Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica D 189 (2004) 115–129
PHYSICA D
www.elsevier.com/locate/physd

Model reduction for compressible flows using POD and Galerkin projection

Clarence W. Rowley [a,*], Tim Colonius [b], Richard M. Murray [b]

[a] *Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544, USA*
[b] *Mechanical Engineering, California Institute of Technology, Pasadena, CA 91125, USA*

Received 25 March 2002; received in revised form 27 January 2003; accepted 25 March 2003

Abstract

We present a framework for applying the method of proper orthogonal decomposition (POD) and Galerkin projection to compressible fluids. For incompressible flows, only the kinematic variables are important, and the techniques are well known. In a compressible flow, both the kinematic and thermodynamic variables are dynamically important, and must be included in the configuration space. We introduce an energy-based inner product which may be used to obtain POD modes for this configuration space. We then obtain an approximate version of the Navier–Stokes equations, valid for cold flows at moderate Mach number, and project these equations onto a POD basis. The resulting equations of motion are quadratic, and are much simpler than projections of the full compressible Navier–Stokes equations.

PACS: 47.40; 02.70.D; 47.11

Keywords: Proper orthogonal decomposition; Compressible flows; Galerkin projection; Model reduction

1. Introduction

The tools of proper orthogonal decomposition (POD) and Galerkin projection have been used for some time to obtain low-dimensional models of fluids [1–3], but most of the applications have been to incompressible flows. This paper presents a method for applying these techniques to compressible flows.

In an incompressible flow, the velocity is the only flow variable that is dynamically important. The pressure acts only to enforce the incompressibility constraint, and may in fact be eliminated completely from the equations of motion, as we review in Section 3. By contrast, in a compressible flow, the pressure becomes dynamically important, along with all the other thermodynamic variables. As a result, additional evolution equations must be specified for these new variables. Furthermore, even the choice of inner product is no longer obvious, as it is not immediately clear how to combine both thermodynamic and kinematic variables in a rational way (e.g., such that the projected equations behave well, or such that the induced norm has a physical meaning, such as energy). In

* Corresponding author.
E-mail address: cwrowley@princeton.edu (C.W. Rowley).

0167-2789/$ – see front matter © 2003 Published by Elsevier B.V.
doi:10.1016/j.physd.2003.03.001

addition, the equations of motion become much more complicated, and Galerkin projections of the fully compressible Navier–Stokes equations are rather unwieldy.

In this paper, we first give an overview of the general method of POD and Galerkin projection, in the context of an abstract Hilbert space. We briefly outline how these tools have been applied to incompressible flows in Section 3, and in Section 4, we discuss the difficulties that arise for compressible flow, and present the details of our new method. We introduce a family of energy-based inner products, and present a simplified version of the compressible Navier–Stokes equations, valid for flows with small temperature gradients and moderate Mach numbers, that greatly simplifies the resulting Galerkin projections.

2. Model reduction

The tools of POD and Galerkin projection provide a systematic way for producing reduced-order models from data. The central idea of POD is to determine a nested family of subspaces, of increasing (finite) dimension, that optimally span the data, in the sense that the error in the projection onto each subspace is minimized. Galerkin projection then determines dynamics on each subspace, by orthogonal projection of the governing equations.

In this section, we give an overview of these methods, in the context of an abstract Hilbert space. The abstract setting is used for several reasons. In subsequent sections, we will want to consider different inner products, and keeping the exposition general allows one to see precisely where the dependence on the inner product lies. This is especially useful for computations, as one can write a single subroutine to compute the inner product, and if a different inner product is desired, only the one routine needs to be changed. In the usual development of POD [1,3], the standard L^2 inner product is usually assumed, and it is not always clear where this assumption has been used, and what routines need to be modified if a different inner product is desired.

2.1. Proper orthogonal decomposition

Let H be a Hilbert space, with inner product $\langle \cdot, \cdot \rangle$. The goal is, given an ensemble of data $\{u_k \in H | k = 1, \ldots, m\}$, find a subspace S of fixed dimension $n < m$, such that the error $E(\|u_k - P_S u_k\|)$ is minimized. Here, $\| \cdot \|$ is the induced norm on H, P_S is the orthogonal projection onto the subspace, and $E(\cdot)$ denotes an average over k. The data $\{u_k\}$ could be thought of as an ensemble of many different experiments, or as a time average, with u_k representing snapshots of a function $u(t)$ at different times $t = t_k$. Note that minimizing the error $E(\|u_k - P_S u_k\|)$ is equivalent to maximizing $E(\|P_S u_k\|^2)$, the "energy" in the projection, since $\|u\|^2 = \|u - Pu\|^2 + \|Pu\|^2$ for any orthogonal projection P.

For our purposes, the space H will consist of functions on some spatial domain Ω in which a fluid evolves, for instance, $H = L^2(\Omega)$. These functions may be vector-valued, and in applications we will explicitly state which inner product we use.

Solving this optimization problem leads to the eigenvalue problem:

$$R\varphi = \lambda\varphi, \tag{2.1}$$

where $R : H \to H$ is the linear operator given by

$$R = E(u_k \otimes u_k^*). \tag{2.2}$$

Here, $u^* \in H^*$ denotes the dual of u, the functional given by $u^*(\cdot) = \langle \cdot, u \rangle$, and $\otimes$ is the usual tensor product [4]. Thus, $(u \otimes v^*)(\psi) = u\langle \psi, v \rangle$ for any $u, v, \psi \in H$. It follows easily from the definition that R is self-adjoint, so

the eigenfunctions φ may indeed be chosen to be orthonormal. Furthermore, we see from (2.1) by taking an inner product with φ that

$$\lambda = E(|\langle u_k, \varphi \rangle|^2), \tag{2.3}$$

so the eigenvalue λ is the average energy in the projection of the ensemble onto φ, where the energy is in the sense of the induced norm. From (2.3), we conclude that R is positive semi-definite ($\lambda \geq 0$), and the functions φ_j which maximize $E(\|P_S u_k\|^2)$ are the eigenfunctions corresponding to the largest n eigenvalues of R. These eigenfunctions are called *POD modes*.

2.1.1. Computation: method of snapshots

The method of snapshots provides an alternate way of computing the POD modes, and is often more efficient than the *direct method*, where one directly solves the eigenvalue problem (2.1).

The main idea is to write the POD modes as linear combinations of the snapshots u_k. This can always be done, because the range of $R = E(u_k \otimes u_k^*)$ is contained within the span of the ensemble $\{u_k\}$. Thus, any eigenfunction φ may be written as

$$\varphi = \sum_k c_k u_k \tag{2.4}$$

for some coefficients $c_k \in \mathbb{R}$ (or $\mathbb{C}$). Define the average $E(\cdot)$ as a weighted average over the snapshots u_k:

$$E(f(u)) = \sum_{k=1}^{m} \alpha_k f(u_k),$$

where the weights $\alpha_k > 0$ satisfy $\sum_k \alpha_k = 1$ (typically, $\alpha_k = 1/m$, for equal weighting). The eigenvalue problem $R\varphi = \lambda\varphi$ may then be rewritten in terms of the coefficients c_k:

$$Uc = \lambda c, \tag{2.5}$$

where $c = (c_1, \ldots, c_m)$ and U is an $m \times m$ matrix with $U_{ij} = \alpha_i \langle u_j, u_i \rangle$.

The direct solution of (2.1) involves solving an eigenvalue problem on the space H, which may be infinite-dimensional. For instance, if the snapshots $u_k \in H$ are data from a simulation, the dimension of H will be the number of gridpoints (or spectral modes) in the simulation, typically large. By contrast, the method of snapshots involves solving an m-dimensional eigenvalue problem (2.5), where m is the number of snapshots in the ensemble. The method of snapshots is thus more efficient whenever the number of snapshots is smaller than the number of gridpoints.

2.2. Galerkin projection

Consider a dynamical system which evolves in a Hilbert space H. In particular, for $u(t) \in H$, $u(t)$ satisfies

$$\dot{u} = X(u), \tag{2.6}$$

where X is a vector field on H. For instance, for a partial differential equation governing a variable $u(x, t)$, defined on some spatial domain $x \in \Omega$, H will be a space of functions defined on Ω, and X will be a spatial differential operator. Given a finite-dimensional subspace S of H, Galerkin projection specifies a dynamical system which evolves on S and approximates (2.6) in some sense. This approximate dynamical system is obtained by orthogonal projection of the vector field X onto the subspace, and is denoted

$$\dot{r} = P_S X(r), \tag{2.7}$$

where $r(t) \in S$ and $P_S : H \to S$ is the orthogonal projection map.

To apply this method to computations, we need to write (2.7) in coordinates. Let $\{\varphi_k \in H | k = 1, \ldots, n\}$ be an orthonormal basis for the subspace S. (For instance, such a basis may be obtained from POD on a set of data, as described in the preceding section.) Writing $r(t)$ in coordinates $a_k(t)$ with respect to this basis, we have

$$r(t) = \sum_{k=1}^{n} a_k(t)\varphi_k. \tag{2.8}$$

From (2.7), we have $P_S(\dot{r} - X(r)) = 0$, and the equations of motion become

$$\dot{a}_k(t) = \langle X(r(t)), \varphi_k \rangle, \quad k = 1, \ldots, n, \tag{2.9}$$

where we have used orthonormality of the φ_k.

For many types of equations, the ODEs given by (2.9) may be determined analytically, in terms of the coordinates a_k. This is particularly useful computationally, as the inner product in (2.9) does not need to be computed at every timestep. For instance, if $X(u)$ is quadratic, given by

$$X(u) = L(u) + Q(u, u),$$

where $L : H \to H$ is linear and $Q : H \times H \to H$ is bilinear, the projected ODEs (2.9) become

$$\dot{a}_k(t) = \sum_i a_i(t)\langle L(\varphi_i), \varphi_k \rangle + \sum_{i,j} a_i(t)a_j(t)\langle Q(\varphi_i, \varphi_j), \varphi_k \rangle,$$

where the inner products are constants (independent of t) which may be determined before integrating the ODEs.

2.3. Stability and choice of inner product

It is well known that under some circumstances, Galerkin projections can perform poorly, and even produce unstable equilibrium points and limit cycles where the full system possesses stable equilibrium points and limit cycles. Possible reasons for this are illustrated with an elegant example in [5]. We show here that if an "energy-based" inner product is used (defined more precisely below), then Galerkin projection preserves the stability of an equilibrium point at the origin.

For simplicity, we restrict ourselves to the finite-dimensional case, where a dynamical system $\dot{x} = f(x)$ evolves on $\mathbb{R}^n$. Suppose the origin is a stable equilibrium point, and suppose that $V(x) = x^{\mathrm{T}}Qx$ is a quadratic Liapunov function for this system, in some neighborhood U of the origin. That is, Q is symmetric, positive definite, and $\dot{V}(x(t)) \leq 0$ along trajectories, for $x(t) \in U$, which implies

$$(f(x))^{\mathrm{T}}Qx + x^{\mathrm{T}}Qf(x) \leq 0 \quad \forall\, x \in U. \tag{2.10}$$

For instance, if the governing equation conserves energy or is dissipative, then we may choose $V(x)$ to be the total energy. Now, define an "energy-based" inner product on $\mathbb{R}^n$ by $\langle x, y \rangle = x^{\mathrm{T}}Qy$, and consider an orthogonal projection P onto some arbitrary subspace of $\mathbb{R}^n$. Recall that an orthogonal projection must be self-adjoint, and so

$$\langle x, Py \rangle = \langle Px, y \rangle \quad \forall\, x, y \in \mathbb{R}^n \Rightarrow QP = P^{\mathrm{T}}Q.$$

The Galerkin projection of the equations is $\dot{r} = Pf(r)$, where r is in the subspace (so $Pr = r$). It follows that $V(r) = r^{\mathrm{T}}Qr$ is a Liapunov function for the projected system as well, since

$$\dot{V}(r) = \dot{r}^{\mathrm{T}}Qr + r^{\mathrm{T}}Q\dot{r} = f(r)^{\mathrm{T}}P^{\mathrm{T}}Qr + r^{\mathrm{T}}QPf(r) = f(r)^{\mathrm{T}}Q(Pr) + (Pr)^{\mathrm{T}}Qf(r) \leq 0 \quad \forall\, r \in U$$

by (2.10), since $Pr = r$. Asymptotic stability may also be deduced if $\dot{V}(x)$ is strictly negative.

Thus, an "energy-based" inner product has the advantage that stability of an equilibrium point at the origin (e.g., stability of the linearization) is preserved by Galerkin projection. This is obviously not the case if a non-energy-based inner product is used. For instance, using the standard Euclidean inner product, orthogonal projection of the stable linear system:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 3 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

onto the x_1 subspace gives $\dot{x}_1 = x_1$, which is unstable. Note, however, that an energy-based inner product does not guarantee that stability of other equilibrium points will be preserved, nor that stability of limit cycles is preserved, as we shall see in Section 5.

3. Incompressible fluids

Before applying these methods to compressible flows, we first outline the well-established methods for incompressible flows, and highlight the differences in the next section.

The definition of incompressibility is that the velocity $\mathbf{u} = (u, v, w)$ is divergence-free, $\operatorname{div}\mathbf{u} = 0$. The motion of the fluid satisfies the Navier–Stokes equations [6], written in Cartesian coordinates as

$$\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} = -\nabla p + \nu\nabla^2\mathbf{u}, \tag{3.1}$$

where ν is the viscosity, p the pressure, and $\mathrm{D}/\mathrm{D}t = \partial/\partial t + \mathbf{u}\cdot\nabla$ is the material derivative. Velocities have been normalized by some velocity scale U, lengths by a length scale L, time by U/L, pressure by ρU^2 where ρ is the density, and viscosity by ρUL (thus, ν is the reciprocal of the Reynolds number). These equations may be written as

$$\dot{\mathbf{u}} = N(\mathbf{u}) - \nabla p, \tag{3.2}$$

where $N(\mathbf{u}) = -(\mathbf{u}\cdot\nabla)\mathbf{u} + \nu\nabla^2\mathbf{u}$.

For incompressible flows, we write the velocity $\mathbf{u}$ as an expansion in POD modes $\boldsymbol{\varphi}(\mathbf{x})$, defined on a spatial domain Ω in which the fluid evolves:

$$\mathbf{u}(\mathbf{x}, t) = \sum_{j=1}^{n} a_j(t)\boldsymbol{\varphi}_j(\mathbf{x}). \tag{3.3}$$

Our Hilbert space H is just the space of smooth (C^∞), divergence-free, vector-valued functions on Ω, with the standard inner product:

$$\langle \mathbf{u}, \mathbf{v}\rangle = \int_\Omega \mathbf{u}(\mathbf{x})\cdot\mathbf{v}(\mathbf{x})\,\mathrm{d}V. \tag{3.4}$$

(Additionally, we restrict the space H to contain only functions with finite norm.)

Inserting the expansion (3.3) into the Navier–Stokes equation (3.2), and taking an inner product with $\boldsymbol{\varphi}_k$ gives

$$\dot{a}_k = \langle N(\mathbf{u}), \boldsymbol{\varphi}_k\rangle - \langle \nabla p, \boldsymbol{\varphi}_k\rangle. \tag{3.5}$$

The pressure term on the right-hand side may be rewritten:

$$\langle \nabla p, \boldsymbol{\varphi}_k\rangle = \int_\Omega \boldsymbol{\varphi}_k\cdot\nabla p\,\mathrm{d}V = \int_\Omega \operatorname{div}(p\boldsymbol{\varphi}_k)\,\mathrm{d}V = \int_{\partial\Omega} p\boldsymbol{\varphi}_k\cdot\mathbf{n}\,\mathrm{d}S,$$

wherein the second equality we have used that div $\boldsymbol{\varphi}_k = 0$. Thus, this term depends only on the pressure on the boundary $\partial\Omega$. Furthermore, if velocity is zero along the boundary (for instance, at a wall, or in the farfield of an open flow), then $\boldsymbol{\varphi}_k = 0$ on $\partial\Omega$, and the pressure term vanishes altogether. If the boundary is not a wall, but rather an "artificial" boundary we impose (i.e., we consider only a limited portion of the whole flow), then the pressure term represents the influence of the rest of the flow on the domain we are considering, and must be specified as a boundary condition.

The technique for incompressible flows is well known. Many enhancements to the basic theory exist (see [3] for a thorough discussion), but the essential elements are the same as given here. The main feature that is normally included, but is not mentioned above, is that the velocity is usually decomposed into mean and fluctuating components ($\mathbf{u} = \bar{\mathbf{u}} + \mathbf{u}'$). If the mean $\bar{\mathbf{u}}$ is constant in time, the technique is the same as above, with the linear projection (3.3) replaced by an affine projection (as discussed in Section 4.3 for the compressible equations). The mean may also be slowly varying in time, in which case it is usually modeled in terms of the fluctuations $\mathbf{u}'$, which give rise to Reynolds stresses, as discussed in [2]. This is an important distinction, and gives rise to cubic equations, while Eq. (3.5) are only quadratic. Another enhancement to the basic theory is the modeling of the energy transfer to the higher modes. In this paper, we take the mean flow to be constant in time, without attempting to model the Reynolds stress contributions, and we neglect the energy transfer to the higher modes. These extensions could presumably be added to the compressible theory as well.

4. Compressible fluids

The main result of this paper is the application of the above techniques to compressible flows. The distinction is a fundamental one. On a superficial level, the constraint div $\mathbf{u} = 0$ no longer holds. On a more fundamental level, in a compressible flow, the thermodynamic variables are *dynamically* important, and must be included in the actual dynamics, not merely as a constraint. For incompressible flows, as discussed above, the pressure drops out completely, or appears only as an imposed boundary condition, and the only dynamical variable is the velocity. For compressible flows, this is not the case, and evolution equations must be given for one or more thermodynamic variables (e.g., ρ, p, entropy s, enthalpy h), as well as the velocity. This introduces further questions of whether to treat these variables completely separately from the velocity, or together as a single vector-valued "configuration" variable (e.g., $\mathbf{q}(x) = (\rho, u, v, w, p)(x)$).

Both these alternatives are considered in [7,8], in developing reduced-order models for the compressible flow past a rectangular cavity. A scalar-valued method was used, in which the flow variables are considered separately, as coupled dynamical systems, with separate POD modes for each variable. This method was compared to a vector-valued method, in which the flow variables are taken together as a single configuration vector, with a single set of vector-valued POD modes for the configuration space. The scalar-valued method appeared to offer no advantages, requiring higher-order models, and furthermore demonstrating poor long-time behavior for the particular example considered. Here, we consider only the vector-valued method.

To compute POD modes, we first need to define an appropriate inner product on our configuration space—the standard inner product may not be a sensible choice. For instance, if we use as flow variables $\mathbf{q} = (\rho, u, v, w, p)$, defined on the fluid domain Ω, the standard inner product is

$$\langle \mathbf{q}_1, \mathbf{q}_2 \rangle = \int_\Omega (\rho_1\rho_2 + u_1u_2 + v_1v_2 + w_1w_2 + p_1p_2)\,\mathrm{d}V, \tag{4.1}$$

which does not make dimensional sense, since one cannot add a velocity and a pressure (one may, of course, nondimensionalize the variables, but then the choice of nondimensionalization becomes critical). We introduce an inner product suitable for compressible flows in Section 4.2.

4.1. Governing equations

The fully compressible Navier–Stokes equations (see Appendix A) are significantly more complicated than the incompressible equations, so we make some approximations to obtain a simpler set of equations. Using assumptions valid for cold flows at moderate Mach number, and with a careful choice of thermodynamic variables, we obtain *quadratic* equations, which makes Galerkin projections particularly simple.

We consider a *cold flow* ($T_{\text{wall}} = T_\infty$) and note that if the Mach number is not too high, density gradients will remain small and will be dominated by pressure changes [9]. This is consistent with the neglect of the viscous dissipation and heat conduction in the energy equation, and thus we treat the flow as *isentropic*. However, we retain the viscous diffusion in the momentum equation, but in this term we assume dynamic and kinematic viscosities are constant, again under the approximation that temperature gradients are small. Under these assumptions, the equations of motion become

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} + \rho \operatorname{div} \mathbf{u} = 0, \qquad \frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} + \frac{1}{\rho}\nabla p = \nu \nabla^2 \mathbf{u}, \qquad \mathrm{d}s = 0, \tag{4.2}$$

where $\mathbf{u}$ is the velocity in three dimensions, ρ the density, p the pressure, s the entropy, and $\nu = \mu/\bar{\rho}$ a constant. We still have more variables than equations, so to close the system we require an equation of state. With an equation of state, we may write the equations in terms of a single thermodynamic variable, since the entropy is constant.

Using the ideal gas relation $p = \rho RT$, and the Gibbs equation $\mathrm{d}h = T\,\mathrm{d}s + \mathrm{d}p/\rho$ with $\mathrm{d}s = 0$, Eq. (4.2) may be written with the enthalpy h as the only thermodynamic variable:

$$\frac{\mathrm{D}h}{\mathrm{D}t} + (\gamma - 1)h \operatorname{div} \mathbf{u} = 0, \qquad \frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} + \nabla h = \nu \nabla^2 \mathbf{u}. \tag{4.3}$$

This procedure will work for any thermodynamic variable, but the enthalpy is particularly convenient, because Eq. (4.3) is *quadratic*. If we had used the density, for instance, the momentum equation would become

$$\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} + \rho^{\gamma-2}\nabla\rho = \nu \nabla^2 \mathbf{u}.$$

For air, $\gamma = 1.4$, so the exponent $\gamma - 2$ in the above equation is fractional, which would cause difficulty when we perform Galerkin projections, described in Section 4.3.

Another choice of thermodynamic variable which yields quadratic equations is the local sound speed a. Using the ideal gas relation $a^2 = (\gamma - 1)h$, Eq. (4.3) become

$$\frac{\mathrm{D}a}{\mathrm{D}t} + \frac{\gamma - 1}{2} a \operatorname{div} \mathbf{u} = 0, \qquad \frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} + \frac{2}{\gamma - 1} a \nabla a = \nu \nabla^2 \mathbf{u}. \tag{4.4}$$

Eqs. (4.3) and (4.4) are equivalent, and we refer to them (in either form) as the *isentropic Navier–Stokes equations*. We focus on Eq. (4.4) in terms of sound speed, because it is straightforward to obtain a physically based inner product for these variables, as we see in the next section.

4.2. Inner products for compressible flow

In order to obtain vector-valued POD modes, and perform Galerkin projections as in the previous section, we must first define an inner product. Here we introduce a family of inner products useful for compressible flow problems. As mentioned earlier, the usual inner product (e.g., (4.1)) may not make dimensional sense when both thermodynamic and kinematic variables are included. Of course, one could simply nondimensionalize the variables, but then the sense in which projections are "optimal" is rather arbitrary, and depends on the nondimensionalization

(essentially the choice of "coordinates"), and hence only indirectly on the underlying physics. Some choices of nondimensionalization may be more relevant than others; for instance, [10] suggests a nondimensionalization based on the variance of the density and velocity fluctuations in a dataset used for POD.

The inner product presented here has a direct physical interpretation, in that the "energy" defined by the induced norm is a meaningful physical quantity, either the integrated stagnation energy or the integrated stagnation enthalpy. There are several reasons for using an inner product based on the flow physics. First, as shown in Section 2.3, if an energy-based inner product is used, then stability of an equilibrium point at the origin will be preserved by Galerkin projection. Second, since the underlying phenomena and equations of motion have a physical interpretation, it is natural to introduce a physically motivated inner product as well. However, we point out that in many situations an energy-based inner product may not be appropriate, for instance, in jet noise problems where acoustic waves of interest contain a very small fraction of the energy. Other choices of inner products for these flows are considered in [11,12].

For incompressible flows, the norm induced by the standard inner product is directly related to the kinetic energy, with $T = \|\mathbf{u}\|^2/2$. For compressible flows, both thermodynamic and kinematic variables contribute to the total energy. For instance, the stagnation (or total) enthalpy of the flow is given by

$$h_0 = h + \tfrac{1}{2}(u^2 + v^2),$$

where h is the static enthalpy. Analogously, the stagnation energy is $e = E + (1/2)(u^2 + v^2)$, where $E = h/\gamma$ is the internal energy per unit mass. (For the remainder of the paper, we restrict ourselves to two spatial dimensions, but the extension to three dimensions is obvious.) Motivated by these physical quantities based on energy, we look for inner products which have induced norms of the form:

$$\frac{1}{2}\|\mathbf{q}\|_\alpha^2 = \int_\Omega \left(\alpha h + \frac{1}{2}(u^2 + v^2)\right) \mathrm{d}V, \tag{4.5}$$

where $\mathbf{q}$ is the vector of flow variables, and $\alpha > 0$ a constant. Note that the right-hand side of (4.5) is not quadratic, because h appears linearly, but this is easily remedied by transforming to the flow variables $\mathbf{q} = (u, v, a)$, since $a^2 = (\gamma - 1)h$. Thus, we define a family of inner products:

$$\langle \mathbf{q}_1, \mathbf{q}_2 \rangle_\alpha = \int_\Omega \left(u_1 u_2 + v_1 v_2 + \frac{2\alpha}{\gamma - 1} a_1 a_2\right) \mathrm{d}V, \tag{4.6}$$

which has induced norm given by (4.5).

Choosing $\alpha = 1$ corresponds to using the integral of the stagnation enthalpy as the norm, and taking $\alpha = 1/\gamma$ corresponds to using the stagnation energy as the norm. However, note that neither the integral of the stagnation enthalpy or the stagnation energy is actually a conserved quantity. The conserved quantity is the *total energy*, given by

$$\int_\Omega \left(\rho E + \frac{1}{2}\rho(u^2 + v^2)\right) \mathrm{d}V. \tag{4.7}$$

(Of course, this is only conserved if there is no energy flux through the boundary of Ω, as one obtains when, for example, the velocity vanishes on $\partial\Omega$.) Though this norm would perhaps be the most natural from an energy point of view, it is not obvious how to choose configuration variables for which the equations of motion are tractable. The choice $\mathbf{q} = (\sqrt{\rho E}, \sqrt{\rho}u, \sqrt{\rho}v)$ may be used to compute POD modes, but the physical significance of these variables is not clear, and the equations of motion in these variables will of course be quite involved.

4.3. Projected equations

Here, we compute the coefficients in the Galerkin projection of the isentropic Navier–Stokes equation (4.4), which may be written in two dimensions as

$$u_t = -uv_x - vu_y - \frac{2}{\gamma-1}aa_x + \nu(u_{xx}+u_{yy}), \qquad v_t = -uv_x - vv_y - \frac{2}{\gamma-1}aa_y + \nu(v_{xx}+v_{yy}),$$
$$a_t = -ua_x - va_y - \frac{\gamma-1}{2}a(u_x+v_y). \tag{4.8}$$

Writing $\mathbf{q} = (u, v, a)$, the equations take the form:

$$\dot{\mathbf{q}} = \nu L(\mathbf{q}) + Q(\mathbf{q}, \mathbf{q}), \tag{4.9}$$

where

$$L(\mathbf{q}) = \begin{pmatrix} u_{xx}+u_{yy} \\ v_{xx}+v_{yy} \\ 0 \end{pmatrix}, \tag{4.10}$$

$$Q(\mathbf{q}^1, \mathbf{q}^2) = -\begin{pmatrix} u^1u_x^2 + v^1u_y^2 + \frac{2}{\gamma-1}a^1a_x^2 \\ u^1v_x^2 + v^1v_y^2 + \frac{2}{\gamma-1}a^1a_y^2 \\ u^1a_x^2 + v^1a_y^2 + \frac{\gamma-1}{2}a^1(u_x^2+v_y^2) \end{pmatrix}. \tag{4.11}$$

Next, we expand $\mathbf{q}$ in terms of POD modes (or any other orthogonal basis functions), as

$$\mathbf{q}(\mathbf{x}, t) = \bar{\mathbf{q}}(\mathbf{x}) + \sum_{j=1}^{n} a_j(t)\boldsymbol{\varphi}_j(\mathbf{x}), \tag{4.12}$$

where $\bar{\mathbf{q}}$ is fixed, and typically taken to be the mean of all the snapshots used for POD, though other choices are possible (e.g., a nearby steady solution of Navier–Stokes). For the affine projection above, the mean $\bar{\mathbf{q}}$ must be subtracted from the original data before the POD modes are computed. The resulting Galerkin equations are given by

$$\dot{a}_k = \nu b_k^1 + b_k^2 + \sum_{i=1}^{n}(\nu L_{ik}^1 + L_{ik}^2)a_i + \sum_{i,j=1}^{n} Q_{ijk}a_ia_j, \tag{4.13}$$

where the coefficients:

$$b_k^1 = \langle L(\bar{\mathbf{q}}), \boldsymbol{\varphi}_k\rangle, \qquad b_k^2 = \langle Q(\bar{\mathbf{q}}, \bar{\mathbf{q}}), \boldsymbol{\varphi}_k\rangle, \qquad L_{ik}^1 = \langle L(\boldsymbol{\varphi}_i), \boldsymbol{\varphi}_k\rangle,$$
$$L_{ik}^2 = \langle Q(\bar{\mathbf{q}}, \boldsymbol{\varphi}_i) + Q(\boldsymbol{\varphi}_i, \bar{\mathbf{q}}), \boldsymbol{\varphi}_k\rangle, \qquad Q_{ijk} = \langle Q(\boldsymbol{\varphi}_i, \boldsymbol{\varphi}_j), \boldsymbol{\varphi}_k\rangle$$

are constants which may be computed before solving the reduced system. Note that the affine terms in (4.13) vanish if $\bar{\mathbf{q}}$ is a steady solution of Navier–Stokes (so that $\nu L(\bar{\mathbf{q}}) + Q(\bar{\mathbf{q}}, \bar{\mathbf{q}}) = 0$).

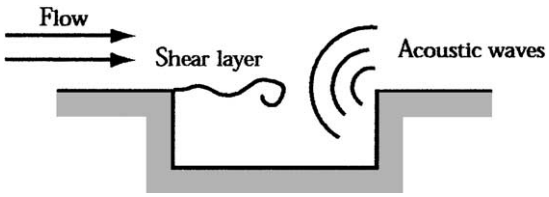

Fig. 1. Basic configuration of a cavity flow.

5. Example: cavity flow oscillations

To demonstrate the effectiveness of the approach described above, we show some results of the method applied to the flow past a rectangular cavity. The geometry of the flow is indicated in Fig. 1. Cavity flows often exhibit self-sustained oscillations, usually explained by the following mechanism: small disturbances in the shear layer at the leading edge are amplified exponentially, and produce acoustic waves when they impinge on the downstream corner; these acoustic waves propagate upstream, and excite further disturbances in the shear layer, creating a natural feedback mechanism which often leads to finite-amplitude oscillations.

We have performed an extensive set of simulations of the two-dimensional compressible flow, presented in detail in [13]. The fully compressible Navier–Stokes equations were solved using sixth-order compact finite differences in space and fourth-order Runge–Kutta in time, resulting in a scheme with very low dispersion and dissipation, necessary to resolve acoustic waves. The flow was two-dimensional, and hence laminar, and typical runs used roughly 500,000 gridpoints. The parameters for the run presented are $L/D = 2$, $L/\theta = 58$, $Re_\theta = 69$, $M = 0.6$, where L and D are the cavity length and depth, respectively, θ the momentum thickness of the boundary layer at the leading edge of the cavity, $M = U_\infty/a_\infty$ the freestream Mach number, and $Re_\theta = \rho_\infty U_\infty \theta/\mu$ the Reynolds number based on θ.

We computed POD modes from the simulation data, using 51 snapshots taken after transients had decayed. The initial condition for the simulation was a Blasius boundary layer spanning the cavity at $t = 0$, and the snapshots were taken at evenly spaced times between $tU/L = 67.8$ and 77.5. The mean of the snapshots was subtracted from the ensemble before computing POD modes. The inner product used in computing POD modes is given by Eq. (4.6), with $\alpha = 1$, with the integral evaluated over the portion of the computational domain with $-1 \leq x/D \leq 3$, $-1 \leq y/D \leq 2$. The computational domain extends well beyond this, but only the nearfield was used for computing POD modes. Other values of the parameter α were examined in [7], and the results were qualitatively identical to those with $\alpha = 1$.

The first four POD modes are shown in Fig. 2, and together they capture over 98.7% of the energy in the fluctuations (where the energy is in the sense of the induced norm). The vorticity shows the structures in the shear layer, and the dilatation reveals the presence of acoustic waves, radiated from the downstream corner. Note that the modes occur in pairs, shifted in phase by 90°, as is characteristic of flows with convecting structures.

We then project the isentropic Navier–Stokes equations (4.4) onto these POD modes, as described in Section 4.3, to obtain reduced-order models. Fig. 3 shows the time traces from several reduced-order models, retaining between 2 and 20 states, compared with projections of the snapshots from the full simulation. The initial condition for the reduced-order models was obtained by projecting a snapshot from the full simulation onto the POD modes. For the short times shown in Fig. 3, the agreement is excellent: the frequencies and amplitudes of the oscillations are captured very well by all the models.

Fig. 4 illustrates the long-time behavior of the models. All the reduced-order models possess a stable limit cycle except for the 2-mode model, which has a (non-physical) stable fixed point at $(-1.47, -3.26)$. For the higher-order

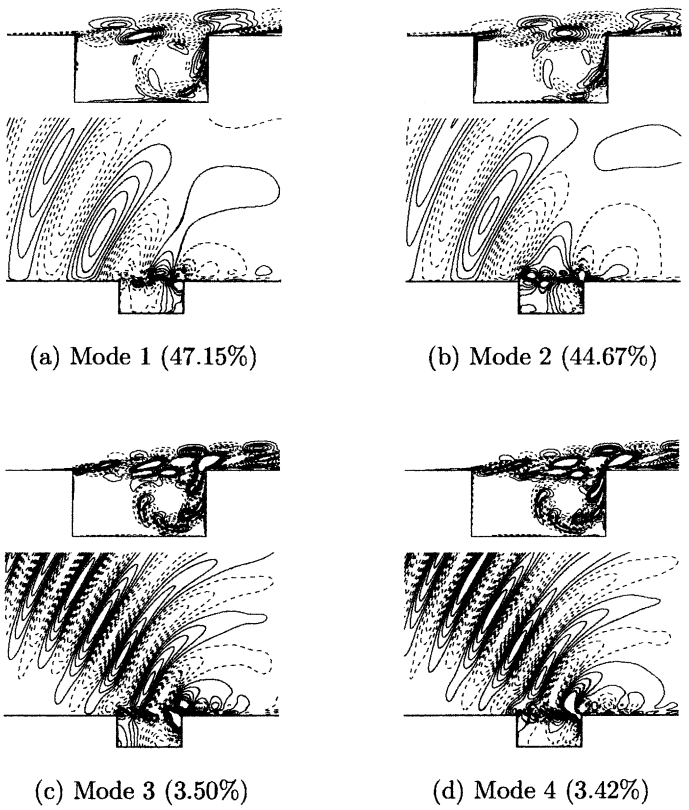

Fig. 2. POD modes from cavity flow simulation: vorticity (top) and dilatation (bottom), and percent energy captured.

models, the final amplitude of the limit cycle agrees better with the data from the full simulation as the number of modes increases from 4, to 6, to 11 modes, but when more than 11 modes are retained, the behavior becomes more complex and non-physical.

6. Discussion

We have presented a method for applying POD and Galerkin projection to compressible flows. For Galerkin projection, we use an approximate, isentropic version of the Navier–Stokes equations that is valid for cold flows (small temperature gradients) at moderate Mach number, and we have chosen appropriate flow variables so that only quadratic terms appear in the equations. By contrast, the full compressible Navier–Stokes equations contain cubic terms, and furthermore, the time derivatives may not be solved for explicitly, which complicates Galerkin projections.

An inner product is necessary for both POD and Galerkin projection, and we have introduced an inner product that is physically meaningful, based on stagnation enthalpy or stagnation energy. This inner product thus has the property that the stability of an equilibrium point at the origin will be preserved under Galerkin projection, and furthermore it yields projected equations that are easy to implement.

We have applied the method to a flow past a rectangular cavity, and have found that the POD modes identify both vortical structures in the shear layer and acoustic waves in the farfield. The method works well for the flow studied,

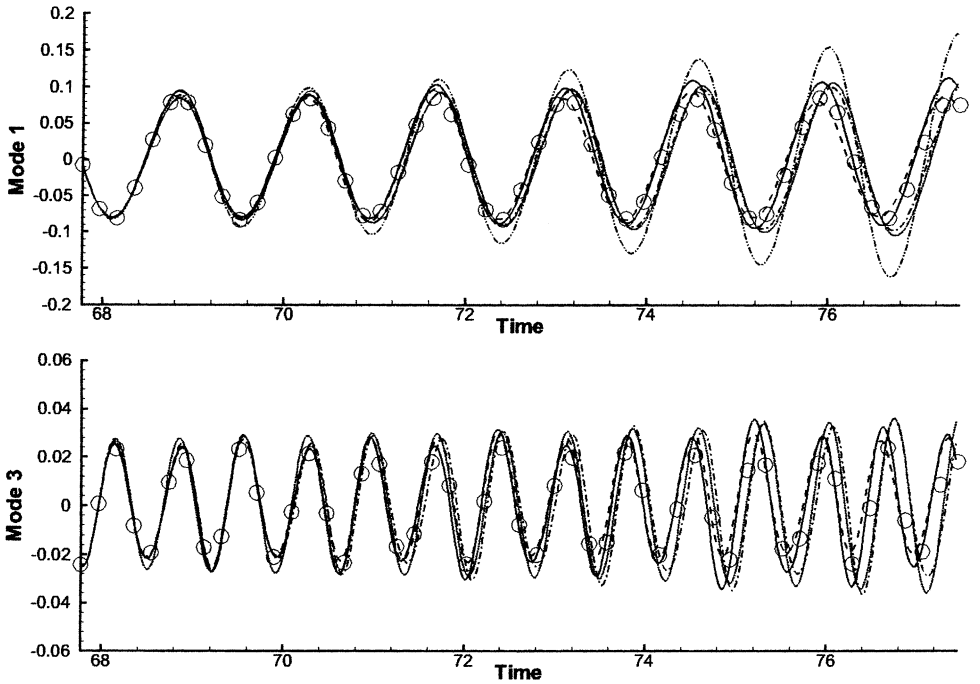

Fig. 3. Time traces from Galerkin projections: coefficients of POD modes 1 and 3, from reduced-order simulations with 2 modes (–··–), 4 (· · ·), 6 (–·–), 11 (- - -), and 20 modes (—), and projection of simulation data (○).

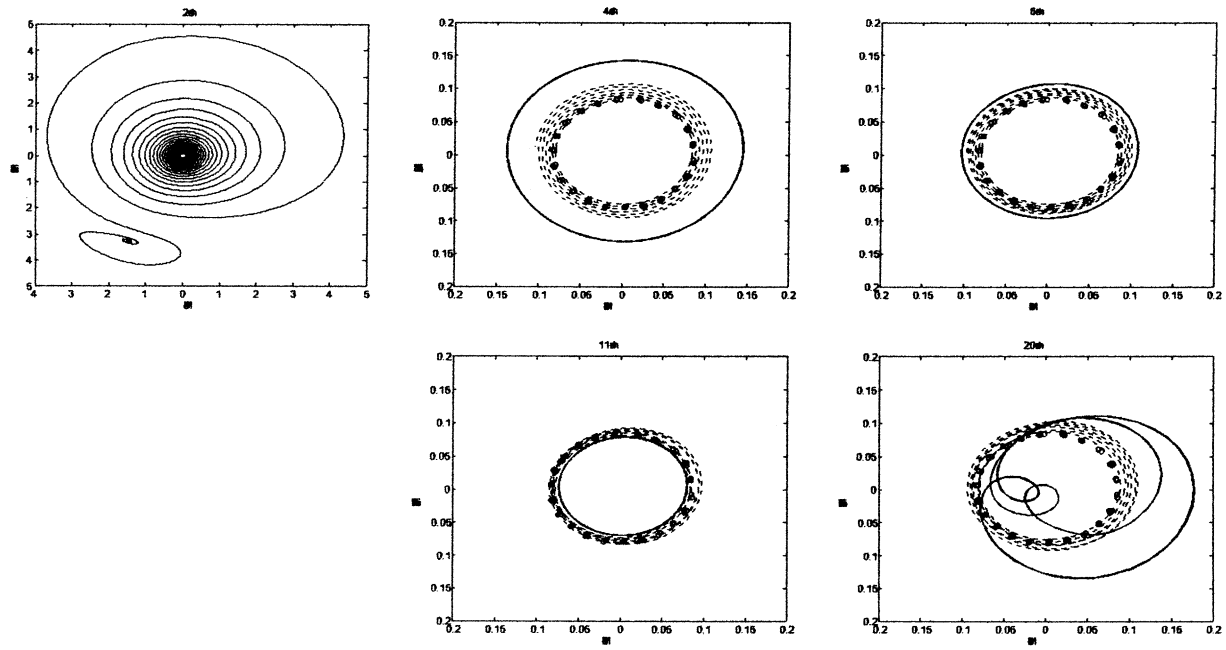

Fig. 4. Trajectories of Galerkin models, with 2–20 modes, for the first 10 time units (- - -), and for the final limit cycle after transients have decayed (—), and data from full simulation (○).

at low Reynolds number and moderate Mach number, capturing over 98% of the energy in the fluctuations with the first four POD modes. Galerkin projections of strikingly low dimension are obtained, and give results which closely match those from the full simulations, even when as few as four modes are retained.

Acknowledgements

This work was supported by AFOSR under grants F49620-98-1-0095 and F49620-03-1-0081, with technical monitors Dr. Thomas Beutner and Dr. John Schmisseur.

Appendix A. Galerkin projection of fully compressible Navier–Stokes

A.1. Equations of motion

The fully compressible Navier–Stokes equations may be written as follows:

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} + \operatorname{div}\mathbf{u} = 0, \qquad \rho\frac{\mathrm{D}u_i}{\mathrm{D}t} = -\frac{\partial p}{\partial x_i} + \frac{1}{Re}\frac{\partial}{\partial x_j}\left(2S_{ij} - \frac{2}{3}\delta_{ij}\operatorname{div}\mathbf{u}\right),$$
$$\rho\frac{\mathrm{D}T}{\mathrm{D}t} + (\gamma-1)\rho T\operatorname{div}\mathbf{u} = \frac{\gamma}{Re}\left(2S_{ij}S_{ij} - \frac{2}{3}(\operatorname{div}\mathbf{u})^2\right) + \frac{\gamma}{Re\,Pr}\nabla^2 T,$$

where $p = (\gamma-1)\rho T$, and $S_{ij} = (1/2)(\partial u_i/\partial x_j + \partial u_j/\partial x_i)$ is the rate-of-strain tensor (the symmetric part of the velocity gradient). Here the usual compressible nondimensionalization has been used, using a length scale L, the ambient sound speed a_∞, and ambient density ρ_∞. Temperature is nondimensionalized by a_∞^2/c_p, and pressure by $\rho_\infty a_\infty^2$. These equations are of the form:

$$\frac{\partial\rho}{\partial t} = R(\mathbf{q}), \qquad \rho\frac{\partial u_i}{\partial t} = U_i(\mathbf{q}), \qquad \rho\frac{\partial T}{\partial t} = \Theta(\mathbf{q}),$$

where $\mathbf{q} = (\rho, u_1, u_2, u_3, T)$, and R, U_i, and Θ are the nonlinear differential operators given by

$$R(\mathbf{q}) = -(\mathbf{u}\cdot\nabla)\rho - \operatorname{div}\mathbf{u}, \qquad U_i(\mathbf{q}) = -\rho(\mathbf{u}\cdot\nabla)u_i - \frac{\partial p}{\partial x_i} + \frac{1}{Re}\frac{\partial}{\partial x_j}\left(2S_{ij} - \frac{2}{3}\delta_{ij}\operatorname{div}\mathbf{u}\right),$$
$$\Theta(\mathbf{q}) = -(\gamma-1)\rho T\operatorname{div}\mathbf{u} + \frac{\gamma}{Re}\left(2S_{ij}S_{ij} - \frac{2}{3}(\operatorname{div}\mathbf{u})^2\right) + \frac{\gamma}{Re\,Pr}\nabla^2 T.$$

These equations may be written more concisely as

$$A(\mathbf{q})\dot{\mathbf{q}} = \mathbf{f}(\mathbf{q}), \tag{A.1}$$

where $A(\mathbf{q}) = \operatorname{diag}(1, \rho, \rho)$ and $\mathbf{f}(\mathbf{q}) = (R(\mathbf{q}), U_1(\mathbf{q}), \Theta(\mathbf{q}))$. Here, only one spatial dimension is used, but the extension to higher spatial dimensions is trivial. Note that A is affine in $\mathbf{q}$, and may be written

$$A(\mathbf{q}) = B + L(\mathbf{q}) = \operatorname{diag}(1, 0, 0) + \operatorname{diag}(0, \rho, \rho).$$

Furthermore, $\mathbf{f}$ is cubic and may be written

$$\mathbf{f}(\mathbf{q}) = \mathbf{f}_1(\mathbf{q}) + \mathbf{f}_2(\mathbf{q}, \mathbf{q}) + \mathbf{f}_3(\mathbf{q}, \mathbf{q}, \mathbf{q}),$$

where $\mathbf{f}_{1,2,3}$ are multilinear functions (linear in each argument).

A.2. Galerkin projection

Now, let $\{\boldsymbol{\varphi}_k\}$ be a basis for the function space containing $\mathbf{q}$, and express $\mathbf{q}$ in terms of this basis:

$$\mathbf{q}(t) = \sum_k a_k(t)\boldsymbol{\varphi}_k. \tag{A.2}$$

Inserting this expansion into the governing equation (A.1), we have

$$\left[B + L\left(\sum_l a_l\boldsymbol{\varphi}_l\right)\right]\sum_k \dot{a}_k\boldsymbol{\varphi}_k = \mathbf{f}(\mathbf{q}).$$

Taking an inner product with $\boldsymbol{\varphi}_j$ then gives

$$\sum_k \dot{a}_k\left(\langle\boldsymbol{\varphi}_j, B\boldsymbol{\varphi}_k\rangle + \sum_l a_l\langle\boldsymbol{\varphi}_j, L(\boldsymbol{\varphi}_l)\boldsymbol{\varphi}_k\rangle\right) = \langle\boldsymbol{\varphi}_j, \mathbf{f}(\mathbf{q})\rangle. \tag{A.3}$$

This equation may be solved for $\dot{a}_k$ by inverting the $n \times n$ matrix on the left-hand side (n is the number of basis elements, or modes). In matrix form, this equation is

$$M(\mathbf{a})\dot{\mathbf{a}} = \tilde{\mathbf{f}}(\mathbf{a}), \tag{A.4}$$

where $\mathbf{a} = (a_1, \ldots, a_n)$, and

$$M(\mathbf{a})_{jk} = \langle\boldsymbol{\varphi}_j, B\boldsymbol{\varphi}_k\rangle + \sum_l a_l\langle\boldsymbol{\varphi}_j, L(\boldsymbol{\varphi}_l)\boldsymbol{\varphi}_k\rangle,$$

$$\tilde{f}_j(\mathbf{a}) = \langle\boldsymbol{\varphi}_j, \mathbf{f}(\mathbf{q})\rangle = \sum_l a_l\langle\boldsymbol{\varphi}_j, \mathbf{f}_1(\boldsymbol{\varphi}_l)\rangle + \sum_{l,m} a_l a_m\langle\boldsymbol{\varphi}_j, \mathbf{f}_2(\boldsymbol{\varphi}_l, \boldsymbol{\varphi}_m)\rangle + \sum_{l,m,n} a_l a_m a_n\langle\boldsymbol{\varphi}_j, \mathbf{f}_3(\boldsymbol{\varphi}_l, \boldsymbol{\varphi}_m, \boldsymbol{\varphi}_n)\rangle.$$

Note that all the inner products may be computed beforehand, as they depend only on the basis functions $\boldsymbol{\varphi}_j$, not on the coefficients a_j. In the constant-density case (with $\rho = 1$), $B = I$ and $L = 0$, so if the basis functions are orthonormal ($\langle\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_k\rangle = \delta_{jk}$) then the matrix M is just the identity, and Eq. (A.4) reduce to $\dot{\mathbf{a}} = \tilde{\mathbf{f}}(\mathbf{a})$.

Though in principle there is no difficulty in solving the full equation (A.4) directly, one first needs to choose a suitable inner product. The approximate, isentropic equations presented in Section 4.1 have a natural choice of inner product, and are simpler to implement than (A.4), since they are quadratic, and do not involve inversion of a "mass matrix" M.

References

[1] L. Sirovich, Turbulence and the dynamics of coherent structures, Parts I–III, Quart. Appl. Math. XLV (3) (1987) 561–590.
[2] N. Aubry, P. Holmes, J.L. Lumley, E. Stone, The dynamics of coherent structures in the wall region of a turbulent boundary layer, J. Fluid Mech. 192 (1988) 115–173.
[3] P. Holmes, J.L. Lumley, G. Berkooz, Turbulence, Coherent Structures, Dynamical Systems and Symmetry, Cambridge University Press, Cambridge, 1996.
[4] R. Abraham, J.E. Marsden, T.S. Ratiu, Manifolds, Tensor Analysis, and Applications, 2nd ed., Applied Mathematical Sciences, No. 75, Springer-Verlag, 1988.
[5] D. Rempfer, On low-dimensional Galerkin models for fluid flow, Theor. Comput. Fluid Dyn. 14 (2) (2000) 75–88.
[6] G.K. Batchelor, An Introduction to Fluid Dynamics, Cambridge University Press, Cambridge, 1967.
[7] C.W. Rowley, Modeling, simulation, and control of cavity flow oscillations, Ph.D. Thesis, California Institute of Technology, 2002.
[8] C.W. Rowley, T. Colonius, R.M. Murray, Dynamical models for control of cavity oscillations, AIAA Paper 2001-2126, May 2001.
[9] G.P. Zank, W.H. Matthaeus, The equations of nearly incompressible fluids. I. Hydrodynamics, turbulence, and waves, Phys. Fluids 3 (1) (1991) 69–82.

[10] J.L. Lumley, A. Poje, Low-dimensional models for flows with density fluctuations, Phys. Fluids 9 (7).
[11] T. Colonius, J.B. Freund, Reconstruction of large-scale structures and acoustic radiation from a turbulent $M = 0.9$ jet using proper orthogonal decomposition, in: I.P. Castro, P.E. Hancock, T.G. Thomas (Eds.), Advances in Turbulence IX, CIMNE, Barcelona, 2002.
[12] J.B. Freund, T. Colonius, POD analysis of sound generation by a turbulent jet, AIAA Paper 2002-0072, 2002.
[13] C.W. Rowley, T. Colonius, A.J. Basu, On self-sustained oscillations in two-dimensional compressible flow over rectangular cavities, J. Fluid Mech. 455 (2002) 315–346.