An Introduction to Optimal Control

Carlo Cossu

An Introduction to Optimal Control

Lecture notes from the FLOW-NORDITA Summer School on Advanced Instability Methods for Complex Flows, Stockholm, Sweden, 2013

Carlo Cossu
Directeur de Recherche CNRS,
Institut de Mécanique des Fluides de Toulouse,
CNRS—Université de Toulouse,
Allée Camille Soula,
Toulouse 31400, France
e-mail: carlo.cossu@imft.fr

The goal of these lecture notes is to provide an informal introduction to the use of variational techniques for solving constrained optimization problems with equality constraints and full state information. The use of the Lagrangian augmented cost function and variational techniques by which the adjoint equation and the optimality condition are found are introduced by the use of examples starting from steady finite-dimensional problems to end with unsteady initial-boundary value problems. Gradient methods based on sensitivity and adjoint equation solutions are also mentioned.

1 Introduction

In many applications, one wants to drive a given system to a desired state or to a better performance by acting on design parameters or control variables. For instance, aeronautical engineers are interested in finding the wing shape for which a prescribed pressure distribution is obtained. They may also want to reduce the viscous drag of a wing or of a fin by delaying transition to turbulence. To this end, they can tailor an appropriate pressure distribution by changing the wing shape or, for a given shape, they can use blowing and suction at the wall to modify the boundary layer stability properties. The question here would be what is the best wing shape attaining the minimum drag? And to attain the most stable boundary layer that remains laminar up to the largest Reynolds number? Operational meteorologists run global numerical simulations of the atmosphere to produce weather forecasts. With time increasing, numerical solutions deviate from the observed weather because of their sensitive dependence on initial conditions. Weather prediction simulations therefore need to be regularly reinitialized, taking into account observations of the weather evolution coming from multiple data sources such as satellites and buoys. In this case, the question is what is the initial condition providing the numerical weather prediction that best matches the observations over a selected period of time? Similar problems exist in virtually all fields of theoretical and applied science, ranging from, e.g., economics to epidemiology.

All these very different problems often have a common mathematical structure. The system is identified by its *state*. In an incompressible viscous flow for instance the system state is given by the velocity and pressure fields in a prescribed spatial domain and temporal interval. In numerical simulations of the flow, these variables would be known only in a large, but finite, number of grid points, so that the state would be given by a finite dimensional vector. Similar considerations apply to weather forecasts. The *control* is the set of parameters or variables trough which we can act on the state. In the case of the drag minimization, these could be, e.g., the distribution of wall-normal velocity that is applied on the wing skin, or the function defining the wing shape or may be only a single scalar denoting an amplitude of wall suction or shape change with prescribed shape. In the case of data assimilation in numerical weather forecast, the control variable is the initial condition given to the numerical simulation. The state and the control variables are ruled by the *state equation* that can be used to compute the state for a given control. In the cited examples, the state equation is essentially given by the Navier–Stokes equations supplemented by appropriate initial and boundary conditions. The goal of acting on the control can be often stated as a minimization problem for the *cost or performance or objective function(al)*, that depends on the state and on the control of the system. In the wing example, the objective could, e.g., be the viscous total drag or the inverse of the critical Reynolds number where the boundary layer becomes linearly unstable penalized by a measure of the cost of the control, etc. (see, e.g., Ref. [1]). In the meteorological data assimilation example, the objective to be minimized is the mean deviation of the numerical predictions from the observational data (see, e.g., Ref. [2]).

The scope of these lecture notes is to provide an informal introduction to some techniques aimed at finding the optimal control that minimizes a cost function under the constraint given by the state equation. As the main scope here is to be pedagogical and concise, these techniques are mainly introduced through examples. Many good monographs already exist, as the ones in Refs. [3–7] among others, that are more complete and rigorous than the present notes. In the following, no proofs of existence or uniqueness of the optimal control solution are provided and it will be implicitly assumed that the necessary hypotheses are satisfied, ensuring that the solution exists and can be computed. These techniques are then extended to unsteady finite-dimensional problems in Sec. 4. In Sec. 5, after introducing the definitions of inner product and linear adjoint operator, the Lagrangian definition and the optimality system are reformulated in these more abstract terms. This allows to easily extend the introduced optimization techniques to space-dependent problems in Sec. 6. A brief introduction to feedback control in the case of linear evolution problems with quadratic cost functions is provided in Sec. 7. Some final comments are made in the last section, Sec. 8.

2 Preliminaries and Some Notation

In an (*unconstrained) optimization* problem the goal is to find the system *state* $\mathbf{q}$ (which is a vector or a function or a vector field etc.) that minimizes the real scalar *cost* or performance or objective $\mathcal{J}(\mathbf{q})$. The minimum of $\mathcal{J}$ can be attained on the boundary of the domain of definition of the state or in internal points where, assuming $\mathcal{J}$ continuous with continuous derivatives, the gradient $\partial\mathcal{J}/\partial\mathbf{q}$ necessarily vanishes[1].

Many methods are available for the numerical solution of unconstrained optimization problems and are usually classified into *gradient-based methods* that require the knowledge of $\partial\mathcal{J}/\partial\mathbf{q}$ to improve the solution and *gradient-free methods* that only require evaluations of $\mathcal{J}$. Probably the most intuitive and "natural" of gradient methods is the *steepest descent* where, given an initial guess $\mathbf{q}^{(0)}$, the guess is improved by following the direction of steepest descent in successive small steps of length λ: $\mathbf{q}^{(p+1)} = \mathbf{q}^{(p)} - \lambda^{(p)}(\partial\mathcal{J}/\partial\mathbf{q})^{(p)}$. In many situations, however, the

[1] In the following, the notation $\partial\mathcal{J}/\partial\mathbf{q}$ is used to denote the gradient of the function $\mathbf{q}\mapsto\mathcal{J}$. Similarly, $\partial\mathbf{q}/\partial\mathbf{g}$ is used to denote the Jacobian of the function $\mathbf{g}\mapsto\mathbf{q}$.

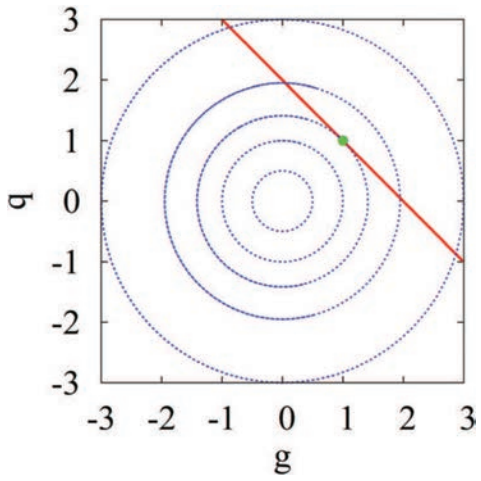

Fig. 1 Example of constrained optimization in the control-state plane. The level-sets of the cost function $\mathcal{J}(q,g)$ are the concentric circles (dotted line) with outward increasing values. Solutions satisfying the constraint $F(q,g)=0$ lie on the straight (solid) line. The constrained minimum (small filled circle) is attained in a point where the constraint curve is tangent to one of the level sets of $\mathcal{J}$.

convergence of this method can be quite slow if, e.g., the descent path enters long narrow "valleys" of slow slope (see, e.g., Ref. [8]). Other techniques, using, e.g., conjugate gradient methods, or Newton and quasi-Newton methods, are then available to improve the speed of convergence in those cases (again, see, e.g., Ref. [8]).

In unconstrained optimization, it is assumed that the state $\mathbf{q}$ can be directly changed in order to reach the optimal solution. In most situations, however, this is not possible and one can only act on a set of control variables $\mathbf{g}$. The state and the control satisfy the *state equation* $\mathbf{F}(\mathbf{q},\mathbf{g})=\mathbf{0}$ (state equation, evolution equation, etc.) that can be used to determine the state for given enforced control. The goal of *constrained optimization* is to minimize the cost $\mathcal{J}(\mathbf{q},\mathbf{g})$ by acting on $\mathbf{g}$ under the constraint $\mathbf{F}(\mathbf{q},\mathbf{g})=\mathbf{0}$. In the following, it is assumed that $\mathbf{F}$ and $\mathcal{J}$ and their derivatives are continuous. The solution of the constrained problem is usually very different from the solution of the unconstrained problem as seen from the example below.

Example: Consider the cost function $\mathcal{J}(q,g)=q^2+g^2$, where the state and the control variables are scalars defined on the whole real line. The unconstrained minimum of $\mathcal{J}$ is zero and is obtained in the origin of the $q-g$ plane. If the constraint $F(q,g)=q+g-2=0$ is enforced, then the unconstrained minimum cannot be reached. In the $q-g$ plane, the level sets of $\mathcal{J}$ are concentric circles with outward increasing values, while the solution curve of the constraint is a straight line. From Fig. 1, it is seen that the lowest level-set (circle) intersecting the straight line is the one tangent to the straight line. The optimal points can be found by, e.g., replacing the constraint in the cost function in order to have $\mathcal{J}[q(g),g]=2g^2-4g+4$. The minimum is found enforcing $D\mathcal{J}/Dg=4(g-1)=0$ and verifying the sign of the second order derivative, which gives the minimum constrained value $\mathcal{J}=2$ attained in $g=1$, $q=1$. Remark that in order to solve this problem we have explicitly solved $F(q,g)=0$ in terms of the state variable to get $q(g)$ that has then been replaced into $\mathcal{J}$ to compute the total derivative $D\mathcal{J}/Dg$. In general this explicit solution is not available (think about, e.g., $F=13q^7-4g^{11}+11qg^3-17=0$). In this case one must keep as variables both q and g in the minimization procedure. A possible solution technique (see Sec. 3.1 below) is then to compute the total derivative of $\mathcal{J}$ using the chain rule $D\mathcal{J}/Dg=(\partial\mathcal{J}/\partial q)(dq/dg)+(\partial\mathcal{J}/\partial g)$. The sensitivity $dq/dg=-(\partial F/\partial q)^{-1}\partial F/\partial g$ is found by differentiating $F=0$ which gives $dF=(\partial F/\partial q)dq+(\partial F/\partial g)dg=0$. The two equations $D\mathcal{J}/Dg=0$ and $F=0$ are then solved in the two unknowns g and q to look for local extremal points. For the present case $\partial F/\partial g=1$, $\partial F/\partial q=1$, $\partial\mathcal{J}/\partial g=2g$, $\partial\mathcal{J}/\partial q=2q$, the sensitivity is $dq/dg=-1$ and $D\mathcal{J}/Dg=2(g-q)$, which solved with $F=g+q-2=0$ gives the already found optimal $\mathcal{J}=2$, $g=1$, $q=1$. We'll see in the following that still other solution methods exist. ■

The choice made in these lecture notes is to introduce concepts on specific cases, starting with the simplest and then extending the ideas to more complex situations. The basic ideas and methods of constrained optimization will therefore be introduced in Sec. 3, dedicated to the optimization of a finite-dimensional system in the time-independent case. The introduced concepts will then be extended to time- and space-dependent cases.

3 Constrained Optimization in Finite-Dimensional Time-Independent Systems

Consider the constrained optimization problem in the case where both the state and the control variables do not depend on time. Given the state vector $\mathbf{q}\in\mathbb{R}^N$ and the control vector $\mathbf{g}\in\mathbb{R}^K$, we wish to minimize the cost function $\mathcal{J}(\mathbf{q},\mathbf{g})$ subject to the state equation $\mathbf{F}(\mathbf{q},\mathbf{g})=\mathbf{0}$ where $\mathbf{F}\in\mathbb{R}^N$. Different approaches to solve constrained optimization problems are introduced below for this specific type of problems.

3.1 Gradient Methods Based on Sensitivity. Iterative gradient methods can be used to solve constrained optimization problem. At the p-th iteration an approximation of the optimal control $\mathbf{g}$ is available. In order to compute the cost $\mathcal{J}$, the state equation must be solved and $\mathbf{q}$ explicitly obtained[2]. An improved value for $\mathbf{g}$ is then obtained, at the $(p+1)$th iteration making use of the *total derivative*[3] $D\mathcal{J}/D\mathbf{g}$, i.e., the K derivatives $D\mathcal{J}/Dg_k$ with $k=1,\ldots,K$, where, denoting $\mathbf{e}_k$, the k-th vector of an orthonormal basis $\mathbf{g}=\sum_k g_k\mathbf{e}_k$. The straightforward method to compute $D\mathcal{J}/D\mathbf{g}$ is to use finite difference approximations obtained incrementing each control variable separately by a small increment Δg_k (with $k=1,2,\ldots,K$).

$$\frac{D\mathcal{J}}{Dg_k}\approx\frac{\mathcal{J}[\mathbf{q}(\mathbf{g}+\Delta g_k\mathbf{e}_k),\mathbf{g}+\Delta g_k\mathbf{e}_k]-\mathcal{J}[\mathbf{q}(\mathbf{g}),\mathbf{g}]}{\Delta g_k} \tag{1}$$

The use of Eq. (1) requires to solve the state equation for each Δg_k to get the corresponding $\mathbf{q}(\mathbf{g}+\Delta g_k\mathbf{e}_k)$. The computation of $D\mathcal{J}/D\mathbf{g}$ using this straightforward but "naïve" approach therefore requires $K+1$ solutions of the state equation. This is, however, a problem because the most expensive part of the computations usually is the solution of the state equation which, in general, is nonlinear. Each solution, if solved, e.g., using a Newton method, would require a finite number of inversions of the Jacobian matrix $\partial\mathbf{F}/\partial\mathbf{q}$ recomputed at each Newton iteration. In those cases, one might want to numerically compute $D\mathcal{J}/D\mathbf{g}$ using alternative methods with a reduced number of solutions of the state equation. In order to do that, as a first step, the chain rule is used to express the total derivative of the cost function

$$\frac{D\mathcal{J}}{Dg_k}=\frac{\partial\mathcal{J}}{\partial\mathbf{q}}\cdot\frac{d\mathbf{q}}{dg_k}+\frac{\partial\mathcal{J}}{\partial g_k} \tag{2}$$

[2]We consider a solution strategy where it is required that the state equation is exactly satisfied, not only by the optimal solution, but also during any iterate needed to reach it. However, the count is different if other strategies are used where it is not required that $\mathbf{F}=\mathbf{0}$ during iterations.

[3]The derivative of the cost function $\mathcal{J}(\mathbf{q},\mathbf{g})$ with respect to $\mathbf{q}$ and $\mathbf{g}$ considered as independent variables will be denoted by $\partial\mathcal{J}/\partial\mathbf{q}$ and $\partial\mathcal{J}/\partial\mathbf{g}$, respectively. However, the cost function can also be considered as a composed function of $\mathbf{g}$ alone $\mathcal{J}[\mathbf{q}(\mathbf{g}),\mathbf{g}]$, where $\mathbf{q}(\mathbf{g})$ is obtained from $\mathbf{F}(\mathbf{q},\mathbf{g})=\mathbf{0}$. In this case, the derivative of the cost function with respect to $\mathbf{g}$ is labeled total derivative $D\mathcal{J}/D\mathbf{g}$ and can be obtained with the chain rule as $D\mathcal{J}/D\mathbf{g}=(\partial\mathcal{J}/\partial\mathbf{q})\cdot(d\mathbf{q}/d\mathbf{g})+(\partial\mathcal{J}/\partial\mathbf{g})$

where usually the derivatives $\partial\mathcal{J}/\partial\mathbf{q}$ and $\partial\mathcal{J}/\partial g_k$ are explicitly known. If finite differences are used to compute $d\mathbf{q}/dg_k$, $K+1$ solutions of the state equation would again be necessary. Instead, one can differentiate the state equation. As $\mathbf{F}=\mathbf{0}$ for all solutions, then $D\mathbf{F}/D\mathbf{g}=\mathbf{0}$ and therefore,

$$\frac{\partial\mathbf{F}}{\partial\mathbf{q}}\frac{d\mathbf{q}}{dg_k}+\frac{\partial\mathbf{F}}{\partial g_k}=0 \tag{3}$$

The above is the *sensitivity equation*. The *sensitivities* $d\mathbf{q}/dg_k$ can be found solving K *linear systems* all needing the inversion of the same Jacobian matrix computed for the same state.

$$\frac{d\mathbf{q}}{dg_k}=-\left(\frac{\partial\mathbf{F}}{\partial\mathbf{q}}\right)^{-1}\frac{\partial\mathbf{F}}{\partial g_k} \tag{4}$$

The computation of $D\mathcal{J}/D\mathbf{g}$ based on sensitivity therefore requires one solution of the state equation plus K solutions of $N\times N$ linear systems all with the same linear operator to be inverted. In some situations, this may be more convenient than solving K nonlinear N-dimensional systems.

3.2 Variational Formulation. Consider the simple example already discussed in Sec. 2 where $\mathcal{J}(q,g)=q^2+g^2$ and $F(q,g)=g+q-2=0$. The goal of the constrained optimization is to descend as low as possible on the $\mathcal{J}$ level curves while remaining on the path given by $F=0$. If both $\mathcal{J}$ and F are continuous with continuous derivatives, at the point where the minimum is reached, the path is tangent[4] to the level curve of the optimal $\mathcal{J}$ as clearly seen in Fig. 1. This implies that in the optimal point, the gradient of the cost function and the gradient of F are parallel, i.e., by components in the $q-g$ plane,

$$\left\{\begin{array}{c}\partial\mathcal{J}/\partial g\\ \partial\mathcal{J}/\partial q\end{array}\right\}=a\left\{\begin{array}{c}\partial F/\partial g\\ \partial F/\partial q\end{array}\right\}$$

The optimality system therefore corresponds to the above conditions and the "stay-on-the-path" condition $F=0$, which gives a system of three equations in the three unknowns q, g and a

$$\frac{\partial\mathcal{J}}{\partial g}-a\frac{\partial F}{\partial g}=0 \tag{5}$$

$$\frac{\partial\mathcal{J}}{\partial q}-a\frac{\partial F}{\partial q}=0 \tag{6}$$

$$F=0 \tag{7}$$

For reasons that will become clear in the following, the first condition takes the name of *optimality condition*, while the second is the *adjoint equation*.

Lagrange remarked that an "augmented" cost function $\mathcal{L}=\mathcal{J}-aF$ can be defined such that the optimality system of the constrained problem coincides with the optimality system of an unconstrained problem defined for $\mathcal{L}(q,g,a)$ (where, q, g, and a must be considered as independent). The function $\mathcal{L}$ is the *Lagrangian* or *augmented cost function* and a is the *Lagrange multiplier* or *costate*.

Example: For the considered simple example $\partial\mathcal{J}/\partial q=2q$, $\partial\mathcal{J}/\partial g=2g$, $\partial F/\partial q=1$, $\partial F/\partial g=1$ and therefore the conditions in Eqs. (5)–(7) read: $2g-a=0$, $2q-a=0$, $q+g-2=0$ that admits the solution $(g,q,a)=(1,1,2)$ for which the minimum $\mathcal{J}=2$ is attained. This is the same solution found in Sec. 2. ■

Lagrange's approach can be generalized to the case of an N-dimensional state vector and a K-dimensional control vector. In that case there are N scalar components of the state equation that have to be each one multiplied by the corresponding Lagrange multiplier resulting in a term of the type a_iF_i (where Einstein convention of implicit summation is used), that can be rewritten[5] as $\mathbf{a}\cdot\mathbf{F}$ where the N-dimensional costate (or Lagrange multiplier) vector $\mathbf{a}$ is introduced.

$$\mathcal{L}(\mathbf{q},\mathbf{g},\mathbf{a})=\mathcal{J}(\mathbf{q},\mathbf{g})-\mathbf{a}\cdot\mathbf{F}(\mathbf{q},\mathbf{g}) \tag{8}$$

Extremality conditions are enforced on $\mathcal{L}$ by considering $\mathbf{q}$, $\mathbf{g}$, and $\mathbf{a}$ as independent variables.

$$\frac{\partial\mathcal{L}}{\partial\mathbf{q}}=\mathbf{0},\quad \frac{\partial\mathcal{L}}{\partial\mathbf{g}}=\mathbf{0},\quad \frac{\partial\mathcal{L}}{\partial\mathbf{a}}=\mathbf{0} \tag{9}$$

Using the Lagrangian definition, the above conditions can be made explicit

$$\frac{\partial\mathcal{L}}{\partial q_i}=\frac{\partial\mathcal{J}}{\partial q_i}-a_j\frac{\partial F_j}{\partial q_i}=0\quad(i=1,\ldots,N) \tag{10}$$

$$\frac{\partial\mathcal{L}}{\partial g_k}=\frac{\partial\mathcal{J}}{\partial g_k}-a_j\frac{\partial F_j}{\partial g_k}=0\quad(k=1,\ldots,K) \tag{11}$$

$$\frac{\partial\mathcal{L}}{\partial a_i}=-F_i=0\quad(i=1,\ldots,N) \tag{12}$$

which in vector form, denoting by ^{T} the transpose, gives

$$\frac{\partial\mathcal{L}}{\partial\mathbf{q}}=\mathbf{0}\Rightarrow\left(\frac{\partial\mathbf{F}}{\partial\mathbf{q}}\right)^T\mathbf{a}=\frac{\partial\mathcal{J}}{\partial\mathbf{q}}\quad\text{(adjoint equation)} \tag{13}$$

$$\frac{\partial\mathcal{L}}{\partial\mathbf{g}}=\mathbf{0}\Rightarrow\left(\frac{\partial\mathbf{F}}{\partial\mathbf{g}}\right)^T\mathbf{a}=\frac{\partial\mathcal{J}}{\partial\mathbf{g}}\quad\text{(optimality condition)} \tag{14}$$

$$\frac{\partial\mathcal{L}}{\partial\mathbf{a}}=\mathbf{0}\Rightarrow\mathbf{F}=\mathbf{0}\quad\text{(state equation)} \tag{15}$$

The first two vector equations are, respectively, the *adjoint equation* and the *optimality condition*, while the last is the usual state equation.

The optimality system in Eqs. (13)–(15) can also be derived using a variational approach that has the advantage to be directly generalized to the case where the state and/or the control are functions. In particular, it is required that first order variations $\delta\mathcal{L}$ induced by small variations $\delta\mathbf{q}$, $\delta\mathbf{g}$, or $\delta\mathbf{a}$ are zero, i.e., $\delta\mathcal{L}/\delta\mathbf{q}=\mathbf{0}$, $\delta\mathcal{L}/\delta\mathbf{g}=\mathbf{0}$ $\delta\mathcal{L}/\delta\mathbf{a}=\mathbf{0}$. Now, the variation $\delta\mathcal{L}$ induced by, e.g., a small variation $\delta\mathbf{a}=\varepsilon\tilde{\mathbf{a}}$ of the adjoint is, at first order, given by the directional derivative

$$\frac{\partial\mathcal{L}}{\partial\mathbf{a}}\tilde{\mathbf{a}}:=\lim_{\varepsilon\to0}\frac{\mathcal{L}(\mathbf{q},\mathbf{g},\mathbf{a}+\varepsilon\tilde{\mathbf{a}})-\mathcal{L}(\mathbf{q},\mathbf{g},\mathbf{a})}{\varepsilon}$$

which, using the Lagrangian definition in Eq. (8), gives $(\partial\mathcal{L}/\partial\mathbf{a})\tilde{\mathbf{a}}=-\tilde{\mathbf{a}}\cdot\mathbf{F}$. Enforcing that this variation is zero $\forall\tilde{\mathbf{a}}$ requires that $\mathbf{F}=\mathbf{0}$, i.e., that Eq. (15) is satisfied. Setting to zero the variation of $\mathcal{L}$ with respect to the state gives the adjoint Eq. (13)

[4] This condition of course does not apply when the constrained and the unconstrained optimals coincide. In that case, the path goes through the unconstrained optimal. Formally, one can think that in this case the path is tangent to the level curve of zero diameter.

[5] Using the usual dot product notation $\mathbf{a}\cdot\mathbf{F}=a_iF_i$ for the inner product of vectors $\langle\mathbf{a},\mathbf{F}\rangle$.

$$\frac{\partial \mathcal{L}}{\partial \mathbf{q}}\tilde{\mathbf{q}} := \lim_{\varepsilon\to 0}\frac{\mathcal{L}(\mathbf{q}+\varepsilon\tilde{\mathbf{q}},\mathbf{g},\mathbf{a})-\mathcal{L}(\mathbf{q},\mathbf{g},\mathbf{a})}{\varepsilon}$$
$$= \frac{\partial \mathcal{J}}{\partial \mathbf{q}}\cdot\tilde{\mathbf{q}} - \mathbf{a}\cdot\left[\frac{\partial \mathbf{F}}{\partial \mathbf{q}}\right]\tilde{\mathbf{q}} = \left[\frac{\partial \mathcal{J}}{\partial \mathbf{q}} - \left(\frac{\partial \mathbf{F}}{\partial \mathbf{q}}\right)^T\mathbf{a}\right]\cdot\tilde{\mathbf{q}} = 0$$
$$\forall\tilde{\mathbf{q}} \Rightarrow \frac{\partial \mathcal{J}}{\partial \mathbf{q}} - \left[\frac{\partial \mathbf{F}}{\partial \mathbf{q}}\right]^T\mathbf{a} = \mathbf{0}$$

while setting to zero the variation of $\mathcal{L}$ with respect to the control enforces the optimality condition Eq. (14)

$$\frac{\partial \mathcal{L}}{\partial \mathbf{g}}\tilde{\mathbf{g}} := \lim_{\varepsilon\to 0}\frac{\mathcal{L}(\mathbf{q},\mathbf{g}+\varepsilon\tilde{\mathbf{g}},\mathbf{a})-\mathcal{L}(\mathbf{q},\mathbf{g},\mathbf{a})}{\varepsilon}$$
$$= \frac{\partial \mathcal{J}}{\partial \mathbf{g}}\cdot\tilde{\mathbf{g}} - \mathbf{a}\cdot\left[\frac{\partial \mathbf{F}}{\partial \mathbf{g}}\right]\tilde{\mathbf{g}} = \left[\frac{\partial \mathcal{J}}{\partial \mathbf{g}} - \left(\frac{\partial \mathbf{F}}{\partial \mathbf{g}}\right)^T\mathbf{a}\right]\cdot\tilde{\mathbf{g}} = 0$$
$$\forall\tilde{\mathbf{g}} \Rightarrow \frac{\partial \mathcal{J}}{\partial \mathbf{g}} - \left[\frac{\partial \mathbf{F}}{\partial \mathbf{g}}\right]^T\mathbf{a} = \mathbf{0}$$

The optimality system in Eqs. (13)–(15) is of dimension $2N+K$ and is usually nonlinear. The direct solution of the full system, the so called *one-shot method*, is quite expensive in the case of large N and/or K. Alternative methods to find a solution are therefore discussed below.

3.3 Gradient Methods Based on the Adjoint. Let us now briefly reconsider the gradient-based iterative methods discussed in Sec. 2 to show that the use of the adjoint equation greatly reduces the number of operations needed to compute $D\mathcal{J}/D\mathbf{g}$. This reduction is at the core of so-called *adjoint methods* for constrained optimization.

The adjoint equation, Eq. (13), gives an expression for the derivative of the state equations with respect to the state vector $\partial\mathcal{J}/\partial q_i = (\partial F_j/\partial q_i)a_j$ that can be replaced in the expression of the total derivative of the cost (see Eq. (2)) $D\mathcal{J}/Dg_k = (\partial\mathcal{J}/\partial q_i)(dq_i/dg_k) + \partial\mathcal{J}/\partial g_k$ to get

$$\frac{D\mathcal{J}}{Dg_k} = a_j\frac{\partial F_j}{\partial q_i}\frac{dq_i}{dg_k} + \frac{\partial \mathcal{J}}{\partial g_k}$$

Replacing Eq. (3) $(\partial F_j/\partial q_i)(d\mathbf{q}_i/dg_k) = -(\partial F_j/\partial g_k)$ in the above equation gives: $D\mathcal{J}/Dg_k = -a_j(\partial F_j/\partial g_k) + \partial\mathcal{J}/\partial g_k$ that, rewritten in vector form with the adjoint equation, gives the equations of the adjoint method

$$\left[\frac{\partial \mathbf{F}}{\partial \mathbf{q}}\right]^T\mathbf{a} = \frac{\partial \mathcal{J}}{\partial \mathbf{q}} \tag{16}$$

$$\frac{D\mathcal{J}}{D\mathbf{g}} = -\left[\frac{\partial \mathbf{F}}{\partial \mathbf{g}}\right]^T\mathbf{a} + \frac{\partial \mathcal{J}}{\partial \mathbf{g}} \tag{17}$$

where usually the explicit expressions of $\partial\mathcal{J}/\partial\mathbf{q}$, $\partial\mathcal{J}/\partial\mathbf{g}$ are available. The computation of the gradient using the above expressions requires to compute $(\partial\mathbf{F}/\partial\mathbf{q})^T$ and $(\partial\mathbf{F}/\partial\mathbf{g})$ either analytically or, e.g., by finite differences, one solution of the state equation and the solution of the adjoint equation, which is a $N\times N$ linear system. This operation count makes the method of the adjoint superior to sensitivity-based methods where K linear systems solutions were needed. We will see in the following that this superiority becomes a matter of debate for time-dependent problems.

3.4 An Example: Optimal Amplification of Forcing Supported by a Linear System. Consider the problem of a (steady) N-dimensional linear system subject to a (steady) forcing $\mathbf{Lu}+\mathbf{f}=\mathbf{0}$ where we look for the $\mathbf{f}$ that maximizes the energy amplification $R = \mathbf{u}\cdot\mathbf{u}/\mathbf{f}\cdot\mathbf{f}$.

In this problem, the control variable is the forcing while the state of the system is the response, i.e., $\mathbf{g}=\mathbf{f}$ and $\mathbf{q}=\mathbf{u}$. The dimension of the control K is therefore equal to the dimension of the state N. One can transform the maximization problem into a minimization problem by minimizing the inverse of R, i.e., the ratio of the norm of the forcing to the norm of the response.[6] The problem is formalized in the following way:

$$\mathbf{F}(\mathbf{q},\mathbf{g}) = \mathbf{Lq}+\mathbf{g} = \mathbf{0} \tag{18}$$
$$\mathcal{J}(\mathbf{q},\mathbf{g}) = \mathbf{g}\cdot\mathbf{g}/\mathbf{q}\cdot\mathbf{q} \tag{19}$$

The goal is to reach a local minimum of $\mathcal{J}$ acting on $\mathbf{g}$. For the present problem $\partial\mathbf{F}/\partial\mathbf{q}=\mathbf{L}$, $\partial\mathbf{F}/\partial\mathbf{g}=\mathbf{I}$ (the identity matrix), $\partial\mathcal{J}/\partial\mathbf{q} = -2\mathbf{q}(\mathbf{g}\cdot\mathbf{g})/(\mathbf{q}\cdot\mathbf{q})^2$ and $\partial\mathcal{J}/\partial\mathbf{g} = 2\mathbf{g}/(\mathbf{q}\cdot\mathbf{q})$, so that the explicit optimality system in Eqs. (13)–(15) reads

$$[\partial\mathbf{F}/\partial\mathbf{q}]^T\mathbf{a} = \partial\mathcal{J}/\partial\mathbf{q} \Rightarrow \mathbf{L}^T\mathbf{a} = -2\mathbf{q}(\mathbf{g}\cdot\mathbf{g})/(\mathbf{q}\cdot\mathbf{q})^2 \tag{20}$$
$$[\partial\mathbf{F}/\partial\mathbf{g}]^T\mathbf{a} = \partial\mathcal{J}/\partial\mathbf{g} \Rightarrow \mathbf{g} = \mathbf{a}(\mathbf{q}\cdot\mathbf{q})/2 \tag{21}$$
$$\mathbf{F}=\mathbf{0} \Rightarrow \mathbf{Lq}+\mathbf{g}=\mathbf{0} \tag{22}$$

The full optimality system is of dimension $2N+K$ (in this specific case $3N$ as $K=N$). Even if the state equation is linear, the full system is not because of the cost-function. A "one-shot" method, i.e., the direct solution of the full optimality system taken as a whole will probably be based on modified Newton iterations, that would require the solution of $3N\times 3N$ linear systems at each iteration. This is clearly not convenient and other methods must be used, such as, e.g., gradient-based methods where the minimum of $\mathcal{J}$ is sought using information on the total gradient $D\mathcal{J}/D\mathbf{g}$ or iterative methods where the equations of the optimality system are solved sequentially and not simultaneously.

Exercise: Consider solving the above problem using a gradient-based algorithm. Detail the procedure to follow if the gradient method is based on sensitivities and the procedure for the adjoint-based gradient method. ■

Example: The considered problem can also be solved using an iterative method where the three equations (Eqs. (20)–(22)) are solved sequentially. The method is initialized by giving an initial guess on $\mathbf{g}$, then the loop is:

(1) Given the p-th guess for the optimal forcing $\mathbf{g}^{(p)}$, compute the corresponding response $\mathbf{q}^{(p)}$ solving the state equation $\mathbf{Lq}^{(p)} = -\mathbf{g}^{(p)}$.
(2) Compute $\mathcal{J}$ and its relative increment wrt the previous iteration. If convergence is reached stop, if not continue.
(3) Compute the adjoint state $\mathbf{a}^{(p)}$ solving the adjoint equation.
(4) Finally update the control vector and get $\mathbf{g}^{(p+1)}$ using the optimality condition and go to No. 1.

As an application, we use this iterative method to compute the optimal forcing energy amplification, R, supported by the system defined by the linear operator

$$\mathbf{L} = \begin{bmatrix} -1/\mathrm{Re} & 0 \\ 1 & -3/\mathrm{Re}\end{bmatrix}$$

for a Reynolds number $\mathrm{Re}=40$ using the sample programs reported below. From Table 1, it is seen how the iterations converge very fast to the optimal solution. The optimal forcing is

[6] The optimization procedure considered here should be considered as purely pedagogical. Using the standard definition of the norm of linear operators $R = \| -\mathbf{L}^{-1} \|^2$ and the considered problem is therefore one of computation of vector-induced L_2 matrix norm. For matrix norm computations, other highly efficient algorithms already exists, that are, e.g., coded into the `norm()` functions in MATLAB, OCTAVE, and SCILAB and are used as comparison in the exercises below.

Table 1 Convergence history of the optimal forcing example

It.	R	g_1	g_2
1	12,106	0.63127	0.77556
2	28,622	0.99969	0.024857
3	28,622	0.99969	0.024853

found to be almost parallel to the vector of components $\begin{Bmatrix} 1 \\ 0 \end{Bmatrix}$, while the optimal response is almost parallel to the vector of components $\begin{Bmatrix} 1 \\ 0 \end{Bmatrix}$. An OCTAVE–MATLAB program for this problem is given below.

```
% Define system
  Rey=40.0
  L=[ -1.0/Rey 0; 1-3.0/Rey]
% Exact solution using the norm function
  R_exact= (norm(inv(-L)))^2
% Define tolerance and initialize iterations
  tol=10^(-8);
  g=[ rand; rand] ; % (random initial guess)
  g=g/norm(g); % (normalize)
  J=10^23; dJrel=10^23; it=0;
% Iteration loop
  while (dJrel>tol)
  it=it+1; Jold=J;
  q=-inv(L)*g; % (solve state equation)
  g2=g'*g; q2=q'*q;
  J=g2/q2; % (objective function)
  dJrel=abs((J-Jold)/J);
  a=-2*(inv(L')*q)*(g2)/q2^2;%(solve adjoint equation)
  g=a*q2/2.0; % (enforce optimality eq.)
  g=g/norm(g); % (normalize)
  end % (end of iteration loop)
% optimal amplification
  R=1.0/J;
% print results
  it, R % (final iteration and amplification)
  g % (optimal forcing (defined up to a constant)
  q % (optimal response)
```

The same program translated in SCILAB is:

```
// Define system
  Rey=40.0
  L=[ -1.0/Rey 0; 1-3.0/Rey]
// Exact solution using the norm function
  R_exact= (norm(inv(-L)))^2
// Define tolerance and initialize iterations
  tol=10^(-8);
  g=[ rand(); rand()] ;
  g=g/norm(g); // (normalize)
  J=10^23; dJrel=10^23; it=0;
// Iteration loop
  while (dJrel>tol)
    it=it+1; Jold=J;
    q=-inv(L)*g; // (solve state equation)
    g2=g'*g; q2=q'*q;
    J=g2/q2; // (objective function)
    dJrel=abs((J-Jold)/J);
    a=-2*(inv(L')*q)*(g2)/q2^2;// (solve adjoint equation)
    g=a*q2/2.0; // (enforce optimality eq.)
    g=g/norm(g); // (normalize)
  end // (end of iteration loop)
// optimal amplification
  R=1.0/J;
// print results
  it, R // (final iteration and amplification)
  g // (optimal forcing (defined up to a constant)
  q // (optimal response)
```

■

Exercise: Repeat the analysis of the above example for values of Re ranging from Re = 1 to Re = 1000. Can you find an asymptotic scaling for $R(\mathrm{Re})$? ■

Exercise: Implement a steepest descent method for the example above where the gradient is computed using the adjoint method. Compare the results to the ones obtained above. ■

4 Variational Formulation of Constrained Optimization for Finite-Dimensional Time-Dependent Problems (ODEs)

The variational formulation based on the Lagrangian, discussed in the previous section for steady problems, can be extended without major difficulties to time-dependent problems. Consider an initial value problem (IVP) for a system of ordinary differential equations (ODEs) $d\mathbf{q}/dt = \mathbf{N}(\mathbf{q}, \mathbf{g}, t)$ defined for $t \in [0, T]$ with initial condition $\mathbf{q}(0) = \mathbf{q}_0$. In this case, there are two types of constraints: the one that enforces the evolution equation $\mathbf{F}(\mathbf{q}, \mathbf{g}, t) = d\mathbf{q}/dt - \mathbf{N}(\mathbf{q}, \mathbf{g}, t) = 0$ and the one that enforces the initial condition $\mathbf{F}_0(\mathbf{q}, \mathbf{q}_0) = \mathbf{q}(0) - \mathbf{q}_0 = \mathbf{0}$. The Lagrangian appropriate to the present case is

$$\mathcal{L} = \mathcal{J} - \int_0^T (\mathbf{a} \cdot \mathbf{F}) dt - \mathbf{b} \cdot \mathbf{F}_0 \tag{23}$$

where the costate $\mathbf{a}(t)$, which is a function of time and uses an inner product that includes the time variable, ensures that the evolution equation is respected $\forall t \in [0, T]$. The second Lagrange multiplier $\mathbf{b}$, which is not a function of time and uses the inner product without the time variable, ensures that the initial condition is enforced at $t = 0$. The optimality system is, as usual, obtained by setting to zero the variation of the Lagrangian with respect to all variables considered as independent. $\delta\mathcal{L}/\delta\mathbf{a} = \mathbf{0}$ and $\delta\mathcal{L}/\delta\mathbf{b} = \mathbf{0}$ will enforce the evolution equation and the initial condition respectively, while the adjoint equation will be derived from $\delta\mathcal{L}/\delta\mathbf{q} = \mathbf{0}$ and the optimality condition from $\delta\mathcal{L}/\delta\mathbf{g} = \mathbf{0}$. The meaning of these conditions will be illustrated in the following examples.

Different problem types can be defined depending on the nature of the control (continuous forcing, initial condition, etc.) and the choice of the cost function. Below, two simple examples are considered: (a) The computation of the maximum temporal energy amplification of a linear system, where the control variable is the initial condition. (b) The computation of the optimal continuous forcing, applied in $[0, T]$, needed to drive the final state $\mathbf{q}(T)$ near to the desired target $\mathbf{p}$.

4.1 Optimal Temporal Energy Growth. Consider an unforced evolution equation $d\mathbf{q}/dt = \mathbf{N}(\mathbf{q})$ and the optimal temporal energy amplification $G(T)$ which is the ratio of the "output" energy [the energy of the response at the final time $\mathbf{q}(T)$] to the input energy (the energy of the initial condition $\mathbf{q}_0$). The control variable is the initial condition $\mathbf{g} = \mathbf{q}_0$. One can transform the maximization problem into a minimization problem by minimizing the inverse of G. The optimization problem can therefore be formalized in the following way:

$$\mathbf{F} = \frac{d\mathbf{q}}{dt} - \mathbf{N}(\mathbf{q}) = \mathbf{0} \tag{24}$$

$$\mathbf{F}_0 = \mathbf{q}(0) - \mathbf{g} = \mathbf{0} \tag{25}$$

$$\mathcal{J}(\mathbf{q}, \mathbf{g}) = (\mathbf{g} \cdot \mathbf{g})/(\mathbf{q}(T) \cdot \mathbf{q}(T)) \tag{26}$$

where the goal is to reach a local minimum of $\mathcal{J}$ acting on the initial condition $\mathbf{g}$. Specializing the Lagrangian definition to this specific case (see, e.g., Ref. [9])

$$\mathcal{L}(\mathbf{q},\mathbf{g},\mathbf{a},\mathbf{b}) = \mathcal{J}(\mathbf{q},\mathbf{g}) - \int_0^T \mathbf{a}\cdot\left[\frac{d\mathbf{q}}{dt} - \mathbf{N}(\mathbf{q})\right]dt - \mathbf{b}\cdot[\mathbf{q}(0)-\mathbf{g}] \tag{27}$$

Optimality conditions are obtained by setting to zero the variation of the Lagrangian with respect to $\mathbf{q}$, $\mathbf{g}$, $\mathbf{a}$, and $\mathbf{b}$. Setting to zero the variation wrt the costate (the "field" Lagrange multiplier) requires that

$$\frac{\partial\mathcal{L}}{\partial\mathbf{a}}\tilde{\mathbf{a}} = \lim_{\varepsilon\to 0}\frac{\mathcal{L}(\mathbf{q},\mathbf{g},\mathbf{a}+\varepsilon\tilde{\mathbf{a}},\mathbf{b}) - \mathcal{L}(\mathbf{q},\mathbf{g},\mathbf{a},\mathbf{b})}{\varepsilon} = 0, \quad \forall\tilde{\mathbf{a}}$$

which, using Eq. (23), enforces that

$$\int_0^T \tilde{\mathbf{a}}\cdot\left[\frac{d\mathbf{q}}{dt} - \mathbf{N}(\mathbf{q})\right]dt = 0$$

which is satisfied $\forall\tilde{\mathbf{a}}(t)$ only if the evolution equation $d\mathbf{q}/dt - \mathbf{N}(\mathbf{q}) = \mathbf{0}$ is satisfied $\forall t\in[0,T]$. The condition $\delta\mathcal{L}/\delta\mathbf{b} = \mathbf{0}$ similarly leads to $\tilde{\mathbf{b}}\cdot[\mathbf{q}(0)-\mathbf{g}] = 0$, which is satisfied $\forall\tilde{\mathbf{b}}$ only if $\mathbf{q}(0) = \mathbf{g}$. The condition $\delta\mathcal{L}/\delta\mathbf{q} = \mathbf{0}$ requires that

$$\lim_{\varepsilon\to 0}\frac{\mathcal{L}(\mathbf{q}+\varepsilon\tilde{\mathbf{q}},\mathbf{g},\mathbf{a},\mathbf{b}) - \mathcal{L}(\mathbf{q},\mathbf{g},\mathbf{a},\mathbf{b})}{\varepsilon} = 0, \quad \forall\tilde{\mathbf{q}}$$

that it is easily found to correspond to

$$\frac{\partial\mathcal{J}}{\partial\mathbf{q}(T)}\cdot\tilde{\mathbf{q}}(T) - \int_0^T \mathbf{a}\cdot\left[\frac{d\tilde{\mathbf{q}}}{dt} - \frac{\partial\mathbf{N}}{\partial\mathbf{q}}\tilde{\mathbf{q}}\right]dt - \mathbf{b}\cdot\tilde{\mathbf{q}}(0) = 0$$

that must be satisfied $\forall\tilde{\mathbf{q}}$. Integrating by parts the first term of the integrand and rearranging the linear operator product under integration

$$\left[\frac{\partial\mathcal{J}}{\partial\mathbf{q}(T)}\right]\cdot\tilde{\mathbf{q}}(T) - [\mathbf{a}(T)\cdot\tilde{\mathbf{q}}(T) - \mathbf{a}(0)\cdot\tilde{\mathbf{q}}(0)] - \int_0^T\left[-\frac{d\mathbf{a}}{dt} - \left(\frac{\partial\mathbf{N}}{\partial\mathbf{q}}\right)^T\mathbf{a}\right]\cdot\tilde{\mathbf{q}}dt - \mathbf{b}\cdot\tilde{\mathbf{q}}(0) = 0$$

As $\partial\mathcal{J}/\partial\mathbf{q}(T) = -2\mathbf{q}(T)(\mathbf{g}\cdot\mathbf{g})/(\mathbf{q}(T)\cdot\mathbf{q}(T))^2$, it is therefore finally found that

$$\left[-\mathbf{a}(T) - 2\mathbf{q}(T)\frac{\mathbf{g}\cdot\mathbf{g}}{(\mathbf{q}(T)\cdot\mathbf{q}(T))^2}\right]\cdot\tilde{\mathbf{q}}(T) + [\mathbf{a}(0)-\mathbf{b}]\cdot\tilde{\mathbf{q}}(0) - \int_0^T\left[-\frac{d\mathbf{a}}{dt} - \left(\frac{\partial\mathbf{N}}{\partial\mathbf{q}}\right)^T\mathbf{a}\right]\cdot\tilde{\mathbf{q}}dt$$

which is satisfied $\forall\tilde{\mathbf{q}}$ only if (a) the adjoint evolution equation is satisfied $-d\mathbf{a}/dt = (\partial\mathbf{N}/\partial\mathbf{q})^T\mathbf{a}$ for $t\in[0,T]$, (b) the initial condition (for backward in time integration) on the adjoint equation is given as $\mathbf{a}(T) = -2\mathbf{q}(T)(\mathbf{g}\cdot\mathbf{g})/(\mathbf{q}(T)\cdot\mathbf{q}(T))^2$, and (c) $\mathbf{b}$ is retrieved as the final state of the backward-in-time adjoint integration $\mathbf{b} = \mathbf{a}(0)$. The condition $\delta\mathcal{L}/\delta\mathbf{g} = \mathbf{0}$ requires that

$$\lim_{\varepsilon\to 0}\frac{\mathcal{L}(\mathbf{q},\mathbf{g}+\varepsilon\tilde{\mathbf{g}},\mathbf{a},\mathbf{b}) - \mathcal{L}(\mathbf{q},\mathbf{g},\mathbf{a},\mathbf{b})}{\varepsilon} = 0, \quad \forall\tilde{\mathbf{g}} \tag{28}$$

Making explicit the condition, it is easily found that

$$\left[\frac{\partial\mathcal{J}}{\partial\mathbf{g}}\right]\cdot\tilde{\mathbf{g}} + \mathbf{b}\cdot\tilde{\mathbf{g}} = 0, \quad \forall\tilde{\mathbf{g}} \tag{29}$$

which is satisfied $\forall\tilde{\mathbf{g}}$ only if $[\partial\mathcal{J}/\partial\mathbf{g}] - \mathbf{b} = \mathbf{0}$ As $\partial\mathcal{J}/\partial\mathbf{g} = 2\mathbf{g}/\mathbf{q}(T)\cdot\mathbf{q}(T)$; the optimality condition reads

$$\mathbf{g} = \mathbf{b}\frac{\mathbf{q}(T)\cdot\mathbf{q}(T)}{2} \tag{30}$$

The Lagrange multiplier $\mathbf{b}$ can be eliminated and the full optimality system reads

$$\frac{d\mathbf{q}}{dt} = \mathbf{N}(\mathbf{q}), \quad \mathbf{q}(0) = \mathbf{g} \tag{31}$$

$$-\frac{d\mathbf{a}}{dt} = \left(\frac{\partial\mathbf{N}}{\partial\mathbf{q}}\right)^T\mathbf{a}, \quad \mathbf{a}(T) = -2\mathbf{q}(T)\frac{\mathbf{g}\cdot\mathbf{g}}{(\mathbf{q}(T)\cdot\mathbf{q}(T))^2} \tag{32}$$

$$\mathbf{g} = \mathbf{a}(0)\frac{\mathbf{q}(T)\cdot\mathbf{q}(T)}{2} \tag{33}$$

The first line is the state equation (evolution equation) and the associated initial condition. The second line is the adjoint problem and the third is the optimality condition. The problem stated in the first line is an initial value problem for $\mathbf{q}(t)$ with initial condition given in $t = 0$, while the adjoint problem is an initial value problem that must be integrated backward in time starting from the initial (from this perspective) condition $\mathbf{a}(T)$. The main source of difficulty for this type of problems is that the Jacobian $\partial\mathbf{N}/\partial\mathbf{q}$ being in general a function of $\mathbf{q}$, to solve the adjoint equation backward in time, the whole function $\mathbf{q}(t)$ must be known. In numerical computations, this difficulty is translated into the requirement of large storage capabilities for the whole time history of $\mathbf{q}(t)$. To reduce this storage problem one may, e.g., interpolate $\mathbf{q}$ between somehow large temporal intervals Δt or, e.g., integrate again the direct equations in each one of these intervals, which will reduce the storage requirements at the price of roughly doubling the required CPU time[7]. For this type of problem, adjoint-based methods are therefore not necessarily superior to sensitivity based approaches (but this is, actually, a matter of debate).

Exercise: For the specific problem of optimal transient growth an iterative method, similar to the one used in the second exercise of the previous section, can be used where the three equations of the optimality system are solved sequentially. The method is initialized by giving an initial guess on $\mathbf{g}$, then the loop is as follows:

(1) Given the p-th guess $\mathbf{g}^{(p)}$, compute the solution of the state equation integrating it forward in time from $t = 0$ to $t = T$ with initial condition $\mathbf{q}(0) = \mathrm{g}^{(p)}$.
(2) Compute $\mathcal{J}$ and its relative increment. If convergence is reached, stop, else continue.
(3) Use $\mathbf{q}(T)$ to compute $\mathbf{a}(T)$ and integrate the adjoint system backward in time from $t = T$ to $t = 0$.
(4) Use $\mathbf{a}(0)$ to update the control using the optimality condition and get $\mathbf{g}^{(p+1)}$. Then go to No. 1.

Apply this algorithm to compute the optimal transient growth for the linear problem $d\mathbf{q}/dt = \mathbf{L}\mathbf{q}$ with

$$\mathbf{L} = \begin{bmatrix} -1/\mathrm{Re} & 0 \\ 1 & -3/\mathrm{Re}\end{bmatrix}$$

Find $G_{\max} = \max_T G(T)$ for Re ranging from $\mathrm{Re} = 1$ to $\mathrm{Re} = 1000$. Can you find an asymptotic scaling for $G_{\max}(\mathrm{Re})$? A sample OCTAVE–MATLAB program is given below:

```
% Define system
  Rey=400.0
  T=200.0
  L=[ -1.0/Rey 0; 1-3.0/Rey]
% Exact solution using the norm function
  G_exact=(norm(expm(L*T)))^2
% Define tolerance and initialize iterations
```

[7] This latter technique is called "check-pointing" and is widely used (see, e.g., Ref. [10]).

```
  tol=10^(-8);
  g=[ rand; rand] ; % (random initial guess)
  J=10^23; dJrel=10^23; it=0;
% Iteration loop
  while (dJrel>tol)
    it=it+1; Jold=J;
    Pdir=expm(T*L); % (propagator from 0 to T)
    qT=Pdir*g; % (solve state equation forward in time)
    g2=g'*g; qT2=qT'*qT;
    J=g2/qT2; % (cost function)
    dJrel=abs((J-Jold)/J);
    aT=-2*qT*(g2)/qT2^2; % (IC for adjoint equation)
    Padj=inv(expm(T*L')); % (adjoint propagator from T to 0)
    a0=Padj*aT;
    g=a0*(qT2/2.0); % (optimality equation)
    g=g/sqrt(g'*g); % (normalize)
  end
% end of iteration loop
% print results
  G=1.0/J;
  it, G
  g % optimal initial condition
  qT % optimal response
```

The propagator from $t=0$ to $t=T$ and the adjoint backward propagator from $t=T$ to $t=0$ have been explicitly computed here but that this is not necessary as a simple integration using a time-stepper of, respectively, the state equation forward in time and the adjoint equation backward in time could have been used. This second approach, that, under certain conditions, can be used also in the case where $\mathbf{L}$ depends on t, is implemented in this SCILAB program.

```
//*********************************;
//*              FUNCTIONS              *
//*********************************;
// State equation rhs
  function[f]=StateForw(t,q,L);
  f=L*q
  endfunction;
//*********************************;
// Adjoint equation rhs
  function[f]=AdjntBack(t,a,L);
  f=L'*a
  endfunction;
//*********************************;
//*              MAINPROGRAM            *
//*********************************;
// Define system
  Rey=400.0
  T=200.0
  L=[ -1.0/Rey 0; 1-3.0/Rey]
// Exact solution using the norm function
  G_exact=(norm(expm(L*T)))^2
// Define tolerance and initialize iterations
  tol=10^(-8);
  g=[ rand(); rand()] ; // (random initial guess)
  J=10^23; dJrel=10^23; it=0;
// Iteration loop
  while (dJrel>tol)
    it=it+1; Jold=J;
    // forward integration of evolution eq
    q0=g;[ qT]=ode(q0,0,T, list(StateForw, L));
    g2=g'*g; qT2=qT'*qT;
    J=g2/qT2;
    dJrel=abs((J-Jold)/J);
    // backward integration of adjoint eq
    aT=-2*qT*(g2)/qT2^2;
[ a0]=ode(aT,T,0,list(AdjntBack,L));
    g=a0*(qT2/2.0);// enforce the optimality equation
    g=g/sqrt(g'*g); // normalize
  end
// end of iteration loop
// print results
  G=1.0/J;
  it, G
  g // optimal initial condition
  qT // optimal response
// end of program
```

■

4.2 Drive to a Final Target State by Time-Dependent Forcing. Consider now the problem of driving the final state $\mathbf{q}(T)$ as close as possible to the desired target $\mathbf{p}$ using a control continuously applied to the rhs of the evolution ODE system $d\mathbf{q}/dt=\mathbf{f}(\mathbf{q},\mathbf{g},t)$, where the initial condition $\mathbf{q}(0)=\mathbf{q}_0$ is given. A cost function including the cost of the control is

$$\mathcal{J}=\frac{1}{2}[\mathbf{q}(T)-\mathbf{p}]\cdot[\mathbf{q}(T)-\mathbf{p}]+\frac{\gamma^2}{2}\int_0^T(\mathbf{g}\cdot\mathbf{g})dt \tag{34}$$

where the weight γ^2 is chosen so as to put the desired respective emphasis on the final target objective (small values of γ) or on the cost of the control (large values of γ)[8].The Lagrangian is defined in the usual way

$$\mathcal{L}=\mathcal{J}-\int_0^T\mathbf{a}\cdot\left[\frac{d\mathbf{q}}{dt}-\mathbf{f}(\mathbf{q},\mathbf{g},t)\right]dt-\mathbf{b}\cdot[\mathbf{q}(0)-\mathbf{q}_0] \tag{35}$$

The optimality system is found, as usual, by enforcing zero variation of $\mathcal{L}$:

$$\frac{d\mathbf{q}}{dt}=\mathbf{f}(\mathbf{q},\mathbf{g},t)\ \text{(state eq.)} \tag{36}$$

$$\mathbf{q}(0)=\mathbf{q}_0 \tag{37}$$

$$-\frac{d\mathbf{a}}{dt}=\left(\frac{\partial\mathbf{f}}{\partial\mathbf{q}}\right)^T\mathbf{a},\ \text{(adjoint eq.)} \tag{38}$$

$$\mathbf{a}(T)=\mathbf{q}(T)-\mathbf{p} \tag{39}$$

$$\mathbf{g}=-\frac{1}{\gamma^2}\left[\frac{\partial\mathbf{f}}{\partial\mathbf{g}}\right]^T\mathbf{a}\ \text{(optimality cond.)} \tag{40}$$

Exercise: Derive the optimality system in Eqs. (36)–(40) ■

Example: Imagine that you have just arrived, for the holidays, in a wonderful wooden red-painted Swedish holiday house which is of course very cold after your too long absence. Your goal is to drive the house to the target temperature Θ at time T. The cost of energy has largely increased and you are also very conscious about the necessity to limit CO_2 emissions. The question is therefore: is there a way to minimize the energy consumption by using an optimal temporal distribution of the heating? Assume that, in the absence of (control) heating, the heat exchange with the environment is proportional to the temperature difference with the external temperature θ_e and that the applied heating is proportional to the input power $g(t)$ via a constant B. Assuming a constant heat capacity and defining the state of the system as $q(t)=\theta(t)-\theta_e$, the state equation is

[8] It is possible to also tune the weight to each of the components of the deviation from the target and of the cost of the control defining cost functions of the type $(\mathbf{q}(T)-\mathbf{p})\cdot\mathbf{Q}_q(\mathbf{q}(T)-\mathbf{p})$, $\mathbf{g}\cdot\mathbf{Q}_g\mathbf{g}$ with symmetric definite positive $\mathbf{Q}$s.

$$\frac{dq}{dt} = -Aq + Bg, \quad q(0) = q_0 \tag{41}$$

The cost function, defined as

$$\mathcal{J} = \frac{1}{2}[q(T) - p]^2 + \frac{\gamma^2}{2}\int_0^T g^2 dt \tag{42}$$

is a combination of the desired target performance and the cost of the heating. It is your task to find the right balance between your desire to precisely reach the desired target temperature and the energetic cost you are ready to put in it. Of course, one expects that low values of γ^2 will result in better targets but also in larger mean g consumption. The optimality system in Eqs. (36)–(40) in the present case is

$$\frac{dq}{dt} = -Aq + Bg, \quad q(0) = q_0 \tag{43}$$

$$\frac{da}{dt} = Aa, \quad a(T) = q(T) - p \tag{44}$$

$$g = -\frac{B}{\gamma^2}a \tag{45}$$

For this scalar linear problem, it is easy to find the exact solution of this system[9]. If, for instance, one assumes that at $t = 0$ the house has the outside temperature, i.e., $q_0 = 0$, then the solution is

$$\frac{g_{\text{opt}}}{p} = \left(\frac{AB}{\gamma^2}\right)\frac{e^{At}}{Ae^{AT} + \frac{B^2}{\gamma^2}\sinh(AT)}$$

$$\frac{q_{\text{opt}}}{p} = \left(\frac{B^2}{\gamma^2}\right)\frac{\sinh(At)}{Ae^{AT} + \frac{B^2}{\gamma^2}\sinh(AT)}$$

Consider, for instance, the case $A = B = 1$, $p = 1$, and $T = 6$, for which the optimal control law, state evolution, and final state are reported in Fig. 2 for the selected value of γ^2. Two properties of the solution can be easily remarked. First, as $T = 6$ is larger than the characteristic cooling time of the house, the optimal control action is mostly concentrated near the end of the temporal interval to avoid a useless energy waste if applied too early. Second, if a small weight is put on the control cost, then the target temperature if almost reached at T (e.g., $\approx 95\%$ of the target is reached when $\gamma^2 = 10^{-2}$). However, for larger weights of the control cost, the final state can be far from the desired target (e.g., less than $\approx 40\%$ of the target is reached when $\gamma^2 = 1$).

As it is relatively rare that an exact solution is accessible to calculus, let us consider, for illustrative purposes, the numerical solution of the optimality system. An iterative method can be used where, for a given approximated control law, first the state equation is solved forward in time, to get the state, then the adjoint (or costate) equation is solved backward in time. The new costate is finally used to update the control law and go to the next iteration. As the state equation is linear in the state variable, the Jacobian does not depend on the state and therefore, in this special case, there is no need to store the whole state $q(t)$ to solve the adjoint equation. However, the knowledge of the whole control function

[9]Replacing the optimality condition in the state equation a linear system is found in the ϕ variable with components $\phi = \begin{Bmatrix} g \\ a \end{Bmatrix}$, whose solution is easily found in terms of the initial conditions q_0 and a_0. The initial condition on the costate a_0 can then be expressed in terms of q_0 and p, by using $a(T) = q(T) - p$. This finally gives the complete solution for the state and the costate. The control can then be easily found using the optimality condition.

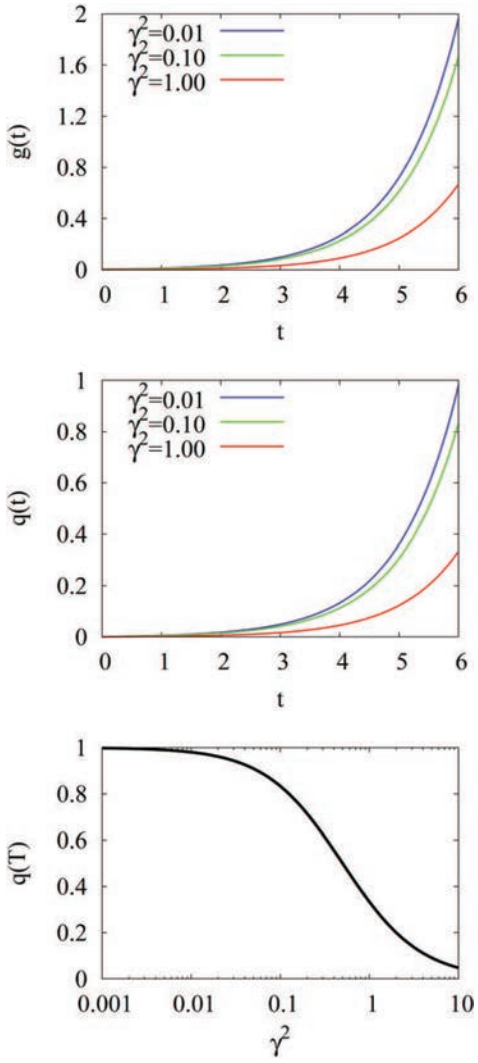

Fig. 2 Example of optimal control to reach target state $p = 1$ at $T = 6$ for different values of the control cost parameter γ^2. Top panel: optimal control laws $g(t)$, middle panel optimal state evolution $q(t)$. Bottom panel dependence of the final state $q(T)$ on the control cost parameter γ^2.

$g(t)$ and therefore costate $a(t)$ in $[0, T]$ is necessary to solve the state equation during iterations. Also, as the simple iterative solution does not converge for small values of γ^2, the control update is under relaxed using the factor α, as exemplified in the SCILAB code reported below. Using this program, it is found that the iterative solution converges, with the relative variation of $\mathcal{J}$ dropping below $\sim 10^{-6}$, in less than ≈ 16 iterations and the converged solution compares well to the exact solution.

```
//*********************************;
//*              FUNCTIONS                *
//*********************************;
// Returns the rhs state equation;
  function[ rhs] =StaEqnRHS(tloc, q, g, A, B);
    rhs=-A*q+B*g;
  endfunction;
//*********************************;
// Returns the rhs of the adjoint;
  function[ rhs] =AdjEqnRHS(tloc, a, A);
  rhs=A*a
  endfunction;
```

```
//*********************************;
//*            MAINPROGRAM          *
//*********************************;
// Here we give a value to the parameters
// and choose the initial conditions
  A=1.0, B=1.0 // coefficients of the equation
  q0=0.0 // initial condition
  p=1.0 // target temperature
  T=6.0 // target time
  gam2=0.01 // weight of control cost γ^2
  maxiter=15 // iterations
  Nt=200 // retained time samples
  α=min(0.5,gam2) // relaxation factor
  for j=1:Nt;
    t(j)=(j-1)*T/(Nt-1); // time grid
  end;
  g=0.0*t; //initialize to zero the control;
// perform iterations up to maxiter
  for iter=1:maxiter
// Integrate state eqn forward in time;
  q[1] =q0; // give IC;
  for j=2:Nt;
    // local time interval and local IC
    t_i=t(j-1); t_f=t(j); q_i=q(j-1);
    // local control in the time interval:
    gloc=0.5*(g(j)+g(j-1));
    // integrate forward and store solution in q_f
   [q_f] =ode(q_i,t_i,t_f,list(StaEqnRHS, gloc, A B));
    q(j)=q_f;
  end
  // compute cost of control and total cost
  g2int=0.5*g(1)^2+sum(g(2:Nt-1)^2)+0.5*g(Nt)^2;
  Jg(iter)=0.5*gam2*(T/(Nt-1))*g2int;
  J(iter)=0.5*(q(Nt)-p)^2+Jg(iter);
  // Integrate the adjoint equations backward in time
    //enforce IC (at T) for backward integration
    a(Nt)=(q(Nt)-p);
    for j=Nt-1:-1:1;
    // local time interval
    t_i=t(j+1); t_f=t(j); a_i=a(j+1);
    // integrate backward and store solution in a_f
   [a_f] =ode(a_i,t_i,t_f,list(AdjEqnRHS, A));
    a(j)=a_f;
  end
// Enforce optimality cond. using under-relaxation;
  g=(1-α)*g+α*(-B/gam2)*a;
// and plot the result of the current iteration;
// Plot q(t) and g(t);
  xset("window",0);
  xtitle("State and control", 't', 'q and g');
  plot2d(t',q',style=3); plot2d(t',g',style=2);
// Plot a(t);
  xset("window",1); xtitle("Costate", 't', 'a');
  plot2d(t', a',style=1);
end // end of the iteration loop
// Compare to exact solutions computed for q0=0;
  DEN=(A*exp(A*T)+(B^2/gam2)*sinh(A*T));
  q_ex=p*(B^2/gam2)*sinh(A*t)/DEN;
  g_ex=p*(A*B/gam2)*exp(A*t)/DEN;
  xset("window", 3); clf();
  xtitle("q versus q_ex", 't', 'q');
  plot2d(t',q_ex', style=-3); plot2d(t',q', style=3);
  xset("window",4); clf();
  xtitle("g versus g_ex", 't', 'g');
  plot2d(t',g_ex',style=-2); plot2d(t', g', style=2);
// Print cost function history and convergence
  for iter=2:maxiter
  dJrel(iter)=abs(1.0-J(iter-1)/J(iter));
end
[[1:maxiter]' J dJrel]
```

■

Exercise: Find the optimality system for the problem

$$\frac{dq}{dt} = -Aq + Bg, \quad q(t=0) = q_0 \tag{46}$$

in the case where the final desired target state is enforced as a constraint $q(T) = p$ and where the cost function is

$$\mathcal{J} = \frac{1}{2}\mathcal{J}_g = \int_0^T g^2 dt \tag{47}$$

Design an iterative method to solve this problem and compare the solution to the solution found in the example above. Describe and comment the results. ■

Exercise: Consider the problem of the drive of a vehicle to the target speed p at T using the control $g(t)$ on the power input in the presence of aerodynamic-type drag

$$\frac{dq}{dt} = -Aq^2 + Bg, \quad q(t=0) = q_0 \tag{48}$$

Write the optimality system in the case where the cost function is defined as

$$\mathcal{J} = \frac{1}{2}[q(T) - \mathbf{p}]^2 + \frac{\gamma^2}{2}\int_0^T g^2 dt \tag{49}$$

Numerically compute the solution using an iterative method. Can you comment the results? Does the final solution depend on the initial guess? ■

5 Inner Product and Adjoint Operators

In many optimization problems, the state is not an N-dimensional vector but a function of a real variable, say a spatial coordinate. To easily extend the variational methods discussed in previous sections to this case, we first need to partially reformulate them in more "general" terms. The dot product of two vectors will, e.g., be seen as their inner product, while a transposed matrix will be seen as the adjoint of the original matrix.

Given the vector space $\mathcal{V}$, an *inner product* (also known as *scalar product*) is defined as a functional associating to each couple of vectors $u, v \in \mathcal{V}$ a scalar denoted by $\langle u, v\rangle$, called the *inner or scalar product of u and v*. The inner product must have the following properties: (a) $\langle v, u\rangle = \langle u, v\rangle^*$, (b) $\langle u, \alpha v\rangle = \alpha\langle u, v\rangle$, $\forall \alpha \in \mathbb{C}$, (c) $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$, (d) $\langle u, u\rangle \geq 0$; $\langle u, u\rangle = 0 \Leftrightarrow u = 0$ where* denotes the complex conjugate.

A vector space $\mathcal{V}$ endowed with an inner product is an *Euclidean space*. Given the linear operator L mapping an Euclidean space $\mathcal{V}$ in itself, the operator $L^\dagger$, *adjoint* of L is defined by the property

$$\forall u, v \in \mathcal{V}\colon \langle u, Lv\rangle = \langle L^\dagger u, v\rangle \tag{50}$$

As an example, consider the vector space of N-dimensional real vectors where the inner product is defined as the standard dot product $\langle \mathbf{p}, \mathbf{q}\rangle := \mathbf{p}\cdot\mathbf{q} = p_j q_j$ (it can be easily verified that this definition respects all the properties of inner products). In this space, linear operators are associated to standard $N \times N$ matrices. To find the adjoint of the linear operator $\mathbf{M}$, associated to an $N \times N$ matrix, we apply the definition given in Eq. (50): $\langle \mathbf{p}, \mathbf{Mq}\rangle = p_i(M_{ij}q_j) = (M_{ij}p_i)q_j = (M^T_{ji}p_i)q_j = \langle \mathbf{M}^T\mathbf{p}, \mathbf{q}\rangle$. The adjoint of a linear operator associated to a real matrix is its transpose: $\mathbf{M}^\dagger = \mathbf{M}^T$.

Exercise: Consider the space of N-dimensional complex vectors for which one can define $\langle \mathbf{p}, \mathbf{q} \rangle := \mathbf{p}^* \cdot \mathbf{q} = p_j^* q_j$ (where *denotes the complex conjugate). Given the linear operator defined by the complex matrix $\mathbf{M}$ show that its adjoint $\mathbf{M}^\dagger$ is the complex conjugate transpose of $\mathbf{M}$. ■

The optimal design and control formulations described up to this point can be reformulated in terms of inner products and adjoint operators, e.g., in the static constrained optimization case the Lagrangian can be defined as $\mathcal{L} = \mathcal{J} - \langle \mathbf{a}, \mathbf{F} \rangle$ and the equation for the adjoint reads $\langle (\partial \mathcal{J}/\partial \mathbf{q}), \tilde{\mathbf{q}} \rangle - \langle \mathbf{a}, (\partial \mathbf{F}/\partial \mathbf{q}) \tilde{\mathbf{q}} \rangle = 0$. Using the definition of the adjoint: $\langle (\partial \mathcal{J}/\partial \mathbf{q}), \tilde{\mathbf{q}} \rangle - \langle (\partial \mathbf{F}/\partial \mathbf{q})^\dagger \mathbf{a}, \tilde{\mathbf{q}} \rangle = 0$ and therefore $(\partial \mathcal{J}/\partial \mathbf{q}) - (\partial \mathbf{F}/\partial \mathbf{q})^\dagger \mathbf{a} = \mathbf{0}$, which explains why the equation was called an "adjoint" equation.

This formalism can be easily extended to the case where the state is not an N-dimensional real vector but a real function of a real variable, say x. In this case, the discrete index j of the vector components is replaced by the continuous variable x, and summations on the index j are replaced by integrals in dx. Considering the space $\mathcal{V}$ of functions $f(x)$ defined for $x \in [\alpha, \beta]$, the following standard inner product is defined $\langle p, q \rangle = \int_\alpha^\beta p(x) q(x) dx$. In a static optimization problem involving a state function $q(x)$ satisfying a state equation in all points of the domain of definition, the Lagrangian will therefore be defined as usual as $\mathcal{L} = \mathcal{J} - \langle a, F \rangle$, where now it is understood that the costate is also a function of space $a(x)$ and that the inner product is a spatial integration.

As an example of linear operator mapping functions of space into functions of space consider the first order x-derivative $L = (\partial/\partial x)$ with boundary condition $q(\alpha) = 0$. To compute the adjoint of L, starting from the definition Eq. (50), we use the explicit definition of the inner product in terms of an integral

$$\left\langle p, \frac{\partial q}{\partial x} \right\rangle = \int_\alpha^\beta p \frac{\partial v}{\partial x} dx = [pq]_\alpha^\beta - \int_\alpha^\beta \frac{\partial p}{\partial x} q dx$$
$$= p(\alpha) q(\alpha) - p(\beta) q(\beta) + \left\langle -\frac{\partial p}{\partial x}, q \right\rangle$$

After integration by parts two terms arise: an inner field term and a boundary contribution. The adjoint will therefore be defined by the formal adjoint (the operator under the integral) plus appropriate adjoint boundary conditions that must "kill" the boundary terms. In our case, the formal adjoint is $L^\dagger = -(\partial/\partial x)$ and, as $q(\alpha) = 0$, the adjoint field p must satisfy the adjoint boundary condition $p(\beta) = 0$ in order to keep to zero the contribution of the boundary terms.

Exercise: Find the adjoint of the first order derivation operator d/dx defined on $[0, \lambda]$ on a vector space of periodic functions $f(x)$ satisfying the periodic boundary conditions $f(\lambda) = f(0)$. ■

We proceed as before considering now the operator $L = (\partial^2/\partial x^2)$ always using the standard inner product $\langle p, q \rangle = \int_\alpha^\beta p(x) q(x) dx$. A double integration by parts gives

$$\left\langle p, \frac{\partial^2 q}{\partial x^2} \right\rangle = \left[p \frac{\partial q}{\partial x} - \frac{\partial p}{\partial x} q \right]_\alpha^\beta + \left\langle \frac{\partial^2 p}{\partial x^2}, q \right\rangle$$

In the interior domain, therefore, $L^\dagger = \partial^2/\partial x^2$. The adjoint boundary conditions are found by setting to zero the boundary terms. If the boundary conditions on L are for instance $q(\alpha) = 0$, $q(\beta) = 0$ then the boundary conditions on the adjoint are $p(\alpha) = 0$, $p(\beta) = 0$ and the operator is *self-adjoint* because not only $L^\dagger = L$, but also the boundary conditions of $L^\dagger$ are the same of the BC of L.

6 Constrained Optimization in Systems Depending on Space and Time (PDE Systems)

Using the inner product and adjoint concepts introduced in the previous section, some easy examples of extension of the variational optimization approach to space-dependent problems are now considered. We will see that the almost trivial part of the extension consists in replacing the dot-products of the finite-dimensional case with integrations in space. However, this is not all, and additional Lagrange multipliers need to be introduced in order to take into account boundary conditions. Also, a new type of control, not considered up to now, consists in applying control via the boundary conditions.

6.1 An Example: Optimal Growth for the Nonparallel Ginzburg–Landau Model.

Consider the following linear Ginzburg–Landau initial value problem

$$\frac{\partial q}{\partial t} = \sigma(x) q - \mathcal{U} \frac{\partial q}{\partial x} + \mu \frac{\partial^2 q}{\partial x^2}, \quad q(x, t = 0) = q_0(x) \tag{51}$$

where $\sigma(x) = \sigma_0 - \sigma_2 x^2/2$ and $\sigma_2 > 0$ and $\mathcal{U}, \mu, \sigma_0, \sigma_2 \in \mathbb{R}$. The state of the system $q(x, t)$ is defined in $x \in]-\infty, \infty[$ and it is assumed that $\lim_{x \to \pm\infty} q = 0$ and that $\langle q, q \rangle$ is finite (and a similar conditions for the derivatives) with, as usual, $\langle p, q \rangle := \int_{-\infty}^{\infty} p(x) q(x) dx$. The state equation of the system can be rewritten as $F = \partial q/\partial t - \mathcal{G} q = 0$ where

$$\mathcal{G} = \sigma(x) \mathcal{I} - \mathcal{U} \frac{\partial}{\partial x} + \mu \frac{\partial^2}{\partial x^2}$$

We now consider the same problem considered in Sec. 4.1 of finding the optimal initial condition maximizing the energy transient growth a time T: $G(T) = \langle q(x, T), q(x, T) \rangle / \langle q_0(x), q_0(x) \rangle$ where the control is the initial condition $g(x) = q_0(x)$. As usual, the maximization problem is transformed into a minimization problem using

$$\mathcal{J}(q, g) = 1/G(T) = \langle g(x), g(x) \rangle / \langle q(x, T), q(x, T) \rangle \tag{52}$$

The variational formulation proceeds in the usual way, but now explicitly using the inner product formulation in the definition of the Lagrangian:

$$\mathcal{L} = \mathcal{J} - \int_0^T \left\langle a, \frac{\partial q}{\partial t} - \mathcal{G} q \right\rangle dt - \langle b, q(x, 0) - g(x) \rangle \tag{53}$$

The optimality system is found following the same kind of procedure already followed in Sec. 4.1. The analogous of Eqs. (31)–(33) is retrieved

$$\frac{\partial q}{\partial t} = \mathcal{G} q, \ \text{(state eq.)} \tag{54}$$

$$q(x, 0) = g(x) \tag{55}$$

$$-\frac{\partial a}{\partial t} = \mathcal{G}^\dagger a, \ \text{(adjoint eq.)} \tag{56}$$

$$a(x, T) = -2 q(x, T) \frac{\langle g(x), g(x) \rangle}{\langle q(x, T), q(x, T) \rangle^2} \tag{57}$$

$$g(x) = a(x, 0) \frac{\langle q(x, T), q(x, T) \rangle}{2} \ \text{(optimality cond.)} \tag{58}$$

Using the definition of adjoint and proceeding with the usual integration by parts, it is easily found that

$$\mathcal{G}^\dagger = \sigma(x) \mathcal{I} + \mathcal{U} \frac{\partial}{\partial x} + \mu \frac{\partial^2}{\partial x^2} \tag{59}$$

where the costate multiplier $a(x)$ must go to 0 for $x \to \pm\infty$. The numerical solution of this type of problem is usually obtained by first discretizing in space the variables and the operators. In this case, the continuous state variable $q(x,t)$ is replaced by the vector of, e.g., its values in some selected grid points $(\mathbf{q})_j(t) = q(x_j,t)$. The iterative method described in Sec. 4.1 can then be used to numerically solve this system.

Exercise: Compute the optimal transient growth and the optimal response to static forcing supported by the linear Ginzburg–Landau problem Eq. (51) in the finite domain $x \in [-X,X]$ using homogeneous boundary conditions for q in $x=-X$ and $x=X$. Initially consider the case with $\mathcal{U}=6$, μ=1, σ_2=0.1, X=100, and a set of values for σ_0. Discretize the problem by considering the discrete state vector $\mathbf{q}(t)$ whose components are the values $q(x_j,t)$ on the grid points $x_j = -X + 2Xj/(N_x)$ with $j = 1,\ldots,Nx-1$ (the points on the boundary are not included in the state vector). Then, discretize the spatial derivatives with the finite difference centered formulae

$$\partial q/\partial x|_j = [q(x_{j+1}) - q(x_{j-1})]/(2\Delta x)$$
$$\partial^2 q/\partial x^2|_j = [q(x_{j+1}) - 2q(x_j) + q(x_{j-1})]/(\Delta x)^2$$

where $\Delta x = 2X/N_x$. Implement the usual iterative solutions of the optimality system. Discuss the results by comparing them to the ones of Ref. [11]. ■

6.2 Another Example: Boundary Control of the Linear Ginzburg–Landau Model. In this second example, the emphasis is on the use of the boundary conditions to control the solution of an initial-boundary-value-problem. Consider the Ginzburg–Landau model Eq. (51) defined in a finite spatial domain $x \in [\alpha,\beta]$ and supplemented with an "upstream" boundary condition given by the control variable and an homogeneous "downstream" boundary condition

$$\frac{\partial q}{\partial t} = \sigma(x)q - \mathcal{U}\frac{\partial q}{\partial x} + \mu\frac{\partial^2 q}{\partial x^2}$$
$$q(x,t=0) = q_0(x), \quad q(\alpha,t) = g(t), \quad q(\beta,t) = 0$$

The goal of the optimization is to reduce the norm of the solution q at time T by the control applied at the upstream boundary. The cost function here is a weighted combination of the target (minimum norm of q at time T) and of the cost of the control

$$\mathcal{J} = \frac{1}{2}\langle q(x,T), q(x,T)\rangle + \frac{\gamma^2}{2}\int_0^T g^2(t)dt \tag{60}$$

with the usual definition $\langle p,q\rangle = \int_\alpha^\beta p(x)q(x)dx$. The Lagrangian is built using the two additional multipliers $c_\alpha(t)$, $c_\beta(t)$ necessary to enforce the boundary conditions in $x=\alpha$ and $x=\beta$, in addition to the usual multipliers $a(x,t)$ (the costate) and $b(x)$

$$\mathcal{L} = \mathcal{J} - \int_0^T \langle a, \frac{\partial q}{\partial t} - \mathcal{G}q\rangle dt - \langle b, q(x,0) - q_0\rangle$$
$$- \int_0^T c_\alpha[q(\alpha,t) - g(t)]dt - \int_0^T c_\beta q(\beta,t)dt \tag{61}$$

As usual, setting to zero the variation of $\mathcal{L}$ with respect to the Lagrange multipliers enforces the state equation and the associated initial and boundary conditions. The remaining adjoint equations and the optimality conditions are explicitly derived below.

The condition $\delta\mathcal{L}/\delta q = 0$ gives

$$\langle q(x,T), \tilde{q}(x,T)\rangle - \int_0^T \left\langle a, \frac{\partial \tilde{q}}{\partial t} - \mathcal{G}\tilde{q}\right\rangle dt - \langle b, \tilde{q}(x,0)\rangle$$
$$- \int_0^T c_\alpha \tilde{q}(\alpha,t)dt - \int_0^T c_\beta \tilde{q}(\beta,t)dt = 0, \quad \forall \tilde{q}$$

Integrating by parts in space and time gives

$$\langle q(x,T), \tilde{q}(x,T)\rangle + \langle a(x,0), \tilde{q}(x,0)\rangle - \langle a(x,T), \tilde{q}(x,T)\rangle$$
$$+ \int_0^T \left[-\mathcal{U}a\tilde{q} + \mu a\frac{\partial \tilde{q}}{\partial x} - \mu\frac{\partial a}{\partial x}\tilde{q}\right]_\alpha^\beta dt - \int_0^T \left\langle \tilde{q}, -\frac{\partial a}{\partial t} - \mathcal{G}^\dagger a\right\rangle dt$$
$$- \langle b, \tilde{q}(x,0)\rangle - \int_0^T \left[c_\alpha\tilde{q}(\alpha,t) + c_\beta\tilde{q}(\beta,t)\right]dt = 0, \quad \forall \tilde{q}$$

and collecting terms

$$\langle q(x,T) - a(x,T), \tilde{q}(x,T)\rangle + \langle a(x,0) - b, \tilde{q}(x,0)\rangle$$
$$- \int_0^T \left\langle -\frac{\partial a}{\partial t} - \mathcal{G}^\dagger a, \tilde{q}\right\rangle dt + \int_0^T \left[-\mathcal{U}a\tilde{q} + \mu a\frac{\partial \tilde{q}}{\partial x} - \mu\frac{\partial a}{\partial x}\tilde{q}\right]_\alpha^\beta dt$$
$$- \int_0^T \left[c_\alpha\tilde{q}(\alpha,t) + c_\beta\tilde{q}(\beta,t)\right]dt = 0, \quad \forall \tilde{q}$$

The first three terms require that

$$-\frac{\partial a}{\partial t} = \mathcal{G}^\dagger a, \quad a(x,T) = q(x,T), \quad b(x) = a(x,0)$$

The boundary terms under temporal integration are

$$\mu a_\beta \frac{\partial \tilde{q}}{\partial x}\bigg|_\beta + \left[\mathcal{U}a + \mu\frac{\partial a}{\partial x} - c\right]_\alpha \tilde{q}(\alpha,t)$$
$$- \mu a_\alpha \frac{\partial \tilde{q}}{\partial x}\bigg|_\alpha - \left[\mathcal{U}a + \mu\frac{\partial a}{\partial x} - c\right]_\beta \tilde{q}(\beta,t)$$

which, setting the variations to zero $\forall\tilde{q}$ requires that $a(\beta,t)=0$, $a(\alpha,t)=0$, $c_\alpha(t) = \mathcal{U}a(\alpha,t) + \mu[\partial a/\partial x](\alpha,t)$, and $c_\beta(t) = -\mathcal{U}a(\beta,t) - \mu[\partial a/\partial x](\beta,t)$ and therefore

$$a(\beta,t)=0, a(\alpha,t)=0, \quad c_\alpha(t) = \mu\frac{\partial a}{\partial x}(\alpha,t), \quad c_\beta(t) = -\mu\frac{\partial a}{\partial x}(\beta,t)$$

The condition $\delta\mathcal{L}/\delta g = 0$ gives

$$\gamma^2\int_0^T g(t)\tilde{g}(t)dt + \int_0^T c_\alpha(t)\tilde{g}(t)dt = 0, \quad \forall\tilde{g}$$

which gives the optimality condition

$$g(t) = -c_\alpha(t)/\gamma^2$$

The full optimality system is therefore given by

$$\frac{\partial q}{\partial t} = \mathcal{G}q, \quad q(x,0) = q_0(x), \quad q(\alpha,t) = g(t), \quad q(\beta,t) = 0 \tag{62}$$

$$-\frac{\partial a}{\partial t} = \mathcal{G}^\dagger a, \quad a(x,T) = q(x,T), \quad a(\alpha,t) = 0, \quad a(\beta,t) = 0 \tag{63}$$

$$g(t) = -\frac{\mu}{\gamma^2}\frac{\partial a}{\partial x}(\alpha,t) \tag{64}$$

No need to mention that this system can be solved using iterative techniques.

Exercise: Find the optimality system for the boundary control of the nonlinear real Ginzburg–Landau equation defined by $\partial q/\partial t = \mathcal{N}(q)$ with

$$\mathcal{N}(q) = \mathcal{G}q - \chi q^3 = \sigma(x)q - \mathcal{U}\frac{\partial q}{\partial x} + \mu\frac{\partial^2 q}{\partial x^2} - \chi q^3$$ ■

Exercise: Find the optimality system for the control by additive forcing of the linear real Ginzburg–Landau equation defined by $\partial q/\partial t = \mathcal{G}q + g$ with homogeneous boundary condition for q in $x = \alpha$ and $x = \beta$. ■

6.3 A Quick Tour of the Derivation of Adjoint Linearized Navier–Stokes Equations in Primitive Variables. Consider the Navier–Stokes equations for an incompressible viscous flow defined in the spatial domain Ω with boundary $\partial\Omega$ and in the time interval $[0, T]$:

$$\nabla \cdot \mathbf{u} = 0 \tag{65}$$

$$\partial \mathbf{u}/\partial t + (\nabla \mathbf{u})\mathbf{u} - \nu \nabla^2 \mathbf{u} + \nabla p = 0 \tag{66}$$

supplemented by appropriate initial and boundary conditions. The continuity (first) equation is scalar and derives from the mass conservation, while the second is a vector equation deriving from momentum conservation. The state of the system is given by the velocity and pressure fields $\mathbf{u}(\mathbf{x}, t), p(\mathbf{x}, t)$.

In the definition of the Lagrangian, to guarantee that the mass and momentum conservation are respected $\forall t$, Eqs. (65) and (66) are projected on the two Lagrange multipliers $\mathbf{u}^\dagger(\mathbf{x}, t)$ and $p^\dagger(\mathbf{x}, t)$ that take the role of the adjoint velocity and pressure fields respectively. By "project" we mean that an inner product is performed, where the inner product of real scalar fields is defined by $\int_\Omega v(\mathbf{x}) r(\mathbf{x}) d\Omega$ and the inner product of real vector fields is defined by $\int_\Omega \mathbf{v} \cdot \mathbf{r} d\Omega$.

The usual integration by parts used to extract the adjoint operators generates field terms in Ω and boundary terms on $\partial\Omega$. Consider, e.g., the part of $\mathcal{L}$ given by the projection of Eq. (65) on $p^\dagger(\mathbf{x}, t)$: $\int_\Omega p^\dagger (\nabla \cdot \mathbf{u}) d\Omega$. The first variation of this term with respect to the state generates a term of the type $\int_\Omega p^\dagger (\nabla \cdot \tilde{\mathbf{u}}) d\Omega$. Considering that $p^\dagger(\nabla \cdot \tilde{\mathbf{u}}) = p^\dagger(\partial \tilde{u}_i/\partial x_i) = (\partial p^\dagger \tilde{u}_i/\partial x_i) - \tilde{u}_i(\partial p^\dagger/\partial x_i) = \nabla \cdot (p^\dagger \tilde{\mathbf{u}}) - \tilde{\mathbf{u}} \cdot \nabla p^\dagger$, the integral is found to be

$$\int_\Omega p^\dagger (\nabla \cdot \tilde{\mathbf{u}}) d\Omega = \int_\Omega \nabla \cdot (p^\dagger \tilde{\mathbf{u}}) d\Omega - \int_\Omega \tilde{\mathbf{u}} \cdot \nabla p^\dagger d\Omega$$

The use of the divergence theorem, which is the "integration by parts" applying to this case, gives

$$\int_\Omega p^\dagger (\nabla \cdot \tilde{\mathbf{u}}) d\Omega = \int_{\partial\Omega} p^\dagger \tilde{\mathbf{u}} \cdot \vec{n} d\Omega - \int_\Omega \tilde{\mathbf{u}} \cdot \nabla p^\dagger d\Omega$$

The first term will give a condition on $p^\dagger$ on the boundary $\partial\Omega$, while the second term under integral gives the adjoint pressure gradient $-\nabla p^\dagger$ that will appear in the adjoint of the momentum conservation equation.

As an example, consider the problem of distributed control of the Navier–Stokes equations[10] with homogeneous boundary conditions

$$\begin{aligned}
&\nabla \cdot \mathbf{u} = 0 \\
&\partial \mathbf{u}/\partial t + (\nabla \mathbf{u})\mathbf{u} - \nu \nabla^2 \mathbf{u} + \nabla p = \mathbf{g} \\
&\mathbf{u}(\mathbf{x} \in \partial\Omega, t) = \mathbf{0}, \quad \mathbf{u}(\mathbf{x}, t = 0) = \mathbf{u}_0(\mathbf{x}) \\
&\int_\Omega p d\Omega = 0
\end{aligned}$$

where the distributed control $\mathbf{g}$ is applied in order to force the system to approach the desired state ($\mathbf{U}$) and the following penalized functional is considered:

$$\mathcal{J} = \frac{1}{2}\int_0^T \int_\Omega |\mathbf{u} - \mathbf{U}|^2 d\Omega dt + \frac{\gamma_1^2}{2}\int_\Omega |\mathbf{u} - \mathbf{U}|^2_{t=T} d\Omega + \frac{\gamma_2^2}{2}\int_0^T \int_\Omega |\mathbf{g}|^2 d\Omega dt$$

For this problem, it is found that the adjoint equations are

$$\begin{aligned}
&\nabla \cdot \mathbf{u}^\dagger = 0 \\
&-\partial \mathbf{u}^\dagger/\partial t + (\nabla \mathbf{u})^T \mathbf{u}^\dagger - (\nabla \mathbf{u}^\dagger)\mathbf{u} - \nu \nabla^2 \mathbf{u}^\dagger - \nabla p^\dagger = \mathbf{u} - \mathbf{U} \\
&\mathbf{u}^\dagger(\mathbf{x} \in \partial\Omega, t) = \mathbf{0}, \mathbf{u}^\dagger(\mathbf{x}, T) = \gamma_1^2(\mathbf{u} - \mathbf{U}) \\
&\int_\Omega p^\dagger d\Omega = 0
\end{aligned}$$

and the optimality condition reads

$$\mathbf{g} = -\mathbf{u}^\dagger/\gamma_2^2 \tag{67}$$

The expression of the adjoint operator is not surprising: first order derivatives (gradient terms) have changed sign, while second-order derivatives have not; the adjoint equation and its initial condition at $t = T$ contain source terms that are proportional to the deviation from the target and would go to zero when the target of the optimization is attained.

7 Feedback Control of Linear Systems With Quadratic Cost-Functions

For unsteady problems, the optimal control approach to the regulation problem discussed, e.g., in Sec. 4.2, has two important shortcomings that can limit its application to a range of practical situations. First, the computed optimal control law $\mathbf{g}(t)$ depends on the specific initial condition $\mathbf{q}_0$ assigned to the problem. This is a problem because a virtually infinite number of control laws must be in principle stored if one wants to take into account all possible initial conditions, not to mention the fact that the full temporal history of the control must be stored. Second, in the design of the control law $\mathbf{g}(t)$ it is assumed that the full temporal history of the full state $\mathbf{q}(t)$ is available for all $t \in [0, T]$. Accessing this complete history for different "scenarios" of control in order to find the optimal one can, however, be too demanding for a wide range of applications. Two very legitimate questions are therefore: can an effective (even if probably more expensive) control law be found using only the current state information? If yes, can this law be made independent of time, at least in time invariant systems, and still provide effective control? In some types of system, the answer to these questions is yes, under some condition, and the used approach is that of *feedback control*.

Consider, e.g., the simple linear time invariant (LTI) scalar system $dq/dt = Aq$ with the scalar $A > 0$ and the time interval $t \in [0, \infty[$. The solution $q(t) = q_0 e^{At}$ to the initial value problem with initial condition $q(0) = q_0$ diverges exponentially with growth rate A, and this for almost all admissible initial conditions. Consider now the goal of driving to zero, at least for $t \to \infty$, the state $q(t)$ and to reduce as much as possible the cost function $\mathcal{J} = \int_0^\infty q(t)^2 dt$ using an additive linear control term. In this case, the evolution equation becomes $dq/dt = Aq + Bg$ with, say $B > 0$. An optimal control, like the one considered in Sec. 4.2, can be designed to drive to zero the state as $t \to \infty$. However, the same result can be reached using a simple linear feedback control $g = -Kq$ where the control is in opposition to the state with gain K. The evolution equation becomes $dq/dt = (A - BK)q$ with solution $q(t) = q_0 e^{(A-BK)t}$. The solution is stabilized for sufficiently large control gains $K > B^{-1}A$. In this easy example, therefore, a time independent linear feedback control law is able to drive the state to zero (which is the objective of the control) with a control based only on information on the present state (and not of its past

[10] This example is taken from Ref [7], p. 77.

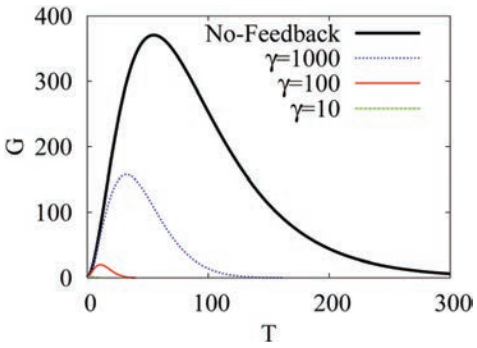

Fig. 3 Example of Riccati-based feedback control of a non-normal linear stable system supporting transient energy growths. The optimal transient growth $G(T)$ of the uncontrolled system (solid line, black) is compared to the controlled cases Riccati-based feedback for decreasing values of the cost parameter γ. For large values of γ the optimal transient growth is reduced and it is completely suppressed when γ is lowered enough. For, e.g., $\gamma = 10$, $G_{\max} = 1$.

or future history). The control law is furthermore effective for all possible initial conditions.

The simple example can be generalized to N-dimensional linear systems with quadratic cost functions (LQ systems), where the state equation and the zero cost-gradient condition, being linear, have unique solutions

$$\frac{d\mathbf{q}}{dt} = \mathbf{A}\mathbf{q} + \mathbf{B}\mathbf{g}; \quad \mathbf{q}(0) = \mathbf{q}_0 \tag{68}$$

$$\mathcal{J}(\mathbf{q}, \mathbf{g}) = \frac{1}{2}\int_0^T \left(\mathbf{q}\cdot\mathbf{Q}\mathbf{q} + \gamma^2 \mathbf{g}\cdot\mathbf{g}\right)dt \tag{69}$$

It is assumed that $\mathbf{Q}$ is symmetric definite positive and that $\mathbf{B}$ has maximum rank. Proceeding in the usual way, the Lagrangian is defined as

$$\mathcal{L} = \mathcal{J} - \int_0^T \mathbf{a}\cdot\left[\frac{d\mathbf{q}}{dt} - \mathbf{A}\mathbf{q} - \mathbf{B}\mathbf{g}\right]dt - \mathbf{b}\cdot[\mathbf{q}(0) - \mathbf{q}_0]$$

and the optimality system formed by the state and adjoint evolution equations and the optimality condition is found to be

$$\frac{d\mathbf{q}}{dt} = \mathbf{A}\mathbf{q} + \mathbf{B}\mathbf{g}, \quad \mathbf{q}(0) = \mathbf{q}_0 \tag{70}$$

$$-\frac{d\mathbf{a}}{dt} = \mathbf{A}^T\mathbf{a} + \mathbf{Q}\mathbf{q}, \quad \mathbf{a}(T) = 0 \tag{71}$$

$$\mathbf{g} = -\frac{1}{\gamma^2}\mathbf{B}^T\mathbf{a} \tag{72}$$

Exercise: Derive Eqs. (70)–(72). ■

Replacing the optimality condition (last equation) in the state equation (first) the structure of the optimization problem is shown to be the linear coupled direct-adjoint system

$$\frac{d\mathbf{q}}{dt} = \mathbf{A}\mathbf{q} - \frac{1}{\gamma^2}\mathbf{B}\mathbf{B}^T\mathbf{a}, \quad \mathbf{q}(0) = \mathbf{q}_0 \tag{73}$$

$$\frac{d\mathbf{a}}{dt} = -\mathbf{A}^T\mathbf{a} - \mathbf{Q}\mathbf{q}, \quad \mathbf{a}(T) = 0 \tag{74}$$

For optimal solutions satisfying this system, stating that the costate is linearly related to the state via $\mathbf{a} = \mathbf{X}\mathbf{q}$ is equivalent to stating that the feedback law is linear via $\mathbf{g} = -\mathbf{K}\mathbf{q}$. Indeed, replacing the linear feedback law into the optimality condition of Eq. (72), it is found that $\mathbf{K} = \mathbf{B}^H\mathbf{X}/\gamma^2$. The optimal linear feedback control law $\mathbf{K}$ can be therefore computed from $\mathbf{X}$. Differentiating wrt to time the relation $\mathbf{a} = \mathbf{X}\mathbf{q}$ and then using Eqs. (73)–(74), the following Riccati equation for $\mathbf{X}$ is found:

$$-\frac{d\mathbf{X}}{dt} = \mathbf{X}\mathbf{A} + \mathbf{A}^T\mathbf{X} + \mathbf{Q} - \frac{1}{\gamma^2}\mathbf{X}\mathbf{B}\mathbf{B}^T\mathbf{X}, \quad \mathbf{X}(T) = \mathbf{0} \tag{75}$$

that must be integrated backward in time starting from T. Taking the transpose of Eq. (75), it is easily seen that if $\mathbf{X}$ is a solution, then $\mathbf{X}^T$ is also a solution of the equation. The solution $\mathbf{X}(t)$ allows us to design the optimal feedback control law $\mathbf{K}(t)$. As mentioned, however, in many applications, storing the whole temporal history of the $N \times N$ operator $\mathbf{K}(t)$ can be not practical or unaffordable. If the time horizon goes to infinity ($T \to \infty$), under some conditions, it can be proved that the solution of the differential Riccati equation converges to the steady real, symmetric solution $\mathbf{X}_\infty$ of the *algebraic Riccati equation*

$$\mathbf{X}_\infty\mathbf{A} + \mathbf{A}^T\mathbf{X}_\infty + \mathbf{Q} - \frac{1}{\gamma^2}\mathbf{X}_\infty\mathbf{B}\mathbf{B}^T\mathbf{X}_\infty = \mathbf{0} \tag{76}$$

Standard methods exist to compute solutions of the algebraic Riccati equation and are implemented in standard software such as OCTAVE, SCILAB, and MATLAB. Once $\mathbf{X}_\infty$ is known, the time-independent feedback operator is easily computed.

Example: Consider the following LQ problem with:

$$\mathbf{A} = \begin{bmatrix} -1/\mathrm{Re} & 0 \\ 1 & -3/\mathrm{Re} \end{bmatrix}; \quad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \quad \mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

We have already seen that the linear operator $\mathbf{A}$ defines a stable system that, however, can sustain large transient energy growths for sufficiently large Re. In the following, we will consider, e.g., $\mathrm{Re} = 100$. We want to verify what is the effect of a Riccati-based optimal feedback control on transient growths supported by this system. The SCILAB program below is used to solve the algebraic Riccati equation associated to this system and compute the (time-constant) Riccati-based feedback matrix $\mathbf{K}$ for selected values $\gamma = 1000, 100, 10$, and 1. To quantify the ability of Riccati-based feedback control to reduce the worst-case transient growths, the optimal transient growth is computed on the closed-loop controlled system whose evolution equation is given by

$$d\mathbf{q}/dt = (\mathbf{A} - \mathbf{B}\mathbf{K})\mathbf{q}$$

From the results reported in Fig. 3, it is seen that the optimal transient growth is reduced by more than one half even for the very large value of the control cost $\gamma = 1000$. For $\gamma = 100$, i.e., assuming that the energy of the control costs $\gamma^2 = 10^4$ times the energy of the controlled state, the maximum transient growth is reduced by a factor larger that 10^2 and it is almost completely suppressed for $\gamma = 10$. Optimal Riccati-based feedback control is therefore quite effective in reducing worst-case transient growths in stable non-normal linear systems. Similar results, complicated by the issue of state observations and that the control is limited to the boundary, have been obtained, e.g., in the case of plane Poiseuille flow, where the system is given by the linearized Navier–Stokes equations [12,13].

```
// define system
  Rey=100
  A=[ -1/Rey, 0; 1, -3/Rey]
  B=[ 1, 0; 0, 1]
  Q=[ 1, 0; 0, 1]
  gam=1000
// solve Riccati equation
  R=B'*B/gam^2;
  X=riccati(A, R, Q, 'c', 'schur')
// compute feedback matrix
```

```
  K=B'*X/gam^2
// linear modal stability of uncontrolled system
  spec(A)
// optimal transient energy growth of uncontrolled
// system is the L2 norm of exp(At)
  Nt=200; t=linspace(0,3*Rey,Nt);
  for j=1:Nt;
    Gfree(j)=norm(expm(A*t(j)))^2;
  end
  Gfreemax=max(Gfree)
// linear modal stability of controlled system
  spec(A-B*K)
// optimal transient energy growth of controlled
// system is the L2 norm of exp((A-BK)t)
  for j=1:Nt
    Gcont(j)=norm(expm((A-B*K)*t(j)))^2;
  end
  Gcontmax=max(Gcont)
  plot2d(t,[(Gfree),(Gcont)])
// end of program
```

■

Exercise: Consider the system defined in the above example. Compute and plot the curves $G_{\max}(\mathrm{Re})$ for Re ranging from 10 to 1000, for the controlled systems obtained for $\gamma = 1000, 100, 10,$ and 1 and compare them to the uncontrolled case curve. For each one of these cases, plot the curve of the raw control cost $J_g = (1/2)\int_0^T \mathbf{g}\cdot\mathbf{g}dt$ computed for the worst case initial condition realizing the maximum transient growth, $G_{\max}$. Plot the results as a curve $J_g(\gamma)$ for the selected RE. ■

Exercise: Consider now the LQ problem defined by

$$\mathbf{A} = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}; \quad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \quad \mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

where the uncontrolled system is unstable. Modify the SCILAB program given in the example above to solve the algebraic Riccati equation associated to this system and compute the (time-constant) Riccati-based feedback matrix $\mathbf{K}$ for the following selected values of the control cost weight parameter $\gamma = 1000, 100, 10,$ and 1. Analyze the linear stability of the controlled system and comment on the results. Compute also which is the worst case initial condition realizing the maximum energy growth and, for this initial condition, compute $J_g = \frac{1}{2}\int_0^T \mathbf{g}\cdot\mathbf{g}dt$. Consider the $J_g(\gamma)$ dependence and comment on the results. ■

8 Conclusion

In these lecture notes, a very informal introduction to constrained optimization techniques has been given. The specific considered case has been the one of equality constraints and differentiability has been (implicitly) assumed on the constraints, the functional, the state, and control variables. Many of these assumptions can be relaxed without major difficulties, but this goes beyond the scope of the these lecture notes.

Other important issues have also been ignored here. For instance, the system we want to control may be not controllable in the sense that a part of the admissible states of the system may not be accessible under any admissible control law. This is a problem, e.g., in iterative methods, if the convergence path crosses one of these regions. In addition, in finding optimal control laws it has been assumed that information about the full state is available. This is rarely the case in practical fluid dynamics applications where only an "output" $\mathbf{y} = \mathbf{Y}(\mathbf{q}, \mathbf{g})$ is accessible. In this case, the state needs to be estimated from the knowledge of $\mathbf{y}$, which of course sets the problem of observability of a state, i.e., to know if all allowed states can be identified from the knowledge of $\mathbf{y}$ and $\mathbf{g}$. Moreover, the system's uncertainty has been ignored. In real systems, noise enters the system as an additional input in addition to the control law, and noise can pollute the observations of the system $\mathbf{y}$. Furthermore, the parameters or the coefficient of the state equation itself can be subject to uncertainty. In this context, it is very important to design a control that is robust to these different uncertainties, i.e., that does not lead to undesired outcomes because of them. A quite complete theoretical framework addressing all these issues exists in the case of linear systems with quadratic cost functions and Gaussian noise. We refer the reader to the excellent monographs dedicated to these issues, such as, e.g., Ref. [6].

Many of the difficulties mentioned above, such as a lack of controllability or observability of some states, often derive from the fact that the choice of the control type, of the type and location of the measures of the state providing the observation vector and of the cost function, has been made without a proper understanding of the physics of the considered system or without taking it into account when posing the optimization problem. If, e.g., you want to drive a house to a desired temperature in a cold winter and you look for an optimal control law to minimize the energy consumption (see Sec. 4.2), the problem will have no solution if, e.g., the maximum available heating power is not able to overcome the natural cooling of the house, which technically would arise as uncontrollability of the system. Also, the optimal control will probably fail if you do not take into account system uncertainties such as someone leaving randomly the entrance door open, or if using a sensor-based control, the thermometer is placed outside the house (unobservability) or near the heater (poor observability) or near the entrance door that you leave often open (random input much larger than state or control input). These considerations also apply to fluid dynamics applications. Consider, e.g., the case where one wants to stabilize a given basic flow against a convective instability, which essentially amplifies downstream perturbations that enter the flow upstream. In such a case, one would, e.g., put actuators upstream of the domain where the cost function is defined. Not doing so would result in trivial uncontrollability.

Another type of problem which has not been taken into account is the one of model reduction. In fluid dynamic applications, the dimension of the state vector can quickly become exceedingly large if, e.g., considering large domains or complex geometries or even moderately large Reynolds numbers. This leads to optimization procedure whose cost could be too large for practical applications, even with the most advanced and efficient control algorithms. Model reduction, in some situations can be a remedy to this problem. The main idea of model reduction is to expand the state into a linear combination of basis functions $q(x,t) = \sum_{j=1}^{M} \hat{q}_j(t)\phi_j(x)$ where the dimension M is smaller than the problem dimension N. An efficient model reduction requires that one is sufficiently smart (or lucky) to be able to find a "good" type of functions and order them in a "good" way, so as to be able with a (relatively) small M to design a control law able to sufficiently approach the optimal $\mathcal{J}$. This reduced system approach of course would, e.g., require the definition of a reduced state equation, but leads to large computational gains. By now, model reduction is a well-developed, almost autonomous field by its own, which is still an active field of research, with well-developed techniques such as the proper orthogonal decomposition (based on the empirical orthogonal eigenfunctions, see, e.g., Ref. [14]), balanced truncation (see, e.g., Refs. [6,15]), Koopman modes (see, e.g., Ref. [16]), etc. Of course, in addition to model reduction, other techniques are available to bypass excessively costly standard optimization procedures applied to large systems typical of fluid dynamics applications (see, e.g., Ref. [17]).

Acknowledgment

I thank KTH-FLOW and NORDITA, where part of these lecture notes have been prepared, for their hospitality.

References

[1] Amoignon, O. G., Pralits, J. O., Hanifi, A., Berggren, M., and Henningson, D. S., 2006, "Shape Optimization for Delay of Laminar-Turbulent Transition," AIAA J., **44**(5), pp. 1009–1024.

[2] Le Dimet, F., and Talagrand, O., 1986, "Variational Algorithms for Analysis and Assimilation: Theoretical Aspects," Tellus, **38A**, pp. 97–110.

[3] Luenberger, D., 1968, *Optimization by Vector Space Methods*, John Wiley & Sons, New York.

[4] Lions, J. L., 1971, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, Berlin.

[5] Pironneau, O., 1984, *Optimal Shape Design for Elliptic Systems*, Springer, New York.

[6] Zhou, K., Doyle, J., and Glover, K., 1996, *Robust and Optimal Control*, Prentice Hall, Englewood Cliffs, NJ.

[7] Gunzburger, M., 2002, *Perspectives in Flow Control and Optimization*, SIAM, Philadelphia, PA.

[8] Dennis, J., and Schnabel, R., 1996, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, SIAM, Philadelphia, PA.

[9] Corbett, P., and Bottaro, A., 2000, "Optimal Perturbations for Boundary Layers Subject to Streamwise Pressure Gradient," Phys. Fluids, **12**, pp. 120–130.

[10] Hinze, M., Walther, A., and Sternberg, J., 2006, "An Optimal Memory-Reduced Procedure for Calculating Adjoints of the Instationary Navier–Stokes Equations," Opt. Cont. Appl. Metods, **27**(1), pp. 19–40.

[11] Cossu, C., and Chomaz, J. M., 1997, "Global Measures of Local Convective Instabilities," Phys. Rev. Lett., **78**(23), pp. 4387–4390.

[12] Joshi, S. S., Speyer, J. L., and Kim, J., 1997, "A Systems Theory Approach to the Feedback Stabilization of Infinitesimal and Finite-Amplitude Disturbances in Plane Poiseuille Flow," J. Fluid Mech., **332**, pp. 157–184.

[13] Kim, J., and Bewley, T. R., 2007, "A Linear Systems Approach to Flow Control," Ann. Rev. Fluid Mech., **39**, pp. 383–417.

[14] Lumley, J., 1970, *Stochastic Tools in Turbulence Academic*, Academic Press, New York.

[15] Rowley, C. W., 2005, "Model Reduction for Fluids Using Balanced Proper Orthogonal Decomposition," Int. J. Bifurc. Chaos, **15**(3), pp. 997–1013.

[16] Rowley, C. W., Mezić, I., Bagheri, S., Schlatter, P., and Henningson, D. S., 2009, "Spectral Analysis of Nonlinear Flows," J. Fluid Mech., **641**, pp. 115–127.

[17] Semeraro, O., Pralits, J., Rowley, C., and Henningson, D., 2013, "Riccati-Less Approach for Optimal Control and Estimation: An Application to Two-Dimensional Boundary Layers," J. Fluid Mech., **731**, pp. 394–417.