Time-stepping and Krylov methods for large-scale instability problems

J.-Ch. Loiseau, M. A. Bucci, S. Cherubini, and J.-Ch. Robinet

Abstract ???

1 Introduction

2 Theoretical framework

Our attention is focused on the characterization of very high-dimensional nonlinear dynamical systems typically arising from the spatial discretization of partial differential equations such as the incompressible Navier-Stokes equations. In general, the resulting dynamical equations are written down as a system of first order differential equations

$$\dot{X}_j = \mathcal{F}_j\left(\{X_i(t);\ i = 1, \cdots, n\}, t\right)$$

where the integer n is the *dimension* of the system, and $\dot{X}_j$ denotes the time-derivative of X_j. Using the notation $\mathbf{X}$ and $\boldsymbol{\mathcal{F}}$ for the sets $\left\{X_j,\ i = 1, \cdots, n\right\}$ and $\left\{\mathcal{F}_j,\ i = 1, \cdots, n\right\}$, this system can be compactly written as

J.-Ch. Loiseau
Laboratoire DynFluid, Arts et Métiers ParisTech, 151 boulevard de l'hôpital, 75013 Paris, France. e-mail: jean-christophe.loiseau@ensam.eu

M. A. Bucci
Laboratoire DynFluid, Arts et Métiers ParisTech, 151 boulevard de l'hôpital, 75013 Paris, France. e-mail: michele.bucci@ensam.eu

S. Cherubini
DMMM, Politecnico di Bari, via Re David 200, 70100 Bari, Italy. e-mail: s.cherubini@gmail.com

J.-Ch. Robinet
Laboratoire DynFluid, Arts et Métiers ParisTech, 151 boulevard de l'hôpital, 75013 Paris, France. e-mail: jean-christophe.robinet@ensam.eu

$$\dot{\mathbf{X}} = \mathcal{F}(\mathbf{X}, t), \tag{1}$$

where $\mathbf{X}$ is the $n \times 1$ *state vector* of the system and t is a continuous variable denoting time. Alternatively, accounting also for temporal discretization gives rise to a discrete-time dynamical system

$$X_{j,k+1} = \Phi_j \left(\left\{ X_{i,k};\ i = 1, \cdots, n \right\}, k \right) \tag{2}$$

or formally

$$\mathbf{X}_{k+1} = \boldsymbol{\Phi}(\mathbf{X}_k, k) \tag{3}$$

where the index k now denotes the discrete time variable. If one uses first-order Euler extrapolation for the time discretization, the relation between $\mathcal{F}$ and $\boldsymbol{\Phi}$ is given by

$$\boldsymbol{\Phi}(\mathbf{X}) = \mathbf{X} + \Delta t \mathcal{F}(\mathbf{X}),$$

where Δt is the time-step and the explicit dependences on t and k have been dropped for the sake of simplicity.

In the rest of this section, the reader will be introduced to the concepts of fixed points and linear stability, two concepts required to characterize a number of properties of the system investigated. Particular attention will be paid to *modal* and *non-modal stability*, two approaches that have become increasingly popular in fluid dynamics over the past decades. Note that the concept of *nonlinear optimal perturbation*, which as raised a lot attention lately, is beyond the scope of the present contribution. For interested readers, please refer to the recent work by [2] and references therein.

Finally, while we will mostly use the continuous-time representation (1) when introducing the reader to the theoretical concepts exposed in this section, using the discrete-time representation (3) will prove more useful when discussing and implementing the different algorithms presented in §3.

2.1 Fixed points

Nonlinear dynamical systems described by Eq. (1) or Eq. (3) tend to admit a number of different equilibria forming the backbone of their phase space. These different equilibria can take the form of fixed points, periodic orbits or strange attractors for instance. In the rest of this work, our attention will be solely focused on fixed points.

For a continuous-time dynamical system described by Eq. (1), fixed points $\mathbf{X}^*$ are solution to

$$\mathcal{F}(\mathbf{X}) = 0. \tag{4}$$

Conversely, fixed points $\mathbf{X}^*$ of a discrete-time dynamical system described by Eq. (3) are solution to

$$\boldsymbol{\Phi}(\mathbf{X}) = \mathbf{X}. \tag{5}$$

It must be emphasized that both Eq. (4) and Eq. (5) may admit multiple solutions. Such a multiplicity of fixed points can easily be illustrated by a simple Duffing oscillator given by

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -\frac{1}{2}y + x - x^3. \end{aligned} \tag{6}$$

Despite its apparent simplicity, this Duffing oscillator admits three fixed points, namely

- a saddle at the origin $\mathbf{X}^* = (0,0)$,
- two linearly stable spirals located at $\mathbf{X}^* = (\pm 1, 0)$.

All of these fixed points, along with some trajectories, are depicted on figure 1 for the sake of illustration. Such a multiplicity of fixed points also occurs in dynamical systems as complex as the Navier-Stokes equations. Determining which of these fixed points is the most relevant one from a physical point of view is problem-dependent and left for the user to decide. Note however that computing these equilibrium points is a prerequisite to all of the analyses to be described in this chapter. Numerical methods to solve Eq. (4) or Eq. (5) will be discussed in §3.2.

2.2 Linear stability analysis

Having computed a given fixed point $\mathbf{X}^*$ of a continuous-time nonlinear dynamical system given by Eq. (1), one may ask whether it corresponds to a stable or unstable equilibrium of the system. Before pursuing, the very notion of *stability* needs to be explained. It is traditionally defined following the concept of Lyapunov stability. Having computed the equilibrium state $\mathbf{X}^*$, the system is perturbed around this state. If it returns back to the equilibrium point, the latter is deemed stable, otherwise, it

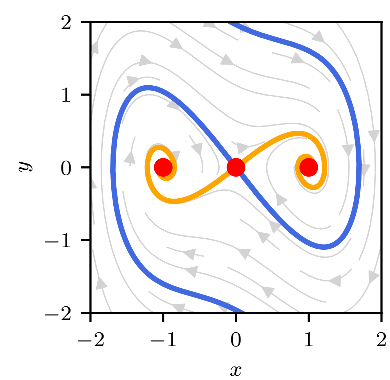

Fig. 1 Phase portrait of the unforced Duffing oscillator (6). The red dots denote the three fixed points admitted by the system. The blue (resp. orange) thick line depicts the stable (resp. unstable) manifold of the saddle point located at the origin. Grey lines highlight a few trajectories exhibited for different initial conditions.

is regarded as unstable. It has to be noted that, in the concept of Lyapunov stability, an infinite time horizon is allowed for the return to equilibrium.

The dynamics of a perturbation $\mathbf{x} = \mathbf{X} - \mathbf{X}^*$ are governed by

$$\dot{\mathbf{x}} = \mathcal{F}(\mathbf{X}^* + \mathbf{x}). \tag{7}$$

Assuming the perturbation $\mathbf{x}$ is infinitesimal, $\mathcal{F}(\mathbf{X})$ can be approximated by its first-order Taylor expansion around $\mathbf{X} = \mathbf{X}^*$. Doing so, the governing equations for the perturbation $\mathbf{x}$ simplify to

$$\dot{\mathbf{x}} = \boldsymbol{\mathcal{A}}\mathbf{x}, \tag{8}$$

where $\boldsymbol{\mathcal{A}}$ is the $n \times n$ Jacobian matrix of $\mathcal{F}$. Starting from an initial condition $\mathbf{x}_0$, the perturbation at time t is given by

$$\mathbf{x}(t) = \exp\left(\boldsymbol{\mathcal{A}}t\right)\mathbf{x}_0. \tag{9}$$

The operator $\boldsymbol{\mathcal{M}}(t) = \exp\left(\boldsymbol{\mathcal{A}}t\right)$ is known as the *exponential propagator*. Introducing the spectral decomposition of $\boldsymbol{\mathcal{A}}$

$$\boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{V}}\boldsymbol{\Lambda}\boldsymbol{\mathcal{V}}^{-1},$$

Eq. (9) can be rewritten as

$$\mathbf{x}(t) = \boldsymbol{\mathcal{V}}\exp\left(\boldsymbol{\Lambda}t\right)\boldsymbol{\mathcal{V}}^{-1}\mathbf{x}_0, \tag{10}$$

where the i$^{\text{th}}$ column of $\boldsymbol{\mathcal{V}}$ is the eigenvector $\mathbf{v}_i$ associated to the i$^{\text{th}}$ eigenvalue $\lambda_i = \boldsymbol{\Lambda}_{ii}$, with $\boldsymbol{\Lambda}$ a diagonal matrix. Assuming that the eigenvalues of $\boldsymbol{\mathcal{A}}$ have been sorted by decreasing real part, it can easily be shown that

$$\lim_{t \to +\infty} \exp\left(\boldsymbol{\mathcal{A}}t\right)\mathbf{x}_0 = \lim_{t \to +\infty} \exp\left(\lambda_1 t\right)\mathbf{v}_1.$$

The asymptotic fate of an initial perturbation $\mathbf{x}_0$ is thus entirely dictated by the real part of the leading eigenvalue λ_1:

- if $\Re\left(\lambda_1\right) > 0$, a random initial perturbation $\mathbf{x}_0$ will eventually grow exponentially fast. Hence, the fixed point $\mathbf{X}^*$ is deemed *linearly unstable*.
- If $\Re\left(\lambda_1\right) < 0$, the initial perturbation $\mathbf{x}_0$ will eventually decay exponentially rapidly. The fixed point $\mathbf{X}^*$ is thus *linearly stable*.

The case $\Re\left(\lambda_1\right) = 0$ is peculiar. The fixed point $\mathbf{X}^*$ is called *elliptic* and one cannot conclude about its stability solely by looking at the eigenvalues of $\boldsymbol{\mathcal{A}}$. In this case, one needs to resort to *weakly non-linear analysis* which essentially looks at the properties of higher-order Taylor expansion of $\mathcal{F}\left(\mathbf{X}\right)$. Once again, this is beyond the scope of the present chapter. Interested readers are referred to [**?**] for more details about such analyses.

Illustration

Let us illustrate the notion of linear stability on a simple example. For that purpose, we will consider the same linear dynamical system as in [3]. This system reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{1}{100} - \frac{1}{Re} & 0 \\ 1 & -\frac{2}{Re} \end{bmatrix}}_{\mathcal{A}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{11}$$

where Re is a control parameter. For such a simple case, it is obvious that the eigenvalues of $\mathcal{A}$ are given by

$$\lambda_1 = \frac{1}{100} - \frac{1}{Re}$$

and

$$\lambda_2 = -\frac{2}{Re}.$$

While λ_2 is constantly negative, λ_1 is negative for $Re < 100$ and positive otherwise. Figure 2 depicts the time-evolution of $\|\mathbf{x}\|_2^2 = x_1^2 + x_2^2$ for two different values of Re. Please note that the short-time ($t < 100$) behavior of the perturbation will be discussed in §2.3. It is clear nonetheless that, for $t > 100$, the time-evolution of the perturbation can be described by an exponential function. Whether this exponential increases or decreases as a function of time is solely dictated by the sign of λ_1, negative for $Re = 50$ and positive for $Re = 100$. For $Re = 50$, the equilibrium point $\mathbf{X}^* = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$ is thus stable, while it is unstable for $Re = 125$.

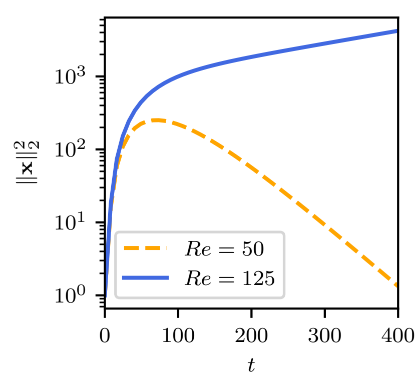

Fig. 2 Evolution as a function of time of $\|\mathbf{x}\|_2^2 = x_1^2 + x_2^2$ for the toy-model (11). For $Re = 50$ (resp. $Re = 125$), the asymptotic fate of $\|\mathbf{x}\|_2^2$ is described by a decreasing (resp. increasing) exponential. For $Re = 50$, the equilibrium point is thus linear stable, while it is linearly unstable for $Re = 125$.

2.3 Non-modal stability analysis

Looking at figure 2, it can be seen that, although the system is linearly stable for $Re = 50$, the perturbation $\mathbf{x}$ can experience a transient growth of its energy for a short period of time, roughly given by $0 < t < 100$ in the present case, before its eventual exponential decay. This behavior is related to the *non-normality* of $\boldsymbol{\mathcal{A}}$, i.e.

$$\boldsymbol{\mathcal{A}}^H \boldsymbol{\mathcal{A}} \neq \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{A}}^H, \tag{12}$$

where $\boldsymbol{\mathcal{A}}^H$ is the Hermitian (also known as the *adjoint*) of $\boldsymbol{\mathcal{A}}$. As a result of this non-normality, the eigenvectors of $\boldsymbol{\mathcal{A}}$ do not form an orthonormal set of vectors[1]. The consequences of this non-orthogonality of the set of eigenvectors can be visualized on figure 3 where the trajectory stemming from a random unit-norm initial condition $\mathbf{x}_0$ is depicted in the phase plane of our toy-model (11). The perturbation $\mathbf{x}(t)$ is first attracted toward the linear manifold associated to the least stable eigenvalue λ_1, causing in the process the transient growth of its energy by a factor 300. Once it reaches the vicinity of the linearly stable manifold, the perturbation eventually decays exponentially fast along this eigendirection of the fixed point. The next sections are devoted to the introduction of mathematical tools particularly useful to characterize phenomena resulting from this non-normality of $\boldsymbol{\mathcal{A}}$, both in the time and frequency domains, when the fixed point considered is stable.

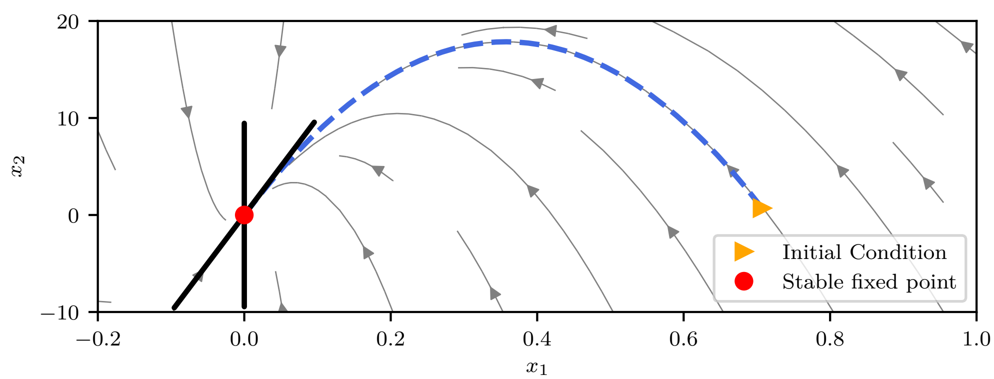

Fig. 3 The blue (dashed) line shows the trajectory stemming from a random unit-norm initial condition $\mathbf{x}_0$. The thick black lines depict the two linear manifolds of the fixed point. The diagonal one corresponds to $\lambda_1 = 1/100 - 1/Re$, while the vertical one is associated to $\lambda_2 = -2/Re$. In the present case, Re is set to 50, thus corresponding to a situation where the fixed point is linearly stable. Note that different scales are used for the horizontal and vertical axes.

[1] Note that the non-normality of $\boldsymbol{\mathcal{A}}$ also implies that its right and left eigenvectors are different. As will be discussed in §4, this observation may have large consequences in fluid dynamics, particularly when addressing the problems of optimal linear control and/or estimation of strongly non-parallel flows.

2.3.1 Optimal perturbation analysis

Having observed that a random initial condition can experience a relatively large transient growth of its energy over a short period of time even though the fixed point is stable, one may be interested in the worst case scenario, i.e. finding which initial condition $\mathbf{x}_0$ is amplified as much as possible before it eventually decays. Searching for such a perturbation is known as *optimal perturbation analysis* and can be addressed by two different methods:

- Optimization,
- Singular Value Decomposition (SVD).

Both approaches will be presented. Although it requires the introduction of additional mathematical concepts, the approach relying on optimization will be introduced first in §2.3.1 as it easier to grasp. The approach relying on singular value decomposition of the exponential propagrator $\boldsymbol{\mathcal{M}} = \exp\left(\boldsymbol{\mathcal{A}}t\right)$ will then be presented in §2.3.1.

Formulation as an optimization problem

The aim of optimal perturbation analysis is to find the unit-norm initial condition $\mathbf{x}_0$ that maximizes $\|\mathbf{x}(T)\|_2^2$, where T is known as the *time horizon*. Note that we here consider only the 2-norm of $\mathbf{x}(T)$ for the sake of simplicity, although one could formally optimize different norms, see [**?**] for examples from fluid dynamics. For a given time horizon T, such a problem can be formulated as the following constrained maximization problem

$$\begin{aligned} &\underset{\mathbf{x}_0}{\text{maximize}}\ \mathcal{J}\left(\mathbf{x}_0\right) \\ &\text{subject to}\ \dot{\mathbf{x}} = \boldsymbol{\mathcal{A}}\mathbf{x} \\ &\qquad\qquad \|\mathbf{x}_0\|_2^2 = 1, \end{aligned} \tag{13}$$

where $\mathcal{J}(\mathbf{x}_0) = \|\mathbf{x}(T)\|_2^2$ is known as the *objective function*. It must be emphasized that problem (13) is not formulated as a convex optimization problem[2]. As such, it may exhibit local maxima. Nonetheless, this constrained maximization problem can be recast into the following unconstrained maximization problem

[2] Formally, a convex optimization problem reads

$$\begin{aligned} &\underset{\mathbf{x}}{\text{minimize}}\ \mathcal{J}\left(\mathbf{x}\right) \\ &\text{subject to}\ g_i\left(\mathbf{x}\right) \leq 0,\ i = 1, \cdots, m \\ &\qquad\qquad h_i\left(\mathbf{x}\right) = 0,\ i = 1, \cdots, p, \end{aligned}$$

where the objective function $\mathcal{J}\left(\mathbf{x}\right)$ and the inequality constraints functions $g_i\left(\mathbf{x}\right)$ are convex. The conditions on the equality constraints functions $h_i\left(\mathbf{x}\right)$ are more restrictive as they need to be affine functions, i.e. of the form $h_i\left(\mathbf{x}\right) = \mathbf{a}_i^T\mathbf{x} + b_i$. See the book by Boyd & Vandenberghe [1] for extensive details about convex optimization.

$$\underset{\mathbf{x},\mathbf{v},\mu}{\text{maximize}} \ \mathcal{L}\left(\mathbf{x},\mathbf{v},\mu\right), \tag{14}$$

where

$$\mathcal{L}\left(\mathbf{x},\mathbf{v},\mu\right) = \mathcal{J}\left(\mathbf{x}_0\right) + \int_0^T \mathbf{v}^T\left(\dot{\mathbf{x}} - \boldsymbol{\mathcal{A}}\mathbf{x}\right)\mathrm{d}t + \mu\left(\|\mathbf{x}_0\|_2^2 - 1\right) \tag{15}$$

is known as the *augmented Lagrangian* function. The additional optimization variables $\mathbf{v}$ and μ appearing in the definition of the augmented Lagrangian $\mathcal{L}$ are called *Lagrange multipliers*. Solutions to problem (14) are identified by vanishing first variations of $\mathcal{L}$ with respect to our three optimization variables. The first variation of $\mathcal{L}$ with respect to $\mathbf{v}$ and μ are simply the constraints of our original problem (13). The first variation of $\mathcal{L}$ with respect to $\mathbf{x}$ on the other hand is given by

$$\mathbf{0} = \left[\nabla_{\mathbf{x}}\mathcal{J} + \mathbf{v}(T)\right]\cdot\delta\mathbf{x}(0) + \int_0^T \left[\dot{\mathbf{v}} - \boldsymbol{\mathcal{A}}^H\mathbf{v}\right]\cdot\delta\mathbf{x}\ \mathrm{dt} + \left[2\mu\mathbf{x}_0 - \mathbf{v}(0)\right]\cdot\delta\mathbf{x}(0). \tag{16}$$

Eq. (16) vanishes only if

$$\dot{\mathbf{v}} = \boldsymbol{\mathcal{A}}^H\mathbf{v} \ \text{ over } t \in (0,T), \tag{17}$$

and

$$\nabla_{\mathbf{x}}\mathcal{J} - \mathbf{v}(T) = 2\mu\mathbf{x}(0) - \mathbf{v}(0) = 0. \tag{18}$$

Note that Eq. (17) is known as the adjoint system of our original linear dynamical system. Maximizing $\mathcal{L}$ is then a problem of simultaneously satisfying (8), (17) and (18). This is in general iteratively by gradient-based algorithms such as gradient descent or the rotation-update gradient algorithm that will be described in §3. For more details about adjoint-based optimization, see [1, 2].

Formulation using SVD

3 Numerical methods

3.1 Krylov methods for for solving linear systems

3.2 Fixed points computation

3.2.1 Selective Frequency Damping

3.2.2 Newton-Krylov method

3.2.3 BoostConv

3.3 Modal stability analysis

3.3.1 Power Iteration method

3.3.2 Arnoldi decomposition

3.3.3 Krylov-Schur decomposition

3.4 Non-modal stability analysis

3.4.1 Optimal perturbation analysis

3.4.2 Resolvent analysis

4 Application to fluid dynamics

5 Conclusions and perspectives

References

1. S Boyd and L Vandenberghe. *Convex optimization*. Cambridge university press, 2004.
2. R R Kerswell, C C T Pringle, and A P Willis. An optimization approach for analysing nonlinear stability with transition to turbulence in fluids as an exemplar. *Reports on Progress in Physics*, 77(8):085901, 2014.
3. P J Schmid and L Brandt. Analysis of fluid systems: Stability, receptivity, sensitivity. Lecture notes from the FLOW-NORDITA summer school on advanced instability methods for complex flows, Stockholm, Sweden, 2013. *Applied Mechanics Reviews*, 66(2):024803, 2014.