

Elements of resolvent methods in fluid mechanics: notes for an introductory short course v0.3

A S Sharma
a.sharma@soton.ac.uk
University of Southampton

July 26, 2019

1 Introduction

This is a collection of notes for part of a short course on modal methods in fluid mechanics held at DAMTP, University of Cambridge, in the summer of 2019. These notes in particular are meant to introduce the reader to resolvent analysis as it is currently used in fluid mechanics. Most of the papers on the topic assume a level of knowledge a bit beyond that of the average beginning PhD student, so there is a need for some introductory material to get new students up to speed quickly. These notes are a step towards providing such material and will serve as a base from which to explore the literature on the topic. The presentation assumes a good working knowledge of Fourier transforms and linear algebra, some familiarity with the incompressible Navier-Stokes equations, and not much else. Some experience with state space systems from an introductory course in control is beneficial. In most cases, rigour and technical detail have been elided in order not to obscure the central point. Inevitably, there will be mistakes in the notes and I would be grateful to be informed of these by email.

The method of analysis described in what follows arose from a desire to have a systematic and well-founded way to form ‘quick and dirty’ approximations to turbulent Navier-Stokes flows from the equations themselves (that is, as far as possible without recourse to simulation or experimental data). It was hoped that such approximations would successively approach the original equations as the detail of the approximation was increased. Fast and simple calculations would then enable the kind of parametric control studies that are expensive with direct numerical simulation.

This kind of approach was inspired by the successful model reduction methods of modern linear control theory, such as balanced truncation. Unfortunately, the existing methods of the time were designed for linear systems, or nonlinear systems that could sensibly be linearised around an operating point. Although many researchers had long practised looking at linear operators formed around the mean flow, it was not then clear to me what it was that was actually being calculated; the classical linearisation theorem taught to undergraduates explains the correspondence between a nonlinear system and its locally valid linearisation around an equilibrium. In contrast, turbulent flows are far from equilibrium, the turbulent mean is not an equilibrium point in phase space, and the turbulent fluctuations are large.

This dissatisfaction ultimately resulted in the present analysis. If it makes sense to speak of lineage in this context, one may draw a line back through the pseudospectra insights of Trefethen and coworkers [1], and the laminar resolvent based work arising from the control theory community [2]. Inevitably, this view and the presentation that follows is my own individual perspective.

These notes begin with an introduction to the singular value decomposition and its operator counterpart, the Schmidt decomposition. A general formulation of the resolvent decomposition is then introduced. A brief discussion of the interpretation as a nonlinear feedback loop is given. The methodology is then applied to the Navier-Stokes equations.

2 The singular value decomposition

The singular value decomposition (SVD) is a particular matrix factorisation that has very useful properties. It is widely used in data and model reduction because it solves the problem of finding the optimal approximation of a linear operator. Since we will be using it extensively, we now review some of its most important properties. In this section, vectors will be represented by lowercase letters, matrices by uppercase, and the conjugate transpose of A by A^* .

Lemma 2.1. Let M be a complex $m \times n$ matrix. The decomposition

$$M = U\Sigma V^* \tag{1}$$

always exists, where U is an $m \times m$ complex matrix, V is an $n \times n$ complex matrix, Σ is a $m \times n$ real and diagonal matrix with elements $\Sigma_{ii} = \sigma_i$ and $\sigma_1 \geq \sigma_2 \geq \dots$. The σ_i are called the singular values and (1) is called the singular value decomposition of M . Matrices U and V are unitary, $UU^* = U^*U = I_m$ and $VV^* = V^*V = I_n$.

$$\underbrace{\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}}_M = \underbrace{\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}}_U \underbrace{\begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & & \vdots \\ & & & \sigma_3 & & \\ & & & & & \vdots \end{bmatrix}}_\Sigma \underbrace{\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}}_{V^*}$$

Figure 2.1: The structure of the singular value decomposition with $n > m$. The linears in Σ and V^* represent the reduced SVD (see Section 2.2)

From the singular value decomposition, we can make the following observations. Since U and V are unitary, the rank of M is equal to the number of nonzero singular values. Notice that the inverse of a unitary matrix is its conjugate transpose. The decomposition is unique up to a constant complex multiplicative factor on each basis and up to the ordering of the singular values. That is, if $U\Sigma V^*$ is a singular value decomposition, so is $(e^{i\theta}U)\Sigma(V^*e^{-i\theta})$. The columns of V and U that span the space corresponding to any exactly repeating singular values may be combined arbitrarily. The structure of the matrix decomposition is illustrated in figure 2.1.

2.1 The maximum gain problem and its relationship with norms

It is helpful to think of M as an operator mapping a complex vector in the domain of M to another in the range of M . The columns of V , v_i , provide a basis which spans the domain. The singular value decomposition of M can be written in terms of the vectors of U and V ,

$$M = U\Sigma V^* = \sum_{i=1}^m \sigma_i u_i v_i^*. \quad (2)$$

Since V is unitary, $v_i^* v_j = \delta_{ij}$, so applying M to v_j gives

$$M v_j = \sum_{i=1}^m \sigma_i u_i v_i^* v_j = \sigma_j u_j. \quad (3)$$

Since V provides a basis for the domain of M , any vector a in the domain of M can itself be expressed in terms of a weighted sum of columns of V . That is, expressing a as

$$a = \sum_{i=1}^n v_i c_i$$

gives

$$\begin{aligned} Ma &= \sum_{i=1}^m u_i \sigma_i v_i^* a \\ &= \sum_{i=1}^m u_i \sigma_i c_i. \end{aligned}$$

We may then pose the question, what is the maximum amplitude of ‘output’ for a given ‘input’ amplitude? This is achieved with the input parallel to v_1 , with a gain of σ_1 . So,

$$\sigma_1 = \max_{a \neq 0} \frac{\|Ma\|}{\|a\|}$$

is achieved with $a/\|a\| = v_1$. Any other choice of a that is not parallel to v_1 would achieve an inferior gain. This is illustrated in figure 2.2 for a of unit length. M maps a circle (ball) of unit radius to an ellipse (hyperellipse). The singular values are the major and minor axes of the ellipse.

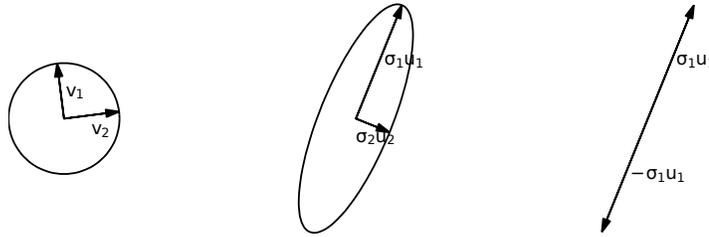


Figure 2.2: Mapping of the unit circle ($\|a\| = 1$, left) to an ellipse (Ma , centre) and mapping of Mv_1 to $\sigma_1 u_1$ (right). If we imagine the locus of points of a with unit length being drawn on a rubber sheet, the effect on M is to rotate and stretch the sheet. The amount of stretching in each direction is given by each singular value, and the directions by the singular vectors.

2.2 The low-rank approximation of matrices

For a non-square or rank-deficient square matrix, some of the singular values will be zero. In this case, the reduced SVD can be defined where the columns of U or V relating to the zero singular values, and the corresponding entries of Σ , can be truncated with the decomposition remaining exact. In this case, though, U (or V) will not be unitary because the columns associated with the null space of M will have been truncated. This is illustrated in Figure 2.1, where the truncated columns of Σ and V are separated from the rest of the matrix by dotted lines.

Since these matrices often arise from numerical calculations, it is natural to ask what to do with singular values that are approximately zero within some defined

threshold. If these are truncated, the decomposition forms an approximation to M .

Where M is approximated by its SVD expansion truncated to order r ,

$$M \simeq M_r = \sum_{i=1}^r u_i \sigma_i v_i^*, \quad (4)$$

it can be seen that the approximation error is equal to the rest of the expansion (the ‘tail’),

$$Ma - M_r a = \sum_{i=r+1}^m u_i \sigma_i v_i^* a, \quad (5)$$

and so is bounded,

$$\|Ma - M_r a\| \leq \sigma_{r+1} \|a\|. \quad (6)$$

The effect of rank reduction can be seen by looking at the following Matlab code extract, which applies the SVD to images.

```
[U, S, V] = svd(img);
approx_img = U(:,1:r) * S(1:r,1:r) * V(:,1:r)';
```

The output for a pair of sample images is shown in Figure 2.3

2.3 The pseudo-inverse

For reference, we briefly mention the pseudo-inverse here. A better and more detailed introduction (with proofs) is presented in [3]. The inverse of a matrix M in terms of its SVD is simply

$$M = U \Sigma V^* \\ M^{-1} = V \Sigma^{-1} U^*.$$

Clearly Σ^{-1} only exists if M is full rank and square. Otherwise, we define the pseudo-inverse (or Moore-Penrose inverse) via the reduced SVD,

$$M^+ = V_r \Sigma_r^{-1} U_r^*$$

where Σ_r is the truncation of Σ to remove all zero singular values, so that Σ_r is invertible. M^+ is a one-sided inverse (which side depends on whether $m > n$; for example, if $m > n$, then $MM^+ = I_n$).

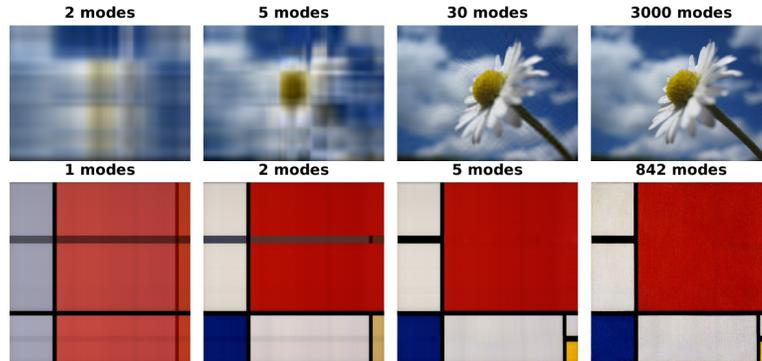


Figure 2.3: The singular value decomposition applied to image reconstruction, showing progressively higher rank approximations of an image. The upper image series requires a high number of modes to capture the detail. The lower image series shows that the Mondrian picture is well-approximated by a very low-rank projection.

Consider the under- or over-determined linear system of equations

$$Mx = b. \tag{7}$$

A least squares solution x' that minimises $\|Mx - b\|$ is given by

$$x' = M^+b.$$

In the case where M does not have linearly independent columns, the solution x' (of all possible solutions) given by the pseudoinverse is the solution that has the minimum length.

2.4 The singular value decomposition for linear operators

The matrix SVD has a direct analogy for linear operators on Hilbert spaces, called either the Schmidt decomposition or sometimes also the singular value decomposition. The reader not used to dealing with functions in and operators on Hilbert spaces may rest assured that the situation is conceptually very similar to the matrix case. A good and detailed reference is [4]. The following statement is equivalent to that in [5]. In the following, $\langle \cdot, \cdot \rangle_X$ represents the inner product on the space X .

Theorem 2.1. If $T : X \rightarrow Y$ is a compact (bounded, linear) operator, where X and Y are Hilbert spaces, then T has the following representation:

$$Tx = \sum_{i=1}^{\infty} \sigma_i \langle x, \phi_i \rangle_X \psi_i, \quad (8)$$

for some $x \in X$ where the set $\{\phi_i\}$ and the set $\{\psi_i\}$ are the eigenvectors of T^*T and TT^* respectively, and $\sigma_i \geq 0$ are the square roots of the eigenvalues. The $\{\phi_i\}$ and $\{\psi_i\}$ form an orthonormal basis for X and Y respectively (so $\langle \psi_i, \psi_j \rangle_Y = \delta_{ij}$ and so on). A pair (ϕ_i, ψ_i) is a Schmidt pair of T with an associated singular value σ_i and (8) is the Schmidt decomposition of T . Further, T is bounded with norm σ_1 , so $\|Tx\| \leq \sigma_1 \|x\|$.

Fortunately, many useful properties and much of the intuition arising from the simpler matrix SVD carry over to the operator case. The most obvious difference is that since X and Y are function spaces, there can be infinitely many singular values. Even then, T can still be approximated by a finite-rank operator with bounded error, in a manner analogous to the matrix case. There are a few points to be aware of in a numerical implementation. Since T is a mapping between Hilbert spaces, inner products will have been defined for X and Y . Therefore, care must be taken to be sure that any discretisation preserves the appropriate inner product in the form of whatever mass matrix, grid weighting or similar is appropriate. This is often a tedious step and not explicitly outlined in papers.

2.5 Further reading

The singular value decomposition is an extremely well established and widely used piece of mathematics. For the matrix case, the reader may find the presentations in [3] or [6] useful and insightful. The practical application in a control setting is well explained in [7]. The operator case is thoroughly covered in [4] and, in an infinite-dimensional linear systems setting, introduced in [5].

2.6 Exercises

Exercise 2.1. Determine the relationship between the SVD of M and the eigenvalues and eigenvectors of MM^* and of M^*M .

Exercise 2.2. Find the singular values of AM where A is a unitary matrix.

Exercise 2.3. Find the minimum singular value of $\lambda I - M$ where λ is an eigenvalue of M .

Exercise 2.4. For the operator case, and using the eigenvalues and eigenvectors of T^*T , show that $\|Tx\| \leq \sigma_1 \|x\|$.

3 Resolvent analysis of dynamical systems

Resolvent analysis in various forms has emerged as a useful tool in fluid dynamics. Here, we present the type introduced in [8] which is applicable to a broad range of systems, including nonlinear systems far from a steady equilibrium, such as turbulence. To prevent the general idea from being obscured by details involving fluid dynamics, we start with the general case, then specialise to the Navier-Stokes equations. The basic idea presented here is relatively simple and can be understood using just Fourier transforms, linear algebra and some familiarity with state space.

The approach differs from classical linearisation in the following way. A linearisation would typically proceed by doing the Taylor expansion about an equilibrium and assuming the perturbations around it to be small, leading us to neglect higher-order terms in the perturbations. This allows some qualitative statements to be made about the region near the equilibrium. Clearly, it is convenient to expand around a point where the equation associated with that point has no time derivatives; this is true at an equilibrium but is also true at the mean. In contrast, although in our case $\frac{d}{dt}\bar{\mathbf{z}}$ is zero, the nonlinear terms are not small, so are kept.

We begin by illustrating the idea using a dynamical system with a finite-dimensional state space. Let the state at time t be $\mathbf{z}(t)$, let $\mathbf{z}(t) \in \mathbb{R}^n$, and the dynamics be given by a function \mathbf{g} , $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^n$;

$$\frac{d}{dt}\mathbf{z}(t) = \mathbf{g}(\mathbf{z}(t)). \quad (9)$$

The aim is to get some sense of what the dynamics of the system look like without having to time-integrate it. The approach is to put the system into a form where we can apply tools from linear algebra, even though the system is nonlinear. To do this, we will go into the frequency domain. Once there, a linear operator arising from an expansion around the mean is formed. The most amplified directions of this operator are found, and are assumed to be excited by the remaining nonlinear terms in the expansion. These most amplified directions represent most favoured motions in the state space for any given frequency.

In the following, we assume we know the long-time average of the state, $\bar{\mathbf{z}}$.¹ The

¹This is obviously a weakness of the method if it needs to be found by time integration, but in some cases it may be known by other means such as experiment.

Taylor expansion of (9) about $\bar{\mathbf{z}}$ is

$$\frac{d}{dt}\mathbf{z}(t) = \frac{d}{dt}\bar{\mathbf{z}} + \frac{d}{dt}(\mathbf{z}(t) - \bar{\mathbf{z}}) \quad (10)$$

$$= \mathbf{g}(\bar{\mathbf{z}}) + \left. \frac{\partial \mathbf{g}}{\partial \mathbf{z}} \right|_{\bar{\mathbf{z}}} (\mathbf{z}(t) - \bar{\mathbf{z}}) + \text{H.O.T.} \quad (11)$$

$$= \mathbf{L}(\mathbf{z}(t) - \bar{\mathbf{z}}) + \mathbf{f}(t) \quad (12)$$

where \mathbf{f} collects together the terms nonlinear in the fluctuations $(\mathbf{z}(t) - \bar{\mathbf{z}})$ and the constant term $\mathbf{g}(\bar{\mathbf{z}})$. \mathbf{L} is the Jacobian of \mathbf{g} about $\bar{\mathbf{z}}$. This expansion is essentially just a change of variables with the new origin being the mean. This splitting into mean and fluctuations is basically a Reynolds decomposition. The expansion around $\bar{\mathbf{z}}$ is then Fourier transformed. This involves the Fourier transform of \mathbf{z} ,²

$$\hat{\mathbf{z}}(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} \mathbf{z}(t) dt \quad (13)$$

with inverse transform

$$\mathbf{z}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} \hat{\mathbf{z}}(\omega) d\omega, \quad (14)$$

and of \mathbf{f} ,

$$\hat{\mathbf{f}}(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} \mathbf{f}(t) dt. \quad (15)$$

The mean, $\bar{\mathbf{z}}$, is closely related to the $\omega = 0$ component. It can be verified by substitution into (14) that

$$\hat{\mathbf{z}}(\omega = 0) = 2\pi\bar{\mathbf{z}}\delta(\omega), \quad (16)$$

with δ being the Dirac delta, and the equation corresponding to $\omega = 0$ is therefore the mean equation.

To proceed, we integrate the expansion (12) against a chosen frequency ω . At any particular $\omega \neq 0$ we then have

$$i\omega\hat{\mathbf{z}}(\omega) = \mathbf{L}\hat{\mathbf{z}}(\omega) + \hat{\mathbf{f}}(\omega). \quad (17)$$

which can be rearranged as

$$\hat{\mathbf{z}}(\omega) = (i\omega I - \mathbf{L})^{-1} \hat{\mathbf{f}}(\omega). \quad (18)$$

²Really, this integral diverges, because our signal is not bounded.

The operator $\mathbf{R}(\omega) := (i\omega I - \mathbf{L})^{-1}$ which maps the Fourier coefficient of the non-linear excitation, $\hat{\mathbf{f}}$ to the Fourier coefficient of the state, $\hat{\mathbf{z}}$, is the resolvent operator (matrix, in this case) of \mathbf{L} . It is essentially a transfer function from $\hat{\mathbf{f}}$ to $\hat{\mathbf{z}}$.

It is worth pausing here to think about what has been done. The benefit of this approach is that it brings us to a point where the familiar and powerful tools of linear algebra can be applied, despite the actual dynamics being highly nonlinear. The nonlinear terms in $\hat{\mathbf{f}}$, rather than being discarded, are essential and act to excite the state. Since the effect of these terms is filtered by the linear part of the dynamics, we will be able to say something about how the filtering affects the state. The tool of choice in this situation is the singular value decomposition, which we will apply to the resolvent.

3.1 The singular value decomposition of the resolvent operator

The question at hand is what typical motions in state space are to be expected. From the preceding discussion we have shown that a matrix \mathbf{R} maps some unknown vector $\hat{\mathbf{f}}$ to the state's Fourier coefficient $\hat{\mathbf{z}}$. This matrix can be found from the governing equations and is the resolvent of the Jacobian formed around the long-time average state. We shall seek to find a basis in which to express $\hat{\mathbf{z}}$ as accurately as possible and with as few coefficients as possible, but without knowing $\hat{\mathbf{z}}$. This means we have to guess what $\hat{\mathbf{z}}$ is going to look like. To do this, we find the optimal approximation to $\mathbf{R}(\omega)$. Since $\mathbf{R}(\omega)$ is just a matrix, this is given by the singular value decomposition at each ω ,

$$\mathbf{R}(\omega) = \mathbf{U}(\omega)\Sigma(\omega)\mathbf{V}^*(\omega). \quad (19)$$

This decomposition at any particular frequency ω induces a basis in which to express $\hat{\mathbf{z}}$, which is the left singular vectors (the columns of $\mathbf{U}(\omega)$). In the absence of any knowledge about $\hat{\mathbf{f}}$, we may truncate this basis to order r with the hope that the relative smallness of the trailing singular values of $\mathbf{R}(\omega)$, that is, $\sigma_j(\omega)$ with $j > r + 1$, act to preclude any component of $\hat{\mathbf{z}}$ in the directions $\hat{\mathbf{u}}_j$, with $j > r + 1$.

The Fourier coefficients of the state can be expanded in terms of the left singular vectors of $\mathbf{R}(\omega)$,

$$\hat{\mathbf{z}}(\omega) = \mathbf{R}(\omega)\hat{\mathbf{f}}(\omega) \quad (20)$$

$$= \sum_{i=1}^r \mathbf{u}_i(\omega)\sigma_i(\omega)\mathbf{v}_i^*(\omega)\hat{\mathbf{f}}(\omega) \quad (21)$$

$$= \sum_{i=1}^r \mathbf{u}_i(\omega)\sigma_i(\omega)c_i(\omega). \quad (22)$$

The truncation to order r defines an optimally reduced space that $\hat{\mathbf{z}}(\omega)$ inhabits. The c_i are the coefficients obtained by the projection of $\hat{\mathbf{f}}(\omega)$ onto the right singular vectors, i.e. $c_i = \mathbf{v}_i^*(\omega)\hat{\mathbf{f}}(\omega)$. To fix the coefficients c_i and calculate a specific trajectory \mathbf{z} would require invoking the nonlinear term.

3.2 Exercises

In [1], the resolvent operator formed around the laminar flow is suggested as a model for transition. Many of the ideas presented in these notes are introduced. In the same paper, the simple conceptual model system is proposed,

$$\frac{d\mathbf{u}}{dt}(t) = \begin{bmatrix} -1/\rho & 1 \\ 0 & -2/\rho \end{bmatrix} + \|\mathbf{u}\| \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \mathbf{u}. \quad (23)$$

Exercise 3.1. Find the eigenvalues of the linear operator in the model system.

Exercise 3.2. Using the same linear operator, form the resolvent operator for $\rho = 25$. Plot the leading singular value as it varies with ω . What happens to the leading singular value when there is an eigenvalue close to the imaginary axis?

4 Nonlinear feedback, solutions to the nonlinear system and the Lur'e decomposition

In this section, we seek to understand the global behaviour when both nonlinear and linear parts of the system co-exist. Some authors eschew this interpretation and instead choose to model $\hat{\mathbf{f}}$ statistically using some prior knowledge obtained through other means. Both approaches are reasonable.

Consider the dynamics in (12),

$$\frac{d}{dt}\tilde{\mathbf{z}}(t) = \mathbf{L}\tilde{\mathbf{z}}(t) + \mathbf{f}(t) \quad (24)$$

where $\tilde{\mathbf{z}}$ is defined to be the fluctuations at time t , $\tilde{\mathbf{z}} := \mathbf{z}(t) - \bar{\mathbf{z}}$. This system can be interpreted as a linear system with external forcing $\mathbf{f}(t)$. As such, it can be integrated from some initial condition to solve for $\tilde{\mathbf{z}}(t)$. We have dealt with $\tilde{\mathbf{z}}(t)$ (the state at some time t) as a vector in \mathbb{R}^n , but we can also talk of the solution being a function $\tilde{\mathbf{z}} : \mathbb{R} \rightarrow \mathbb{R}^n$ which maps a real number (t) to a point in \mathbb{R}^n (the state at

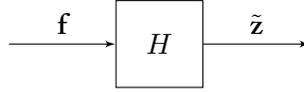


Figure 4.1: A block diagram representation of the open-loop relation H .

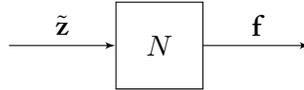


Figure 4.2: A block diagram representation of the open-loop nonlinear relation N .

t).³ This trajectory can be related to the equivalent forcing \mathbf{f} by a linear operator H ,

$$\tilde{\mathbf{z}} = H\mathbf{f}. \quad (25)$$

That is to say, H maps a whole history of $\mathbf{f}(t)$ (the signal \mathbf{f} in the space of such functions) to a whole history of $\tilde{\mathbf{z}}(t)$ (a trajectory in the space of such functions). As such, if \mathbf{f} was known, applying H would result in $\tilde{\mathbf{z}}$, which could give the state at any time t by evaluating $\tilde{\mathbf{z}}(t)$. H could be calculated the time-integration of 24 with an initial condition. This is depicted in Figure 4.1.

The counterpart is the nonlinear relation

$$\mathbf{f} = N(\tilde{\mathbf{z}}) \quad (26)$$

where N calculates the function \mathbf{f} from the state trajectory $\tilde{\mathbf{z}}$. Again, the nonlinear term $\mathbf{f}(t)$ is calculated by supplying t as an argument to \mathbf{f} . In the special case that N is memoryless, $\mathbf{f}(t)$ can be calculated instantaneously from $\tilde{\mathbf{z}}(t)$ without knowing the whole time history.

Similarly to the open loop relation H , N maps a signal $\tilde{\mathbf{z}}$ to an \mathbf{f} . As such, if $\tilde{\mathbf{z}}$ was known, the whole history of $\mathbf{f}(t)$ could be reproduced finding \mathbf{f} and then evaluating at t . This is depicted in Figure 4.2.

A solution to the original equations is a pair $(\mathbf{f}, \tilde{\mathbf{z}})$ that simultaneously satisfies both open-loop relations. The ‘closed-loop’ is such an arrangement, depicted in Figure 4.3. It should be noted that since H is linear, fixing the amplitudes of $\hat{\mathbf{z}}$ can only occur in the closed-loop situation.

³Care should be taken to understand that this function represents a whole trajectory. A confusion often arises at this point because of the common practice of omitting the argument t when discussing the instantaneous state, thus leading to a confusion between $\tilde{\mathbf{z}}$ and $\tilde{\mathbf{z}}(t)$.

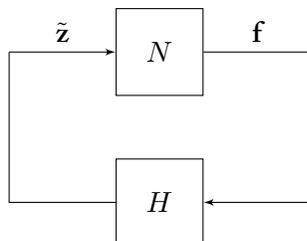


Figure 4.3: A block diagram representation of the closed-loop system consisting of H and N .

Such a decomposition into a linear and memoryless nonlinear parts is called a Lur'e decomposition. Its use is foundational to the study of stability in modern robust control theory and is well explained in two beautiful papers by Zames [9, 10]. The approach to proving global stability in those papers rests on finding conditions where only one pair $(\mathbf{f}, \tilde{\mathbf{z}})$ is consistent with both open-loop elements. In general, fluid systems are not globally stable and so many solutions may arise. The use of this decomposition first appeared in the analysis of Navier-Stokes systems in [11] and [12] which used sector-bounding arguments applied to stability/control and model reduction respectively.

5 Derivation of the resolvent operator for the turbulent NSE

In this section we review how to form the resolvent operator for the incompressible Navier-Stokes equations. There are minor complications arising from the pressure term. In the case where you are able to express the velocity in terms of a divergence-free basis, these difficulties are avoided. The more general case, including pressure, is presented here.

5.1 The Fourier transformed Navier-Stokes-Equations

The non-dimensionalised incompressible Navier-Stokes equations are

$$\begin{aligned} \partial_t \mathbf{u}(\mathbf{x}, t) + \mathbf{u}(\mathbf{x}, t) \cdot \nabla \mathbf{u}(\mathbf{x}, t) &= -\nabla p(\mathbf{x}, t) + \frac{1}{Re} \nabla^2 \mathbf{u}(\mathbf{x}, t) \\ \nabla \cdot \mathbf{u}(\mathbf{x}, t) &= 0, \end{aligned} \quad (27)$$

where $\mathbf{x} \in X$ is a point in the physical flow domain, t is time, $Re = \rho UL/\mu$ is the Reynolds number, U is a characteristic velocity, L is a length scale and ∇ is the gradient operator on X . The density of the fluid is ρ and its viscosity μ . Velocity and

pressure are thus non-dimensional, with velocity scaled by U and pressure scaled by ρU^2 .

As in the general case, first consider the temporal Fourier transform for the state \mathbf{u} ,

$$\hat{\mathbf{u}}(\mathbf{x}, \omega) = \int_{-\infty}^{\infty} e^{-i\omega t} \mathbf{u}(\mathbf{x}, t) dt, \quad (28)$$

From the previous discussion, we know that $\hat{\mathbf{u}}(\mathbf{x}, \omega = 0)$ gives the temporal mean velocity. Define the fluctuations about the mean as $\tilde{\mathbf{u}}(\mathbf{x}, t) = \mathbf{u}(\mathbf{x}, t) - \bar{\mathbf{u}}$. Now, treat the pressure similarly to the velocity,

$$\hat{p}(\mathbf{x}, \omega) = \int_{-\infty}^{\infty} e^{-i\omega t} p(\mathbf{x}, t) dt. \quad (29)$$

Finally, Fourier transform $\tilde{\mathbf{u}}(\mathbf{x}, t) \cdot \nabla \tilde{\mathbf{u}}(\mathbf{x}, t)$ in the same way,

$$\hat{\mathbf{f}}(\mathbf{x}, \omega) = \int_{-\infty}^{\infty} e^{-i\omega t} \tilde{\mathbf{u}}(\mathbf{x}, t) \cdot \nabla \tilde{\mathbf{u}}(\mathbf{x}, t) dt, \quad (30)$$

noticing that $\hat{\mathbf{f}}(\mathbf{x}, \omega = 0)$ yields the time average $\overline{\tilde{\mathbf{u}} \cdot \nabla \tilde{\mathbf{u}}}$, which does not have to be zero, even though the time average of $\tilde{\mathbf{u}}$ itself is.

Using these definitions and taking the Fourier transform of the Navier-Stokes equations,

$$\begin{aligned} & \int_{-\infty}^{\infty} e^{-i\omega t} \left[\partial_t \tilde{\mathbf{u}}(\mathbf{x}, t) \right. \\ & \quad + \tilde{\mathbf{u}}(\mathbf{x}, t) \cdot \nabla \tilde{\mathbf{u}}(\mathbf{x}, t) + \bar{\mathbf{u}}(\mathbf{x}) \cdot \nabla \bar{\mathbf{u}}(\mathbf{x}) \\ & \quad + \bar{\mathbf{u}}(\mathbf{x}) \cdot \nabla \tilde{\mathbf{u}}(\mathbf{x}, t) + \tilde{\mathbf{u}}(\mathbf{x}, t) \cdot \nabla \bar{\mathbf{u}}(\mathbf{x}) \\ & \quad + \nabla (\bar{p}(\mathbf{x}) + \tilde{p}(\mathbf{x}, t)) \\ & \quad \left. - \frac{1}{Re} \nabla^2 (\bar{\mathbf{u}}(\mathbf{x}) + \tilde{\mathbf{u}}(\mathbf{x}, t)) \right] dt = 0. \end{aligned}$$

This yields both the fluctuation equation and the RANS equation. The former ($\omega \neq 0$) is

$$\begin{aligned} i\omega \hat{\mathbf{u}}(\mathbf{x}, \omega) + \bar{\mathbf{u}}(\mathbf{x}) \cdot \nabla \hat{\mathbf{u}}(\mathbf{x}, \omega) + \hat{\mathbf{u}}(\mathbf{x}, \omega) \cdot \nabla \bar{\mathbf{u}}(\mathbf{x}) \\ = -\nabla \hat{p}(\mathbf{x}, \omega) + \frac{1}{Re} \nabla^2 \hat{\mathbf{u}}(\mathbf{x}, \omega) - \hat{\mathbf{f}}(\mathbf{x}, \omega) \end{aligned} \quad (31)$$

and the latter ($\omega = 0$) is

$$\bar{\mathbf{u}}(\mathbf{x}) \cdot \nabla \bar{\mathbf{u}}(\mathbf{x}) = -\nabla \bar{p}(\mathbf{x}) + \frac{1}{Re} \nabla^2 \bar{\mathbf{u}}(\mathbf{x}) - \bar{\mathbf{f}}(\mathbf{x}). \quad (32)$$

These equations include the pressure. The next steps are in order to eliminate it.

For later convenience, define

$$L = \bar{\mathbf{u}}(\mathbf{x}) \cdot \nabla + (\nabla \bar{\mathbf{u}}(\mathbf{x}))^T$$

and the Leray projection [13]

$$\Pi = (I - \nabla(\nabla^2)^{-1}\nabla).$$

We will use the Leray projection to enforce incompressibility allowing the pressure term to disappear. Care must be taken to specify the boundary conditions when inverting the Laplacian. Using incompressibility, taking the divergence of both sides of the NSE gets rid of the time derivative and gives the pressure Poisson equation, which relates the pressure and the velocity field instantaneously,

$$-\nabla^2 p(\mathbf{x}, t) = \nabla \cdot (\mathbf{u}(\mathbf{x}, t) \cdot \nabla \mathbf{u}(\mathbf{x}, t)).$$

Taking the Fourier transform, and using the Fourier transforms defined earlier, the mean component is

$$-\nabla^2 \bar{p}(\mathbf{x}) = \nabla \cdot (\bar{\mathbf{u}}(\mathbf{x}) \cdot \nabla \bar{\mathbf{u}}(\mathbf{x})) + \nabla \cdot \bar{\mathbf{f}}(\mathbf{x})$$

and at other frequencies,

$$\nabla^2 \hat{p}(\mathbf{x}, \omega) = -\nabla \cdot L \hat{\mathbf{u}}(\mathbf{x}, \omega) - \nabla \cdot \hat{\mathbf{f}}(\mathbf{x}, \omega). \quad (33)$$

Thus, \hat{p} is related to $\hat{\mathbf{u}}$ and $\hat{\mathbf{f}}$ via a linear operator. Substituting these into (??), we get

$$\left(i\omega I + \Pi L - \frac{1}{Re} \nabla^2 \right) \hat{\mathbf{u}}(\mathbf{x}, \omega) = -\Pi \hat{\mathbf{f}}(\mathbf{x}, \omega). \quad (34)$$

The operator

$$R(\omega) := - \left(i\omega I + \Pi L - \frac{1}{Re} \nabla^2 \right)^{-1} \Pi \quad (35)$$

is the resolvent of the NSE linearised about the mean, and

$$\hat{\mathbf{u}}(\mathbf{x}, \omega) = R(\omega) \hat{\mathbf{f}}(\mathbf{x}, \omega). \quad (36)$$

Notice that R depends on the time-average velocity field $\bar{\mathbf{u}}$ and on frequency ω , and that $\bar{\mathbf{u}}$ appears naturally and with a consistent interpretation. The importance of the spectrum of the linear operator about $\bar{\mathbf{u}}$ thus has a clear interpretation even in a fully nonlinear flow. The RANS equation (32) and the resolvent equations (36) are connected via the Fourier transform, as shown in Figure 5.1. Notice also that $\mathbf{f} = \bar{\mathbf{u}} \cdot \nabla \bar{\mathbf{u}}$ acts as a memoryless nonlinearity in the sense of Section 4, so the earlier discussion on solutions applies.

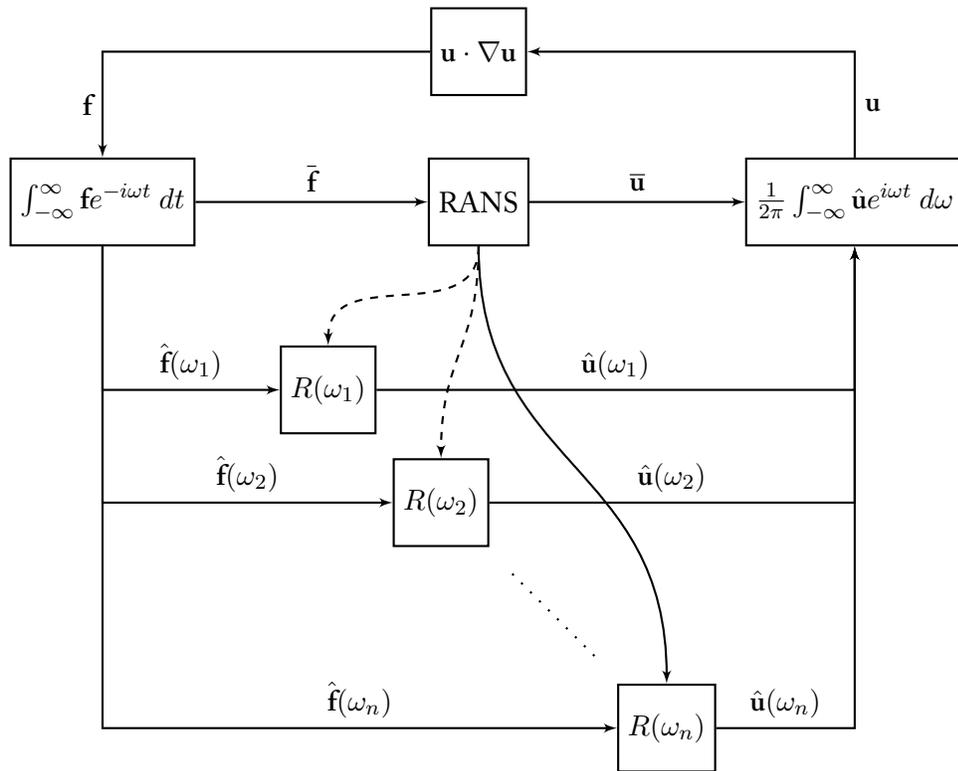


Figure 5.1: A schematic block diagram, showing the network of resolvent operators and the nonlinear terms which compose the frequency-domain representation of the NSE.

6 The application of the singular value decomposition to the resolvent

We should like to find a sensible basis in which to expand $\hat{\mathbf{u}}$. Ideally, the functions should be orthonormal (to make projections and expansions simple), and chosen and ordered in such a way that a truncation of the expansion should still approximate the true $\hat{\mathbf{u}}$ in a quantifiable way.

Like other modal decompositions including proper orthogonal decomposition [14] or dynamic mode decomposition [15], we will use the singular value decomposition, but on a dynamical flow operator instead of a dataset.

In the case that the dynamics are translation-invariant (such as the streamwise direction in an infinite pipe), the Fourier transform already provides a unitary basis, and we should immediately take the Fourier transform in those directions. In non-invariant directions, such as a wall-normal direction, this is not the case, and we proceed as follows.

Notice that $R(\omega)$ is a linear mapping from the Fourier-transformed ‘forcing’ field to the Fourier-transformed ‘velocity’ field. Because there is no reason to expect the nonlinear forcing \mathbf{f} to look like the velocity field, we should expect to use two different bases for $\hat{\mathbf{u}}$ and $\hat{\mathbf{f}}$. One sensible way to form a basis is the Schmidt decomposition of $R(\omega)$, which provides different bases for $\hat{\mathbf{u}}$ and $\hat{\mathbf{f}}$ and is optimal in useful ways. With this choice,

$$\begin{aligned}
 R(\omega)\hat{\mathbf{f}}(\mathbf{x}, \omega) &= \sum_{j \in \mathbb{N}} \sigma_j(\omega) \left\langle \hat{\mathbf{f}}(\mathbf{x}, \omega), \phi_j(\mathbf{x}, \omega) \right\rangle_X \psi_j(\mathbf{x}, \omega), & (37) \\
 \left\langle \psi_j(\mathbf{x}, \omega), \psi_{j'}(\mathbf{x}, \omega) \right\rangle_X &= \delta_{j,j'}, \\
 \left\langle \phi_j(\mathbf{x}, \omega), \phi_{j'}(\mathbf{x}, \omega) \right\rangle_X &= \delta_{j,j'}, \\
 \sigma_j(\omega) &\geq \sigma_{j+1}(\omega). & (38)
 \end{aligned}$$

The pairs ϕ_j and ψ_j at each ω are the Schmidt pairs (singular vectors) in the decomposition. The sets of ϕ_j and ψ_j at each ω each form an orthonormal basis (under the inner product on the spatial domain X), with basis functions ordered by the singular values σ_j . This ordering provides a criterion for truncation. Note that the basis is different for each frequency, which is to be expected, since different motions will ‘resonate’ in the flow at different frequencies.

From Section 2.1 we know that a particular choice of inner product is implicit in the use of the SVD. For the incompressible Navier-Stokes equations, calculations involving energy budgets and the nonlinearity are made simpler by the choice of

the unweighted spatial \mathcal{L}_2 norm, but other systems (such as compressible or reacting flows) require further thought. Care should be taken when implementing the Schmidt decomposition in numerical codes (i.e. translating the Schmidt decomposition to a discrete, matrix SVD) to get the weighting matrix associated with the discretisation scheme employed to correspond correctly to the desired inner product.

In fluids applications, $R(\omega)$ often has very large separation between the leading (one or two) singular values and the next. The physical basis for this is well documented in the literature [8] and (in turbulent shear flows) is largely due to a resonance associated with the critical layer. It also turns out that in shear flows, R is non-normal, so in general $\psi_j(\mathbf{x}, \omega) \neq \phi_j(\mathbf{x}, \omega)$.

The decomposition leads naturally to ordered expansions for both $\hat{\mathbf{u}}$ and $\hat{\mathbf{f}}$ at any particular frequency; expressing each as a weighted superposition of its basis functions gives

$$\hat{\mathbf{u}}(\mathbf{x}, \omega) = \sum_{j=1}^{\infty} \chi_j(\omega) \sigma_j(\omega) \psi_j(\mathbf{x}, \omega), \quad (39)$$

$$\hat{\mathbf{f}}(\mathbf{x}, \omega) = \sum_{j=1}^{\infty} \chi_j(\omega) \phi_j(\mathbf{x}, \omega). \quad (40)$$

The set of ψ_j are sometimes called the response modes (or resolvent modes) and the set of ϕ_j the forcing modes. At this point, we do not know the scalar coefficients χ_i . This decomposition truncated up to rank r is illustrated in Figure 6.1.

If it so happens that $\sigma_1 \gg \sigma_2$, without knowing much about $\hat{\mathbf{f}}$, we may reasonably approximate the Fourier coefficient of $\hat{\mathbf{u}}$ (up to a complex coefficient) by $\hat{\mathbf{u}}(\mathbf{x}, \omega) \cong \psi_1(\mathbf{x}, \omega)$, regardless of our knowledge of $\hat{\mathbf{f}}$. By ‘up to a complex coefficient’ it is meant that, while the functional form of each response mode is known, the complex coefficient of the response mode (determining both the phase and magnitude of the wave) is not determined by the decomposition. However, because the *relative* phase between forcing and response mode pairs is fixed by the decomposition, the phase between different response modes may be fixed either via a direct calculation of the nonlinear forcing, by a projection onto DMD modes, by fitting to a limited set of measurements [16, 17], or by other methods [18, 19].

Retaining just the first mode per frequency (a rank-1 approximation) is essentially the same calculation as the ‘optimal response’ found by various authors [20]. The approximation argument does not hold in reverse; which is to say, to approximate $\hat{\mathbf{f}}$ it would require R^{-1} to be approximately rank-1, which is not usually the case.

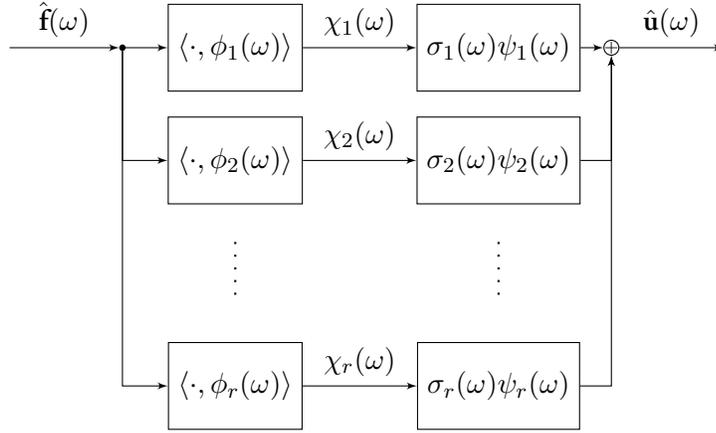


Figure 6.1: The NSE resolvent operator, decomposed into its forcing and response modes and truncated to order r . The forcing $\hat{\mathbf{f}}(\omega)$ at frequency ω is projected onto the forcing modes, to give their scalar coefficients χ_i . These are then amplified by $\sigma_i(\omega)$ and multiplied by the response modes $\psi_i(\omega)$. The superposition of these responses gives the velocity field Fourier coefficient $\hat{\mathbf{u}}(\omega)$.

In some systems, $\chi_j(\omega)$ for the leading few modes may be quite small relative to the following coefficients. This effect may outweigh the effect of any separation of the leading singular values. This may happen, for instance, because the gradient operator involved in calculating the nonlinear term $\hat{\mathbf{f}}$ can act to attenuate the larger scales important at lower frequencies. In such cases, it is reasonable to approximate R by a higher-rank projection induced by (38), with the number of modes retained determining the level of accuracy, as in [18]. The extent to which either scenario applies will depend on the particularities of the system under study.

References

- [1] L. N. Trefethen, A. E. Trefethen, S. C. Reddy, and T. A. Driscoll, “Hydrodynamics stability without eigenvalues,” *Science*, vol. 261, no. 5121, pp. 578–584, 1993.
- [2] M. R. Jovanović and B. Bamieh, “Componentwise energy amplification in channel flows,” *Journal of Fluid Mechanics*, vol. 534, 2005.
- [3] G. Strang, *Linear Algebra and its Applications*. Academic Press, 1976.
- [4] N. Young, *An introduction to Hilbert space*. Cambridge University Press., 1988.

- [5] R. F. Curtain and H. J. Zwart, *An Introduction to Infinite-Dimensional Linear Systems Theory*. New York: Springer-Verlag, 1995.
- [6] L. Trefethen and D. Bau, *Numerical Linear Algebra*. Society for Industrial and Applied Mathematics, 1997.
- [7] W. J. Green and D. J. N. Limebeer, *Linear Robust Control*. New Jersey: Prentice Hall, 1995.
- [8] B. J. McKeon and A. S. Sharma, “A critical-layer framework for turbulent pipe flow,” *Journal of Fluid Mechanics*, vol. 658, p. 336–382, July 2010.
- [9] G. Zames, “On the input-output stability of time-varying nonlinear feedback systems - Part II: Conditions involving circles in the frequency plane and sector nonlinearities,” *IEEE Trans. on Automatic Control*, 1966.
- [10] G. Zames, “On the input-output stability of time-varying nonlinear feedback systems - Part I: Conditions derived using concepts of loop gain, conicity and positivity,” *IEEE Trans. on Automatic Control*, vol. AC-11, pp. 228–238, April 1966.
- [11] A. S. Sharma, D. J. N. Limebeer, B. J. McKeon, and J. F. Morrison, “Stabilising control laws for the incompressible Navier-Stokes equations using sector stability theory,” in *Proceedings of the 3rd AIAA Flow Control Conference, San Francisco, California*, American Institute of Aeronautics and Astronautics, 2005.
- [12] A. S. Sharma, “Model reduction of turbulent fluids flows using the supply rate,” *International Journal of Bifurcation and Chaos*, vol. 19, p. 1267–1278, 2009.
- [13] R. Temam, *Navier-Stokes Equations: Theory and Numerical Analysis*. AMS Chelsea Publishing, 2001.
- [14] P. Holmes, J. L. Lumley, and G. Berkooz, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge, U.K.: Cambridge University Press., first ed., 1996.
- [15] P. J. Schmid, “Dynamic mode decomposition of numerical and experimental data,” *Journal of Fluid Mechanics*, vol. 656, pp. 5 – 28, 2010.
- [16] F. Gómez, H. M. Blackburn, M. Rudman, A. S. Sharma, and B. J. McKeon, “A reduced-order model of three-dimensional unsteady flow in a cavity based on the resolvent operator,” *Journal of Fluid Mechanics*, vol. 798, June 2016.

- [17] S. Beneddine, R. Yegavian, D. Sipp, and B. Leclaire, “Unsteady flow dynamics reconstruction from mean flow and point sensors: an experimental study,” *Journal of Fluid Mechanics*, vol. 824, pp. 174–201, July 2017.
- [18] R. Moarref, M. R. Jovanović, J. A. Tropp, A. S. Sharma, and B. J. McKeon, “A low-order decomposition of turbulent channel flow via resolvent analysis and convex optimization,” *Phys. Fluids*, vol. 26, p. 051701, May 2014.
- [19] R. Moarref, A. S. Sharma, J. A. Tropp, and B. J. McKeon, “Model-based scaling of the streamwise energy density in high-Reynolds-number turbulent channels,” *Journal of Fluid Mechanics*, vol. 734, p. 275–316, Oct. 2013.
- [20] Y. Hwang and C. Cossu, “Amplification of coherent streaks in the turbulent Couette flow: an input–output analysis at low Reynolds number,” *Journal of Fluid Mechanics*, vol. 643, pp. 333–348, 2010.