech.* **918**, A4.

Deng, N., Pastur, L. R., Tuckerman, L. S. & Noack, B. R. 2021*b* Coinciding local bifurcations in the navier-stokes equations. *EPL (Europhysics Letters)* **135** (2), 24002.

Dimitriadis, G. & Li, J. 2009 Bifurcation behavior of airfoil undergoing stall flutter oscillations in low-speed wind tunnel. *AIAA journal* **47** (11), 2577–2596.

Ding, Y. & Kawahara, M. 1999 Three-dimensional linear stability analysis of incompressible viscous flows using the finite element method. *Int. J. Num. Meth. Fluids* **31** (2), 451–479.

Draper, N. R. & Smith, H. 1998 *Applied regression analysis*, , vol. 326. John Wiley & Sons.

Dušek, J., Le Gal, P. & Fraunié, P. 1994 A numerical and theoretical study of the first hopf bifurcation in a cylinder wake. *J. Fluid Mech.* **264**, 59–80.

Fabre, D., Auguste, F. & Magnaudet, J. 2008 Bifurcations and symmetry breaking in the wake of axisymmetric bodies. *Phys. Fluids* **20** (5), 051702.

Fernex, D., Noack, B. R. & Semaan, R. 2021 Cluster-based network modeling—from snapshots to complex dynamical systems. *Sci. Adv.* **7** (25), eabf5006.

Feynman, R. P., Leighton, R. B. & Sands, M. 2011 *The Feynman lectures on physics, Vol. II: Mainly Electromagnetism and Matter*, , vol. 2. Basic books.

Fine, S., Singer, Y. & Tishby, N. 1998 The hierarchical hidden Markov model: Analysis and applications. *Machine learning* **32** (1), 41–62.

Fletcher, C. A. 1984 *Computational Galerkin Methods*, 1st edn. New York: Springer.

Galletti, B., Bruneau, C. H., Zannetti, L. & Iollo, A. 2004 Low-order modelling of laminar flow regimes past a confined square cylinder. *J. Fluid Mech.* **503**, 161–170.

Gao, Y., Qu, X., Zhao, M. & Wang, L. 2019 Three-dimensional numerical simulation on flow past three circular cylinders in an equilateral-triangular arrangement. *Ocean Engineering* **189**, 106375.

Gerhard, J., Pastoor, M., King, R., Noack, B. R., Dillmann, A., Morzyński, M. & Tadmor, G. 2003 Model-based control of vortex shedding using low-dimensional galerkin models. In *33rd AIAA Fluid Dynamics Conference and Exhibit*, p. 4262.

Gomez, F., Blackburn, H. M., Rudman, M., Sharma, A. S. & McKeon, B. J. 2016 A reduced-order model of three-dimensional unsteady flow in a cavity based on the resolvent operator. *J. Fluid Mech.* **798**, R2–1..14.

Gopalakrishnan Meena, M., Nair, A. G. & Taira, K. 2018 Network community-based model reduction for vortical flows. *Phys. Rev. E* **97** (6), 063103.

Gopalakrishnan Meena, M. & Taira, K. 2021 Identifying vortical network connectors for turbulent flow modification. *J. Fluid Mech.* **915**.

Gorban, A. N. & Karlin, I. V. 2005 *Invariant Manifolds for Physical and Chemical Kinetics*. *Lecture Notes in Physics* Vol. 660. Berlin: Springer-Verlag.

Grandemange, M., Cadot, O. & Gohlke, M. 2012 Reflectional symmetry breaking of the separated flow over three-dimensional bluff bodies. *Phys. Rev. E* **86** (3), 035302.

Grandemange, M., Gohlke, M. & Cadot, O. 2013 Turbulent wake past a three-dimensional blunt body. part 1. global modes and bi-stability. *J. Fluid Mech.* **722**, 51–84.

Grandemange, M., Gohlke, M. & Cadot, O. 2014 Statistical axisymmetry of the turbulent sphere wake. *Exp. Fluids* **55** (11), 1838.

Grossmann, S., Lohse, D. & Sun, C. 2016 High–Reynolds number Taylor-Couette turbulence. *Annual Review of Fluid Mechanics* **48** (53).

Guckenheimer, J. & Holmes, P. 2013 *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, , vol. 42. Springer Science & Business Media.

Gumowski, K., Miedzik, J., Goujon-Durand, S., Jenffer, P. & Wesfreid, J. E. 2008 Transition to a time-dependent state of fluid flow in the wake of a sphere. *Phys. Rev. E* **77** (5), 055308.

Hadjighasem, A., Karrasch, D., Teramoto, H. & Haller, G. 2016 Spectral-clustering approach to lagrangian vortex detection. *Phys. Rev. E* **93** (6), 063107.

Hinze, M. & Volkwein, S. 2005 Proper orthogonal decomposition surrogate models for nonlinear dynamical systems: Error estimates and suboptimal control. In *Dimension reduction of large-scale systems*, pp. 261–306. Springer.

Holmes, P., Lumley, J. L., Berkooz, G. & Rowley, C. W. 2012*a* *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, 2nd edn. Cambridge: Cambridge University Press.

Holmes, P., Lumley, J. L., Berkooz, G. & Rowley, C. W. 2012*b* *Turbulence, coherent structures, dynamical systems and symmetry*. Cambridge university press.

Hopf, E. 1948 A mathematical example displaying features of turbulence. *Commun. Pure Appl. Math.* **1**, 303–322.

Huerre, P. & Monkewitz, P. A. 1990 Local and global instabilities in spatially developing flows. *Annu. Rev. Fluid Mech.* **22** (1), 473–537.

Ilak, M. & Rowley, C. 2006 Reduced-order modeling of channel flow using traveling POD and balanced POD. In *3rd AIAA Flow Control Conference*, p. 3194.

Iliescu, T. & Wang, Z. 2014 Variational multiscale proper orthogonal decomposition: Navier-stokes equations. *Numer. Methods Partial Differ. Equ.* **30** (2), 641–663.

Iollo, A., Lanteri, S. & Désidéri, J.-A. 2000 Stability properties of POD–Galerkin approximations for the compressible navier–stokes equations. *Theor. Comput. Fluid Dyn.* **13** (6), 377–396.

Ishar, R., Kaiser, E., Morzyński, M., Albers, M., Meysonnat, P., Schröder, W. & Noack, B. R. 2019 Metric for attractor overlap. *J. Fluid Mech.* **(in print)**.

Jiménez, J. 2018 Coherent structures in wall-bounded turbulence. *J. Fluid Mech.* **842**.

Jordan, D. W. & Smith, P. 1999 *Nonlinear ordinary differential equations: an introduction to dynamical systems*, , vol. 2. Oxford University Press, USA.

Kaiser, E., Noack, B. R., Cordier, L., Spohn, A., Segond, M., Abel, M., Daviller, G., Östh, J., Krajnović, S. & Niven, R. 2014 Cluster-based reduced-order modelling of a mixing layer. *J. Fluid Mech.* **754**, 365–414.

Kirchhoff, G. 1869 Zur theorie freier flüssigkeitsstrahlen. *J. für die Reine und Angew. Math.* **70**, 289–298.

Krueger, P. S., Hahsler, M., Olinick, E. V., Williams, S. H. & Zharfa, M. 2019 Quantitative classification of vortical flows based on topological features using graph matching. *Proc. R. Soc. A* **475** (2228), 20180897.

Kunisch, K. & Volkwein, S. 2002 Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. *SIAM J. Numer. Anal.* **40** (2), 492–515.

Kutz, J. N. 2017 Deep learning in fluid dynamics. *J. Fluid Mech.* **814**, 1–4.

Kutz, J. N., Brunton, S. L., Brunton, B. W. & Proctor, J. L. 2016 *Dynamic mode decomposition: data-driven modeling of complex systems*. SIAM.

Lam, K. & Cheung, W. C. 1988 Phenomena of vortex shedding and flow interference of three cylinders in different equilateral arrangements. *J. Fluid Mech.* **196**, 1–26.

Landau, L. D. 1944 On the problem of turbulence. *C.R. Acad. Sci. USSR* **44**, 311–314.

Landau, L. D. & Lifshitz, E. M. 1987 *Fluid Mechanics*, 2nd edn. *Course of Theoretical Physics* Vol. 6. Oxford: Pergamon Press.

LeGresley, P. A. & Alonso, J. J. 2000 Airfoil design optimization using reduced order models based on proper orthogonal decomposition. *AIAA Paper* (2000-2545).

Lesieur, M. 1987 *Turbulence in fluids: stochastic and numerical modelling*, , vol. 488. Nijhoff Boston, MA.

Lesshafft, L., Semeraro, O., Jaunet, V., Cavalieri, A. V. G. & Jordan, P. 2019 Resolvent-based modeling of coherent wave packets in a turbulent jet. *Phys. Rev. Fluids* **4** (6), 063901.

Li, H., Fernex, D., Semaan, R., Tan, J., Morzyński, M. & Noack, B. R. 2021 Cluster-based network model. *J. Fluid Mech.* **906**.

Liang, Z. & Dong, H. 2014 Virtual force measurement of POD modes for a flat plate in low reynolds number flows. In *52nd Aerospace Sciences Meeting*, p. 0054.

Liang, Z. & Dong, H. 2015 On the symmetry of proper orthogonal decomposition modes of a low-aspect-ratio plate. *Phys. Fluids* **27** (6), 063601.

Loiseau, J. C. & Brunton, S. L. 2018 Constrained sparse galerkin regression. *J. Fluid Mech.* **838**, 42–67.

Loiseau, J. C., Noack, B. R. & Brunton, S. L. 2018*a* Sparse reduced-order modeling: Sensor-based dynamics to full-state estimation. *J. Fluid Mech.* **844**, 459–490.

Loiseau, J.-C., Noack, B. R. & Brunton, S. L. 2018*b* Sparse reduced-order modelling: sensor-based dynamics to full-state estimation. *J. Fluid Mech.* **844**, 459–490.

Luchtenburg, D. M., Günter, B., Noack, B. R., King, R. & Tadmor, G. 2009 A generalized mean-field model of the natural and actuated flows around a high-lift configuration. *J. Fluid Mech.* **623**, 283–316.

Malkus, W. V. R. 1956 Outline of a theory of turbulent shear flow. *J. Fluid Mech.* **1**, 521–539.

Mamun, C. K. & Tuckerman, L. S. 1995 Asymmetry and hopf bifurcation in spherical couette flow. *Phys. Fluids* **7** (1), 80–91.

Manneville, P. 2010 *Instabilities, chaos and turbulence*, , vol. 1. World Scientific.

Mantič-Lugo, V., Arratia, C. & Gallaire, F. 2014 Self-consistent mean flow description of the nonlinear saturation of the vortex shedding in the cylinder wake. *Phys. Rev. Lett.* **113** (8), 084501.

Meliga, P., Chomaz, J.-M. & Sipp, D. 2009 Global mode interaction and pattern selection in the wake of a disk: a weakly nonlinear expansion. *J. Fluid Mech.* **633**, 159–189.

Mendez, M. A., Balabane, M. & Buchlin, J.-M. 2019 Multi-scale proper orthogonal decomposition of complex fluid flows. *J. Fluid Mech.* **870**, 988–1036.

Mezić, I. 2013 Analysis of fluid flows via spectral properties of the koopman operator. *Annu. Rev. Fluid Mech.* **45**, 357–378.

Mittal, R. 1999 Planar symmetry in the unsteady wake of a sphere. *AIAA J.* **37** (3), 388–390.

Mizushima, J. & Shiotani, Y. 2001 Transitions and instabilities of flow in a symmetric channel with a suddenly expanded and contracted part. *J. Fluid Mech.* **434** (1), 355–369.

Morzyński, M., Afanasiev, K. & Thiele, F. 1999 Solution of the eigenvalue problems resulting from global non-parallel flow stability analysis. *Comput. Methods. Appl. Mech. Engrg* **169** (1), 161 – 176.

Murayama, S., Kinugawa, H., Tokuda, I. T. & Gotoda, H. 2018 Characterization and detection of thermoacoustic combustion oscillations based on statistical complexity and complex-network theory. *Phys. Rev. E* **97** (2), 022223.

Nair, A. G. & Taira, K. 2015 Network-theoretic approach to sparsified discrete vortex dynamics. *J. Fluid Mech.* **768**, 549–571.

Nair, A. G., Yeh, C., Kaiser, E., Noack, B. R., Brunton, S. L. & Taira, K. 2019 Cluster-based feedback control of turbulent post-stall separated flows. *J. Fluid Mech.* **875**, 345–375.

Newhouse, S., Ruelle, D. & Takens, F. 1978 Occurrence of strange axiom a attractors near quasi periodic flows ont m, m $\geqq$ 3. *Commun. Math. Phys.* **64** (1), 35–40.

Newman, M. 2018 *Networks*. Oxford university press.

Noack, B. R. 2016 From snapshots to modal expansions – bridging low residuals and pure frequencies. *J. Fluid Mech.* **802**, 1–4.

Noack, B. R., Afanasiev, K., Morzyński, M., Tadmor, G. & Thiele, F. 2003 A hierarchy of low-dimensional models for the transient and post-transient cylinder wake. *J. Fluid Mech.* **497**, 335–363.

Noack, B. R. & Eckelmann, H. 1994*a* A global stability analysis of the steady and periodic cylinder wake. *J. Fluid Mech.* **270**, 297–330.

Noack, B. R. & Eckelmann, H. 1994*b* Theoretical investigation of the bifurcations and the turbulence attractor of the cylinder wake. *Z. angew. Math. Mech.* **74** (5), T396–T397.

Noack, B. R. & Morzyński, M. 2017 The fluidic pinball — a toolkit for multiple-input multiple-output flow control (version 1.0). *Tech. Rep.* 02/2017. Chair of Virtual Engineering, Poznan University of Technology, Poland.

Noack, B. R., Papas, P. & Monkewitz, P. A. 2005 The need for a pressure-term representation in empirical galerkin models of incompressible shear flows. *J. Fluid Mech.* **523**, 339–365.

Noack, B. R., Schlegel, M., Ahlborn, B., Mutschke, G., Morzyński, M., Comte, P. & Tadmor, G. 2008 A finite-time thermodynamics of unsteady fluid flows. *J. Non-Equilibr. Thermodyn.* **33**, 103–148.

Noack, B. R., Stankiewicz, W., Morzyński, M. & Schmid, P. J. 2016 Recursive dynamic mode decomposition of transient and post-transient wake flows. *J. Fluid Mech.* **809**, 843–872.

Noca, F. 1997 On the evaluation of time-dependent fluid-dynamic forces on bluff bodies. PhD thesis, California Institute of Technology.

Noca, F., Shiels, D. & Jeon, D. 1999 A comparison of methods for evaluating time-dependent fluid dynamic forces on bodies, using only velocity fields and their derivatives. *J. Fluids Struct.* **13** (5), 551–578.

Östh, J., Krajnović, S., Noack, B. R., Barros, D. & Borée, J. 2014 On the need for a nonlinear subscale turbulence term in POD models as exemplified for a high Reynolds number flow over an Ahmed body. *J. Fluid Mech.* **747**, 518–544.

Panton, R. W. 1984 *Incompressible Flow*. New York: John Wiley & Sons.

Podvin, B., Pellerin, S., Fraigneau, Y., Evrard, A. & Cadot, O. 2020 Proper orthogonal decomposition analysis and modelling of the wake deviation behind a squareback Ahmed body. *Phys. Rev. Fluids* **5**, 064612.

Prandtl, L. 1921 Applications of modern hydrodynamics to aeronautics. *Tech. Rep.*. NASA Ames Research Center Moffett Field, CA, United States.

Price, S. J. & Paidoussis, M. P. 1984 The aerodynamic forces acting on groups of two and three circular cylinders when subject to a cross-flow. *J. Wind Eng. Indust. Aero.* **17** (3), 329–347.

Protas, B. 2004 Linear feedback stabilization of laminar vortex shedding based on a point vortex model. *Phys. Fluids* **16** (12), 4473–4488.

Protas, B., Noack, B. R. & Östh, J. 2015 Optimal nonlinear eddy viscosity in Galerkin models of turbulent flows. *J. Fluid Mech.* **766**, 337–367.

Raibaudo, C., Zhong, P., Noack, B. R. & Martinuzzi, R. J. 2020 Machine learning strategies applied to the control of a fluidic pinball. *Phys. Fluids* **32** (1), 015108.

Rempfer, D. 1994 On the structure of dynamical systems describing the evolution of coherent structures in a convective boundary layer. *Phys. Fluids* **6** (3), 1402–4.

Rempfer, D. 2000 On low-dimensional galerkin models for fluid flow. *Theoret. Comput. Fluid Dynamics* **14** (2), 75–88.

Rempfer, D. & Fasel, F. H. 1994*a* Dynamics of three-dimensional coherent structures in a flat-plate boundary-layer. *J. Fluid Mech.* **275**, 257–283.

Rempfer, D. & Fasel, H. F. 1994*b* Evolution of three-dimensional coherent structures in a flat-plate boundary-layer. *J. Fluid Mech.* **260**, 351–375.

Reynolds, W. C. & Hussain, A. K. M. F. 1972 The mechanics of an organized wave in turbulent shear flow. Part 3. Theoretical model and comparisons with experiments. *J. Fluid Mech.* **54**, 263–288.

Rigas, G., Morgans, A. S. & Morrison, J. F. 2017*a* Weakly nonlinear modelling of a forced turbulent axisymmetric wake. *J. Fluid Mech.* **814**, 570–591.

Rigas, G., Oxlade, A. R., Morgans, A. S. & Morrison, J. F. 2014 Low-dimensional dynamics of a turbulent axisymmetric wake. *J. Fluid Mech.* **755**.

Rigas, G., Schmidt, O. T., Colonius, T. & Bres, G. A. 2017*b* One-way navier-stokes and resolvent analysis for modeling coherent structures in a supersonic turbulent jet. In *23rd AIAA/CEAS Aeroacoustics Conference*, p. 4046.

Roweis, S. T. & Saul, L. K. 2000 Nonlinear dimensionality reduction by locally linear embedding. *Science* **290** (5500), 2323–2326.

Rowley, C. W. 2005 Model reduction for fluids, using balanced proper orthogonal decomposition. *Int. J. Bifurcat. Chaos* **15** (03), 997–1013.

Rowley, C. W., Colonius, T. & Basu, A. J. 2002 On self-sustained oscillations in two-dimensional compressible flow over rectangular cavities. *J. Fluid Mech.* **455**, 315–346.

Rowley, C. W., Colonius, T. & Murray, R. M. 2004 Model reduction for compressible flows using POD and galerkin projection. *Physica D* **189** (1-2), 115–129.

Rowley, C. W. & Dawson, S. T. 2017 Model reduction for flow analysis and control. *Ann. Rev. Fluid Mech.* **49**, 387–417.

Rowley, C. W., Mezić, I., Bagheri, S., Schlatter, P. & Henningson, D. S. 2009 Spectral analysis of nonlinear flows. *J. Fluid Mech.* **645**, 115–127.

Ruelle, D. & Takens, F. 1971 On the nature of turbulence. *Les rencontres physiciens-mathématiciens de Strasbourg-RCP25* **12**, 1–44.

San, O. & Maulik, R. 2018 Extreme learning machine for reduced order modeling of turbulent geophysical flows. *Phys. Rev. E* **97** (4), 042322.

San, O., Maulik, R. & Ahmed, M. 2019 An artificial neural network framework for reduced order modeling of transient flows. *Commun. Nonlinear Sci. Numer. Simul.* **77**, 271–287.

Sapsis, T. P. 2021 Statistics of extreme events in fluid flows and waves. *Annu. Rev. Fluid Mech.* **53**, 85–111.

Sapsis, T. P. & Lermusiaux, P. F. J. 2009 Dynamically orthogonal field equations for continuous stochastic dynamical systems. *Physica D* **238** (23-24), 2347–2360.

Sapsis, T. P. & Majda, A. J. 2013 Statistically accurate low-order models for uncertainty quantification in turbulent dynamical systems. *Proc. Natl. Acad. Sci. U.S.A.* **110** (34), 13705–13710.

Sayers, A. T. 1987 Flow interference between three equispaced cylinders when subjected to a cross flow. *J. Wind Eng. Indust. Aero.* **26** (1), 1–19.

Schatz, M. F., Barkley, D. & Swinney, H. L. 1995 Instability in a spatially periodic open flow. *Phys. Fluids* **7** (2), 344–358.

Schewe, G. 1983 On the force fluctuations acting on a circular cylinder in crossflow from subcritical up to transcritical Reynolds numbers. *J. Fluid Mech.* **133**, 265–285.

Schlegel, M. & Noack, B. R. 2015 On long-term boundedness of galerkin models. *J. Fluid Mech.* **765**, 325–352.

Schlichting, H. & Gersten, K. 2016 *Boundary-layer theory*. Springer.

Schlueter-Kuck, K. L. & Dabiri, J. O. 2017 Coherent structure colouring: identification of coherent structures from sparse data using graph theory. *J. Fluid Mech.* **811**, 468–486.

Schmid, P. J. 2010 Dynamic mode decomposition for numerical and experimental data. *J. Fluid Mech.* **656**, 5–28.

Schumm, M., Berger, E. & Monkewitz, P. A. 1994 Self-excited oscillations in the wake of two-dimensional bluff bodies and their control. *J. Fluid Mech.* **271**, 17–53.

Shaabani-Ardali, L., Sipp, D. & Lesshafft, L. 2019 Vortex pairing in jets as a global floquet instability: modal and transient dynamics. *J. Fluid Mech.* .

Shaabani-Ardali, L., Sipp, D. & Lesshafft, L. 2020 Optimal triggering of jet bifurcation: an example of optimal forcing applied to a time-periodic base flow. *J. Fluid Mech* **885** (A34).

Shearer, M. 1981 Coincident bifurcation of equilibrium and periodic solutions of evolution equations. *J. Math. Anal. Appl.* **84** (1), 113–132.

Sieber, M., Paschereit, C. O. & Oberleithner, K. 2016 Spectral proper orthogonal decomposition. *J. Fluid Mech.* **792**, 798–828.

Sipp, D. & Lebedev, A. 2007 Global stability of base and mean flows: a general approach and its applications to cylinder and open cavity flows. *J. Fluid Mech.* **593** (1), 333–358.

Smyth, W. D. & Moum, J. N. 2012 Ocean mixing by kelvin-helmholtz instability. *Oceanography* **25** (2), 140–149.

Strogatz, S., Friedman, M., Mallinckrodt, A. J. & McKay, S. 1994 Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering. *Computers Phys.* **8** (5), 532–532.

Strykowski, P. J. & Sreenivasan, K. R. 1990 On the formation and suppression of vortex 'shedding' at low reynolds numbers. *J. Fluid Mech.* **218**, 71–107.

Stuart, J. T. 1958 On the non-linear mechanics of hydrodynamic stability. *J. Fluid Mech.* **4**, 1–21.

Stuart, J. T. 1971 Nonlinear stability theory. *Ann. Rev. Fluid Mech.* **3**, 347–370.

Swift, J. & Hohenberg, P. C. 1977 Hydrodynamic fluctuations at the convective instability. *Phys. Rev. A* **15** (1), 319.

Szaltys, P., Chrust, M., Przadka, A., Goujon-Durand, S., Tuckerman, L. S. & Wesfreid, J. E. 2012 Nonlinear evolution of instabilities behind spheres and disks. *J. Fluids Struct.* **28**, 483–487.

Tadmor, G., Lehmann, O., Noack, B. R., Cordier, L., Delville, J., Bonnet, J.-P. & Morzyński, M. 2011 Reduced order models for closed-loop wake control. *Philos. Trans. R. S. A* **369** (1940), 1513–1524.

Taira, K., Brunton, S. L., Dawson, S. T., Rowley, C. W., Colonius, T., McKeon, B. J., Schmidt, O. T., Gordeyev, S., Theofilis, V. & Ukeiley, L. S. 2017 Modal analysis of fluid flows: An overview. *AIAA J.* pp. 4013–4041.

Taira, K., Nair, A. G. & Brunton, S. L. 2016 Network structure of two-dimensional decaying isotropic turbulence. *J. Fluid Mech.* **795**.

Tatsuno, M., Amamoto, H. & Ishi-i, K. 1998 Effects of interference among three equidistantly arranged cylinders in a uniform flow. *Fluid Dyn. Res.* **22** (5), 297.

Taylor, C. & Hood, P. 1973 A numerical solution of the navier-stokes equations using the finite element technique. *Comput. Fluids* **1**, 73–100.

Towne, A., Schmidt, O. T. & Colonius, T. 2018 Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis. *J. Fluid Mech.* **847**, 821–867.

Tu, J. H., Rowley, C. W., Luchtenburg, D. M., Brunton, S. L. & Kutz, J. N. 2014 On dynamic mode decomposition: Theory and applications. *J. Comput. Dyn.* **1** (2).

Tuckerman, L. S. & Barkley, D. 2000 Bifurcation analysis for timesteppers. In *Numerical methods for bifurcation problems and large-scale dynamical systems*, pp. 453–466. Springer.

Turton, S. E., Tuckerman, L. S. & Barkley, D. 2015 Prediction of frequencies in thermosolutal convection from mean flows. *Phys. Rev. E* **91** (4), 043009.

Wang, Z., Akhtar, I., Borggaard, J. & Iliescu, T. 2012 Proper orthogonal decomposition closure models for turbulent flows: a numerical comparison. *Comput. Methods Appl. Mech. Eng.* **237**, 10–26.

Watson, J. 1960 On the non-linear mechanics of wave disturbances in stable and unstable parallel flows. Part 2. the development of a solution for plane Poiseuille flow and for plane Couette flow. *J. Fluid Mech.* **9**, 371–389.

Watts, D. J. & Strogatz, S. H. 1998 Collective dynamics of 'small-world' networks. *Nature* **393** (6684), 440–442.

Wiener, N. 1948 *Cybernetics or Control and Communication in the Animal and the Machine*. Technology Press.

Willcox, K. & Peraire, J. 2002 Balanced model reduction via the proper orthogonal decomposition. *AIAA journal* **40** (11), 2323–2330.

Yeh, C.-A., Gopalakrishnan Meena, M., M. & Taira, K. 2021 Network broadcast analysis and control of turbulent flows. *J. Fluid Mech.* **910**, A15.

Zaitsev, V. M. & Shliomis, M. I. 1971 Hydrodynamic fluctuations near convection threshold. *Sov. Phys. JETP* **32**, 866.

Zhang, H.-Q., Fey, U., Noack, B. R., König, M. & Eckelmann, H. 1995 On the transition of the cylinder wake. *Phys. Fluids* **7** (4), 779–794.

Zhang, H.-Q., Noack, B. R. & Eckelmann, H. 1994 Numerical computation of the 3-d cylinder wake. *Tech. Rep.* 3/1994. Max-Planck-Institut für Strömungsforschung, Göttingen, Germany.

Zheng, S., Zhang, W. & Lv, X. 2016 Numerical simulation of cross-flow around three equal diameter cylinders in an equilateral-triangular configuration at low Reynolds numbers. *Computers & Fluids* **130**, 94–108.

Zielinska, B. J. A. & Wesfreid, J. E. 1995 On the spatial structure of global modes in wake flow. *Phys. Fluids* **7** (6), 1418–1424.

Publications

JOURNAL PAPERS

Published

Deng, N., Pastur, L. R., Tuckerman L. S., & Noack B. R. 2021 Coinciding local bifurcations in the Navier-Stokes equations. *EPL (Europhysics Letters)* **135(2)**, 24002.

Deng, N., Noack, B. R., Morzyński, M. & Pastur, L. R. 2021 Galerkin force model for transient dynamics of the fluidic pinball. *J. Fluid Mech.* **918**, A34.

Deng, N., Pastur, L. R. & Noack, B. R. 2020 Low-dimensional Flow Models from high-dimensional Flow data with Machine Learning and First Principles. *ERCIM News* **122**: Solving Engineering Problems with Machine Learning, 30-31.

Deng, N., Noack, B. R., Morzyński, M. & Pastur, L. R. 2020 Low-order model for successive bifurcations of the fluidic pinball. *J. Fluid Mech.* **884**, A37.

Cornejo Maceda, G. Y., Noack B. R., Lusseyran, F., Deng, N., Pastur, L. R. & Morzyński, M. 2019 Artificial intelligence control applied to drag reduction of the fluidic pinball. *Proc. Appl. Math. Mech.* **19**(1), article e201900268, 1–2.

Submitted

Deng, N., Noack, B. R., Morzyński, M. & Pastur, L. R. Cluster-based hierarchical network model of the fluidic pinball — Cartographing transient and post-transient, multi-frequency, multi-attractor behaviour.

CONFERENCE CONTRIBUTIONS

Conference Papers

Deng, N., Pastur, L. R., Noack, B. R., Cornejo Maceda, G. Y., Lusseyran, F., Loiseau, J.-C. & Morzyński, M. 2019 Dynamiques transitoires de sillage dans le «pinball fluidique». Paper, *Le 22e rencontre du non-linéaire*, Paris, France, 2019. Pages 15-20.

Cornejo Maceda, G. Y., Noack, B. R., Lusseyran, F., Deng, N., Pastur, L. & Morzyński, M. 2019 Apprentissage automatique de lois de contrôle d'écoulement par programmation génétique. Paper at *24éme Congrés Français de Mécanique*, Brest, France.

Deng, N., Pastur, L. R., Morzyński, M. & Noack, B. R. 2019 Reduced-Order Modeling of the Fluidic Pinball. In: Skiadas C., Lubashevsky I. (eds) *11th Chaotic Modeling and Simulation International Conference*. CHAOS 2018. Springer Proceedings in Complexity. Springer, Cham.

Deng, N., Pastur, L. R., Morzyński, M. & Noack, B. R. 2018 Route to chaos in the pinball fluidique. Paper FEDSM2018-83359, *Proceedings of the ASME 2018 5th Joint US- European Fluids Engineering Summer*, Conference, Montreal, Quebec, Canada

Oral Presentations

Deng, N., Pastur, L. R., Morzyński, M. & Noack, B. R., 2021 Least-order mean-field model for the transient and post-transient dynamics of the fluidic pinball. In: *14th WCCM-ECCOMAS Congress*, Virtual.

Deng, N., Pastur, L. R., Morzyński, M. & Noack, B. R., 2020 Least-order model and Galerkin force model for transient and post-transient dynamics of the fluidic pinball. In: *Division of Fluid Dynamics 73rd meeting*, Virtual.

Deng, N., Pastur, L. R., Morzyński, M. & Noack, B. R., 2020 Galerkin force model for transient and post-transient dynamics of the fluidic pinball. In: *GDR2502 "Flow Separation Control" meeting*, Virtual.

Deng, N., Pastur, L. R., Noack, B. R., Cornejo Maceda, G. Y., Lusseyran, F., Loiseau, J.-C. & Morzyński, M. 2019 Dynamiques transitoires de sillage dans le pinball fluidique. In: *Le 22e rencontre du non-linéaire*, Paris, France.

Deng, N., Pastur, L. R., Noack, B. R. & Morzyński, M. 2019 Least-order mean-field models for two generic instabilities commonly encountered in fluid mechanics: illustration on the fluidic pinball. In: *8th International Symposium on Bifurcations and Instabilities in Fluid Dynamics (BIFD 2019)*, University of Limerick, Ireland.

Deng, N., Noack, B. R., Pastur, L. R. & Morzyński, M. 2019 Reduced-order modeling of the fluidic pinball. In: *Workshop Journée de Dynamique des Fluides du Plateau de Saclay*, LIMSI-CNRS, Orsay, France.

Poster Presentations

Deng, N., Pastur, L. R., Noack, B. R., Cornejo Maceda, G. Y., Lusseyran, F. & Morzyński, M. 2020 Analysis and control of flow with human and machine learning: illustration on the fluidic Pinball. In: *PhD welcome day - IP Paris 2020*, Palaiseau, France.

Deng, N., Pastur, L. R., Noack, B. R. & Morzyński, M. 2020 Least-order model for the transient dynamics of the fluidic pinball. In: *VKI Lecture series 2020: Machine Learning for Fluid Mechanics: Analysis, Modeling, Control and Closures*, Brussels, Belgium.

Deng, N., Pastur, L. R., Noack, B. R., Cornejo Maceda, G. Y., Lusseyran, F., Loiseau, J.-C. & Morzyński, M. 2019 Dynamiques transitoires de sillage dans le pinball fluidique. In: *Le 22e rencontre du non-linéaire*, Paris, France.

ECOLE
DOCTORALE

Titre : Modélisation en champ moyen de bifurcations successives, illustrée sur le pinball fluidique

Mots clés : Champ moyen, stabilité, bifurcation, modèle de Galerkin, modèle de réseau, modèle automatisé.

Résumé : Cette thèse ouvre la voie à une modélisation réduite automatisable en mécanique des fluides, à partir de principes premiers et de techniques d'apprentissage automatique. L'objectif est d'établir un problème de référence présentant les caractéristiques dynamiques principales des écoulements de sillage.
La configuration de référence choisie est l'écoulement de sillage incompressible bidimensionnel autour du "pinball fluidique". À bas nombre de Reynolds, l'écoulement stationaire, symétrique par réflexion miroir, est stable. Lorsque le nombre de Reynolds augmente, le système subit deux bifurcations supercritiques de type Hopf et fourche, associées respectivement à l'instabilité de Bénard-von Kármán et à une instabilité de brisure de symétrie. Une bifurcation de Hopf secondaire conduit à un lâcher de tourbillons asymétrique quasi-périodique, avant de bifurquer finalement vers un régime chaotique. Le régime périodique asymétrique est caractérisé par trois solutions stables et trois cycles limites, ce qui pose un défi à la modélisation automatique.
En prélude à la modélisation, la compréhension des mécanismes sous-jacents est menée par analyse de stabilité linéaire des solutions stationnaires, analyse de Floquet des solutions périodiques et analyse non linéaire de la dynamique asymptotique. Ces analyses, aux côtés de la méthode de Galerkin, constituent le point de départ de notre stratégie de modélisation en champ moyen. Un modèle de champ moyen à cinq dimensions reproduit les six ensembles invariants qui résultent des deux premières bifurcations. Nous dérivons également un modèle de forces aérodynamiques à partir du modèle de Galerkin en champ moyen. Une calibration parcimonieuse permet d'obtenir un modèle facilement interprétable, équilibrant précision et complexité.
Pour les régimes quasi-périodiques et chaotiques, nous proposons un modèle de réseau hiérarchique de clusters (HiCNM), capable de s'adapter à des dynamiques plus complexes présentant un comportement multi-échelles, multi-fréquences et multi-attracteurs. Le HiCNM permet d'identifier les dynamiques transitoires et post-transitoires entre de multiples ensembles invariants d'une manière auto-supervisée, ce qui constitue une étape vers la réduction de modèle automatisée des dynamiques complexes.

Title : Deep mean-field modelling for successive bifurcations exemplified for the fluidic pinball

Keywords : Mean-field model, stability, bifurcation, Galerkin model, network model, automated model.

Abstract : This thesis paves the way to automatable ROM in flow dynamics using first principles and machine learning techniques. The aim is to establish a benchmark problem for the most important dynamical features of wake flows.
The chosen benchmark configuration is the two-dimensional incompressible wake flow around the fluidic pinball. At low Reynolds numbers, this configuration has a stable steady state satisfying the reflectional symmetry. With increasing Reynolds numbers, it undergoes two supercritical bifurcations of Hopf and pitchfork types, associated with the Bénard-von Kármán instability and a symmetry-breaking instability, respectively. A secondary Hopf bifurcation leads to quasi-periodic asymmetric shedding, before finally bifurcating into a chaotic regime. The asymmetric periodic shedding is characterized by three steady solutions and three limit cycles, which evidently poses a challenge to automated modelling.
Before modelling, a comprehensive understanding of the underlying mechanisms is pursued, including linear stability analysis of steady solutions, Floquet analysis of periodic solutions, and nonlinear analysis of asymptotic dynamics. These analyses, together with the Galerkin method, are the starting point of our mean-field modelling strategy. A five-dimensional least-order mean-field model is proposed, resolving the six invariant sets induced by the first two successive bifurcations. In addition, we derive an aerodynamic force model associated with the mean-field Galerkin model. Sparse calibration is applied to balance the accuracy and complexity of the model. These efforts culminate in a sparse human interpretable model for the flow dynamics and a predictive model for the unsteady forces.
For quasi-periodic and chaotic regimes, we propose a hierarchical cluster-based network modelling (HiCNM), adapting to more complex dynamics with multi-scale, multi-frequency, multi-attractor behaviours. The HiCNM enables identifying the transient and post-transient dynamics between multiple invariant sets in a self-supervised manner and steps towards automated ROM of complex dynamics.

Institut Polytechnique de Paris
91120 Palaiseau, France