Optimal control of complex flows

Jan Pralits

Department of Civil, Chemical and Environmental Engineering
University of Genoa, Italy
jan.pralits@unige.it

October 22, 2013

The work presented has been carried out by:

- **Jan Pralits** DICCA, University of Genoa, Italy;
- **Onofrio Semeraro** Post-Doc, Ladhyx, Paris, France;
- **Marco Carini** Research Assistant, Politecnico di Milano, Italy;
- **Paolo Luchini** DIIN, University of Salerno, Italy;
- **Tom Bewley** UCSD, USA;

Introduction

Definitions

Complex flows

here problems with large number of degrees of freedom

Optimal control

The linearized system

$$\frac{\partial \mathbf{x}}{\partial t} = A\mathbf{x} + B\mathbf{u} \quad \text{on} \quad 0 < t < T, \quad \text{with} \quad \mathbf{x} = \mathbf{x}_0 \quad \text{at} \quad t = 0.$$

- $\mathbf{x}$ has dimension n and $\mathbf{u}$ dimension m
- here $n >> m$
- find $\mathbf{u}$ that minimizes a quadratic cost function J
- consider: full state information, no estimation (has been done, references available)

The linear optimal control problem

The classical full-state-information control problem is formulated as: find the control $\mathbf{u}$ that minimizes the cost function

$$J = \frac{1}{2}\int_0^T \left[\mathbf{x}^H Q \mathbf{x} + l^2 \mathbf{u}^H R \mathbf{u}\right] dt,$$

where l is the penalty of the control, and the state $\mathbf{x}$ and the control $\mathbf{u}$ are related via the state equation

$$\frac{\partial \mathbf{x}}{\partial t} = A\mathbf{x} + B\mathbf{u} \quad \text{on} \quad 0 < t < T, \quad \text{with} \quad \mathbf{x} = \mathbf{x}_0 \quad \text{at} \quad t = 0.$$

The solution depends on: $\mathbf{x}_0$, T, Q, R and l.

Solution approaches

- With a feedback rule $\mathbf{u} = K\mathbf{x}$, and a system which is LTI, then the feedback matrix K is computed once off-line
  (convenient since K is independent of $\mathbf{x}_0$).

- Optimal control $\mathbf{u}$ corresponding to the state at each time step is computed in real time, normally with a finite horizon (value of T) to make it tractable.
  (Example: Adjoint-based control optimization.)

Both approaches can be solved using the adjoint of the state equation.

Why introduce Adjoint equations ?

Example gradient computation:

$$J = \mathbf{w}^H \mathbf{x}, \qquad \text{where} \qquad A\mathbf{x} = \mathbf{b}, \qquad \text{Ex. find} \qquad \frac{\partial J}{\partial \mathbf{b}}$$

Why introduce Adjoint equations ?

Example gradient computation:

$$J = \mathbf{w}^H \mathbf{x}, \qquad \text{where} \quad A\mathbf{x} = \mathbf{b}, \quad \text{Ex. find} \quad \frac{\partial J}{\partial \mathbf{b}}$$

Finite difference approach

$$\left(\frac{\partial J}{\partial \mathbf{b}}\right)_j \approx \frac{J(\mathbf{b} + \epsilon \mathbf{e}_j) - J(\mathbf{b})}{\epsilon}$$

Why introduce Adjoint equations ?

Example gradient computation:

$$J = \mathbf{w}^H \mathbf{x}, \qquad \text{where} \quad A\mathbf{x} = \mathbf{b}, \quad \text{Ex. find} \quad \frac{\partial J}{\partial \mathbf{b}}$$

Finite difference approach

$$\left(\frac{\partial J}{\partial \mathbf{b}}\right)_j \approx \frac{J(\mathbf{b} + \epsilon \mathbf{e}_j) - J(\mathbf{b})}{\epsilon}$$

Requires $n + 1$ solutions of $A\mathbf{x} = b$, where n is the dimension of $\mathbf{b}$

Why introduce Adjoint equations ?

Example gradient computation:

$$J = \mathbf{w}^H \mathbf{x}, \qquad \text{where} \quad A\mathbf{x} = \mathbf{b}, \quad \text{Ex. find} \quad \frac{\partial J}{\partial \mathbf{b}}$$

Finite difference approach

$$\left(\frac{\partial J}{\partial \mathbf{b}}\right)_j \approx \frac{J(\mathbf{b} + \epsilon \mathbf{e}_j) - J(\mathbf{b})}{\epsilon}$$

Requires $n+1$ solutions of $A\mathbf{x} = b$, where n is the dimension of $\mathbf{b}$

Alternatively, solve

$$A^H \mathbf{p} = \mathbf{w} \qquad \text{dual problem, adjoint}$$

Why introduce Adjoint equations ?

Example gradient computation:

$$J = \mathbf{w}^H \mathbf{x}, \qquad \text{where} \quad A\mathbf{x} = \mathbf{b}, \quad \text{Ex. find} \quad \frac{\partial J}{\partial \mathbf{b}}$$

Finite difference approach

$$\left(\frac{\partial J}{\partial \mathbf{b}}\right)_j \approx \frac{J(\mathbf{b} + \epsilon \mathbf{e}_j) - J(\mathbf{b})}{\epsilon}$$

Requires $n+1$ solutions of $A\mathbf{x} = b$, where n is the dimension of $\mathbf{b}$

Alternatively, solve

$$A^H \mathbf{p} = \mathbf{w} \qquad \text{dual problem, adjoint}$$

then

$$J = \mathbf{w}^H \mathbf{x} = (A^H \mathbf{p})^H \mathbf{x} = \mathbf{p}^H A \mathbf{x} = \mathbf{p}^H \mathbf{b},$$

Why introduce Adjoint equations ?

Example gradient computation:

$$J = \mathbf{w}^H \mathbf{x}, \qquad \text{where} \quad A\mathbf{x} = \mathbf{b}, \quad \text{Ex. find} \quad \frac{\partial J}{\partial \mathbf{b}}$$

Finite difference approach

$$\left(\frac{\partial J}{\partial \mathbf{b}}\right)_j \approx \frac{J(\mathbf{b} + \epsilon \mathbf{e}_j) - J(\mathbf{b})}{\epsilon}$$

Requires $n + 1$ solutions of $A\mathbf{x} = b$, where n is the dimension of $\mathbf{b}$

Alternatively, solve

$$A^H \mathbf{p} = \mathbf{w} \qquad \text{dual problem, adjoint}$$

then

$$J = \mathbf{w}^H \mathbf{x} = (A^H \mathbf{p})^H \mathbf{x} = \mathbf{p}^H A \mathbf{x} = \mathbf{p}^H \mathbf{b},$$

and

$$\frac{\partial J}{\partial \mathbf{b}} = \mathbf{p} \qquad \text{One Solution, independently of } n$$

Derivation of adjoint I

The adjoint variable $\mathbf{p}$ is introduced as a Lagrange multiplier. The augmented cost function is written

$$J = \int_0^T \frac{1}{2}[\mathbf{x}^H Q\mathbf{x} + l^2 \mathbf{u}^H R\mathbf{u}] - \mathbf{p}^H[\frac{\partial \mathbf{x}}{\partial t} - A\mathbf{x} - B\mathbf{u}]\, dt,$$

linearize + integration by parts and $\delta J = 0$ gives

$$0 = \int_0^T \delta\mathbf{u}^H[B\mathbf{p} + l^2 R\mathbf{u}] + \delta\mathbf{x}^H[\frac{\partial \mathbf{p}}{\partial t} + A^H\mathbf{p} + Q\mathbf{x}]\, dt + [\delta\mathbf{x}^H\mathbf{p}]_0^T,$$

Derivation of adjoint II

The adjoint variable $\mathbf{p}$ is introduced as a Lagrange multiplier. The augmented cost function is written

$$J = \int_0^T \frac{1}{2}[\mathbf{x}^H Q\mathbf{x} + l^2 \mathbf{u}^H R\mathbf{u}] - \mathbf{p}^H [\frac{\partial \mathbf{x}}{\partial t} - A\mathbf{x} - B\mathbf{u}]\, dt,$$

linearize + integration by parts and $\delta J = 0$

$$0 = \int_0^T \delta\mathbf{u}^H [B\mathbf{p} + l^2 R\mathbf{u}] + \delta\mathbf{x}^H [\underbrace{\frac{\partial \mathbf{p}}{\partial t} + A^H\mathbf{p} + Q\mathbf{x}}_{=0}]\, dt + [\delta\mathbf{x}^H\mathbf{p}]_0^T,$$

gives adjoint equations (obs! $\delta\mathbf{x}(0) = 0$)

$$\frac{\partial \mathbf{p}}{\partial t} = -A^H\mathbf{p} - Q\mathbf{x}, \quad \text{with} \quad \mathbf{p}(t = T) = 0,$$

Derivation of adjoint III

The adjoint variable $\mathbf{p}$ is introduced as a Lagrange multiplier. The augmented cost function is written

$$J = \int_0^T \frac{1}{2}[\mathbf{x}^H Q\mathbf{x} + l^2\mathbf{u}^H R\mathbf{u}] - \mathbf{p}^H[\frac{\partial \mathbf{x}}{\partial t} - A\mathbf{x} - B\mathbf{u}]\, dt,$$

linearize + integration by parts and $\delta J = 0$

$$0 = \int_0^T \delta\mathbf{u}^H \underbrace{[B\mathbf{p} + l^2 R\mathbf{u}]}_{=0} + \delta\mathbf{x}^H \underbrace{[\frac{\partial \mathbf{p}}{\partial t} + A^H\mathbf{p} + Q\mathbf{x}]}_{=0}\, dt + [\delta\mathbf{x}^H\mathbf{p}]_0^T,$$

gives adjoint equations (obs! $\delta\mathbf{x}(0) = 0$)

$$\frac{\partial \mathbf{p}}{\partial t} = -A^H\mathbf{p} - Q\mathbf{x}, \quad \text{with} \quad \mathbf{p}(t = T) = 0,$$

and optimality condition

$$\mathbf{u} = -\frac{1}{l^2} R^{-1} B^H \mathbf{p}.$$

Optimal control using feedback

If we consider a feedback rule $\mathbf{u} = K\mathbf{x}$ then

$$\mathbf{u} = K\mathbf{x} = -\frac{1}{l^2} R^{-1} B^H \mathbf{p}.$$

This is commonly solved using a linear relation $\mathbf{p} = X\mathbf{x}$ in order to write the system given by the direct and adjoint equations, as one differential equation for X,

Optimal control using feedback

If we consider a feedback rule $\mathbf{u} = K\mathbf{x}$ then

$$\mathbf{u} = K\mathbf{x} = -\frac{1}{l^2} R^{-1} B^H \mathbf{p}.$$

This is commonly solved using a linear relation $\mathbf{p} = X\mathbf{x}$ in order to write the system given by the direct and adjoint equations, as one differential equation for X,

(usually denoted *differential Riccati equation*).

Optimal control using feedback

If we consider a feedback rule $\mathbf{u} = K\mathbf{x}$ then

$$\mathbf{u} = K\mathbf{x} = -\frac{1}{l^2} R^{-1} B^H \mathbf{p}.$$

This is commonly solved using a linear relation $\mathbf{p} = X\mathbf{x}$ in order to write the system given by the direct and adjoint equations, as one differential equation for X,

(usually denoted *differential Riccati equation*).

How does it work ?

Note that state is often denoted direct

Two-point boundary value problem

Write the direct and adjoint equations on a combined matrix form

$$\frac{d\mathbf{z}}{dt} = Z\mathbf{z} \quad \text{where} \quad Z = Z_{2n\times 2n} = \begin{bmatrix} A & -l^{-2}BR^{-1}B^H \\ -Q & -A^H \end{bmatrix} \tag{1}$$

$$\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \end{bmatrix}, \quad \text{and} \quad \begin{cases} \mathbf{x} = \mathbf{x}_0 & \text{at} \quad t = 0, \\ \mathbf{p} = 0 & \text{at} \quad t = T. \end{cases}$$

Two-point boundary value problem

Write the direct and adjoint equations on a combined matrix form

$$\frac{d\mathbf{z}}{dt} = Z\mathbf{z} \quad \text{where} \quad Z = Z_{2n\times 2n} = \begin{bmatrix} A & -l^{-2}BR^{-1}B^H \\ -Q & -A^H \end{bmatrix} \tag{1}$$

$$\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \end{bmatrix}, \quad \text{and} \quad \begin{cases} \mathbf{x} = \mathbf{x}_0 & \text{at} \quad t = 0, \\ \mathbf{p} = 0 & \text{at} \quad t = T. \end{cases}$$

(Z has a Hamiltonian symmetry, such that eigenvalues appear in pairs of equal imaginary and opposite real part.)

Two-point boundary value problem

Write the direct and adjoint equations on a combined matrix form

$$\frac{d\mathbf{z}}{dt} = Z\mathbf{z} \quad \text{where} \quad Z = Z_{2n\times 2n} = \begin{bmatrix} A & -l^{-2}BR^{-1}B^H \\ -Q & -A^H \end{bmatrix} \tag{1}$$

$$\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \end{bmatrix}, \quad \text{and} \quad \begin{cases} \mathbf{x} = \mathbf{x}_0 & \text{at} \quad t = 0, \\ \mathbf{p} = 0 & \text{at} \quad t = T. \end{cases}$$

(Z has a Hamiltonian symmetry, such that eigenvalues appear in pairs of equal imaginary and opposite real part.)

This linear ODE is a two-point boundary value problem and may be solved using a linear relationship between the state vector $\mathbf{x}(t)$ and adjoint vector $\mathbf{p}(t)$ vi a matrix $X(T)$ such that $\mathbf{p} = X\mathbf{x}$, and inserting this solution ansatz into (1) to eliminate $\mathbf{p}$.

The Riccati equation

It follows that matrix X obeys the differential Riccati equation

$$-\frac{dX}{dt} = A^H X + XA - X l^{-2} B R^{-1} B^H X + Q \quad \text{with} \quad X(T) = 0. \tag{2}$$

The Riccati equation

It follows that matrix X obeys the differential Riccati equation

$$-\frac{dX}{dt} = A^H X + XA - Xl^{-2} BR^{-1} B^H X + Q \quad \text{with} \quad X(T) = 0. \tag{2}$$

Once X is known, the optimal value of $\mathbf{u}$ may then be written in the form of a feedback control rule such that

$$\mathbf{u} = K\mathbf{x} \quad \text{where} \quad K = -l^{-2} R^{-1} B^H X.$$

The Riccati equation

It follows that matrix X obeys the differential Riccati equation

$$-\frac{dX}{dt} = A^H X + XA - X l^{-2} B R^{-1} B^H X + Q \quad \text{with} \quad X(T) = 0. \tag{2}$$

Once X is known, the optimal value of $\mathbf{u}$ may then be written in the form of a feedback control rule such that

$$\mathbf{u} = K\mathbf{x} \quad \text{where} \quad K = -l^{-2} R^{-1} B^H X.$$

Finally, if the system is time invariant (LTI) and we take the limit that $T \to \infty$, the matrix X in (2) may be marched to steady state. This steady state solution for X satisfies the continuous-time algebraic Riccati equation

$$0 = A^H X + XA - X l^{-2} B R^{-1} B^H X + Q,$$

where additionally X is constrained such that $A + BK$ is stable.

The classical way of solution

A linear time-invariant system (LTI) can be solved using its eigenvectors.

The classical way of solution

A linear time-invariant system (LTI) can be solved using its eigenvectors. Assume that an eigenvector decomposition of the $2n \times 2n$ matrix Z is available such that

$$Z = V \Lambda_c V^{-1} \quad \text{where} \quad V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \quad \text{and} \quad \mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \end{bmatrix}$$

and the eigenvalues of Z appearing in the diagonal matrix Λ_c are enumerated in order of increasing real part.

The classical way of solution

A linear time-invariant system (LTI) can be solved using its eigenvectors. Assume that an eigenvector decomposition of the $2n \times 2n$ matrix Z is available such that

$$Z = V \Lambda_c V^{-1} \quad \text{where} \quad V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \quad \text{and} \quad \mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \end{bmatrix}$$

and the eigenvalues of Z appearing in the diagonal matrix Λ_c are enumerated in order of increasing real part. Since

$$\mathbf{z} = V e^{\Lambda_c t} V^{-1} \mathbf{z}_0$$

the solutions $\mathbf{z}$ that obey the boundary conditions at $t \to \infty$ are spanned by the first n columns of V.

The classical way of solution

A linear time-invariant system (LTI) can be solved using its eigenvectors. Assume that an eigenvector decomposition of the $2n \times 2n$ matrix Z is available such that

$$Z = V \Lambda_c V^{-1} \quad \text{where} \quad V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \quad \text{and} \quad \mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \end{bmatrix}$$

and the eigenvalues of Z appearing in the diagonal matrix Λ_c are enumerated in order of increasing real part. Since

$$\mathbf{z} = V e^{\Lambda_c t} V^{-1} \mathbf{z}_0$$

the solutions $\mathbf{z}$ that obey the boundary conditions at $t \to \infty$ are spanned by the first n columns of V. The direct ($\mathbf{x}$) and adjoint ($\mathbf{p}$) parts of the these columns are related as $\mathbf{p} = X\mathbf{x}$, where

$$[\mathbf{p}_1, \, \mathbf{p}_2, \cdots, \mathbf{p}_n] = X[\mathbf{x}_1, \, \mathbf{x}_2, \cdots, \mathbf{x}_n] \quad \to \quad X = V_{21} V_{11}^{-1}$$

Motivation

- Optimal control via application of modern control algorithms (Riccati equation) is intractable because of the very large number of degrees of freedom deriving from the discretization of the Navier-Stokes equations.

Motivation

- Optimal control via application of modern control algorithms (Riccati equation) is intractable because of the very large number of degrees of freedom deriving from the discretization of the Navier-Stokes equations. $X = X_{n \times n}$

Motivation

- Optimal control via application of modern control algorithms (Riccati equation) is intractable because of the very large number of degrees of freedom deriving from the discretization of the Navier-Stokes equations. $X = X_{n \times n}$

- One common approach is to use reduced-order models (ROM).

Motivation

- Optimal control via application of modern control algorithms (Riccati equation) is intractable because of the very large number of degrees of freedom deriving from the discretization of the Navier-Stokes equations. $X = X_{n \times n}$

- One common approach is to use reduced-order models (ROM).

- Here, we present two exact methods which do not rely on such modeling,

  where $$J = \frac{1}{2} \int_0^T [\mathbf{x}^H Q \mathbf{x} + l^2 \mathbf{u}^H R \mathbf{u}] \, dt$$

Motivation

- Optimal control via application of modern control algorithms (Riccati equation) is intractable because of the very large number of degrees of freedom deriving from the discretization of the Navier-Stokes equations. $X = X_{n\times n}$

- One common approach is to use reduced-order models (ROM).

- Here, we present two exact methods which do not rely on such modeling,

  where $$J = \frac{1}{2}\int_0^T [\mathbf{x}^H Q \mathbf{x} + l^2 \mathbf{u}^H R \mathbf{u}]\, dt$$

  1. In the limit that $l^2 \to \infty$,

Motivation

- Optimal control via application of modern control algorithms (Riccati equation) is intractable because of the very large number of degrees of freedom deriving from the discretization of the Navier-Stokes equations. $X = X_{n\times n}$

- One common approach is to use reduced-order models (ROM).

- Here, we present two exact methods which do not rely on such modeling,

$$\text{where} \qquad J = \frac{1}{2}\int_0^T \left[\mathbf{x}^H Q\mathbf{x} + l^2 \mathbf{u}^H R\mathbf{u}\right] dt$$

  1. In the limit that $l^2 \to \infty$, MCE: Minimal Control Energy

Motivation

- Optimal control via application of modern control algorithms (Riccati equation) is intractable because of the very large number of degrees of freedom deriving from the discretization of the Navier-Stokes equations. $X = X_{n\times n}$
- One common approach is to use reduced-order models (ROM).
- Here, we present two exact methods which do not rely on such modeling,

  where $$J = \frac{1}{2}\int_0^T [\mathbf{x}^H Q\mathbf{x} + l^2\mathbf{u}^H R\mathbf{u}]\,dt$$

  1. In the limit that $l^2 \to \infty$, MCE: Minimal Control Energy
  2. For any value of l^2, more general,

Motivation

- Optimal control via application of modern control algorithms (Riccati equation) is intractable because of the very large number of degrees of freedom deriving from the discretization of the Navier-Stokes equations. $X = X_{n\times n}$
- One common approach is to use reduced-order models (ROM).
- Here, we present two exact methods which do not rely on such modeling,

  where $$J = \frac{1}{2}\int_0^T [\mathbf{x}^H Q\mathbf{x} + l^2\mathbf{u}^H R\mathbf{u}]\, dt$$

  1. In the limit that $l^2 \to \infty$, MCE: Minimal Control Energy
  2. For any value of l^2, more general, ADA: Adjoint of the Direct-Adjoint

Minimal Control Energy

Minimal-energy control feedback I

In the limit that $l^2 \to \infty$ we consider

$$J = \int_0^T \frac{1}{2}[l^{-2}\mathbf{x}^H Q\mathbf{x} + \mathbf{u}^H R\mathbf{u}]$$

With this defintion the same derivation as before leads to

$$\frac{d\mathbf{z}}{dt} = Z\mathbf{z} \quad \text{where} \quad Z = Z_{2n\times 2n} = \begin{bmatrix} A & -BR^{-1}B^H \\ -l^{-2}Q & -A^H \end{bmatrix}$$

Taking the limit $l^2 \to \infty$ we get

Minimal-energy control feedback II

In the limit that $l^2 \to \infty$ we consider

$$J = \int_0^T \frac{1}{2}[l^{-2}\mathbf{x}^H Q\mathbf{x} + \mathbf{u}^H R\mathbf{u}]$$

With this defintion the same derivation as before leads to

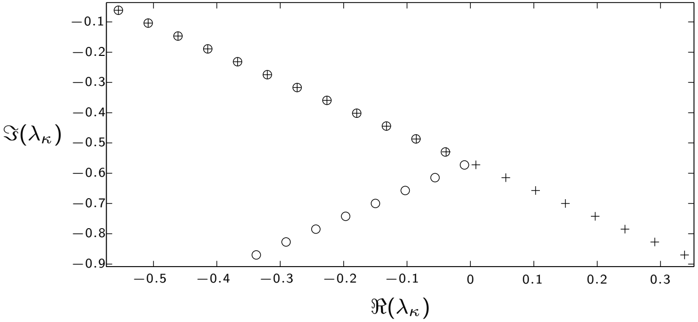
$$\frac{d\mathbf{z}}{dt} = Z\mathbf{z} \quad \text{where} \quad Z = Z_{2n\times 2n} = \begin{bmatrix} A & -BR^{-1}B^H \\ 0 & -A^H \end{bmatrix}$$

Z becomes block triangular. The direct and adjoint equations are

$$\frac{\partial \mathbf{x}}{\partial t} = A\mathbf{x} + B\mathbf{u}, \qquad \mathbf{u} = -R^{-1}B^H\mathbf{p}, \qquad \frac{\partial \mathbf{p}}{\partial t} = -A^H\mathbf{p} + 0$$

Minimal-energy control feedback III

The eigenvalues of this system is given by the union of the eigenvalues of A and $-A^H$.



The eigenvalues of (+) the discretized open-loop system, and (o) the closed-loop system $A + BK$ after minimal-energy control is applied.

Minimal-energy control feedback IV

Here we know the eigenvalues and only need to compute

$$X = V_{21} V_{11}^{-1}$$

It can be shown that X is only function of V_{21}. K is finally given as a function of the unstable eigenvalues and corresponding
left eigenvectors.

$$K = -B^H T_u F^{-1} T_u^H$$

where F has elements

$$f_{ij} = c_{ij} / (\lambda_i + \lambda_j^*)$$

and

$$C = T_u^H B B^H T_u$$

T_u is the matrix containing unstable left eigenvectors

ADA: Adjoint of the Direct-Adjoint

Riccati-less optimal control I

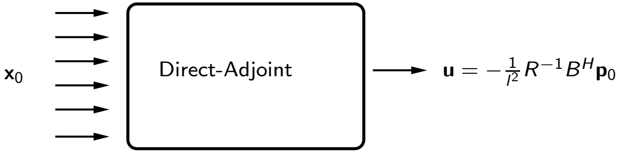
The aim is to compute the solution for K, which is independent of x_0 and time invariant. This can be solved using an iterative procedure to "try" different $\mathbf{x}_0$ (computationally expensive).

ALTERNATIVELY

For a converged solution at $t = 0$ we can write

$$\mathbf{u} = K\mathbf{x}_0 = -\frac{1}{l^2} R^{-1} B^H \mathbf{p}_0.$$

This is a linear relation between the input $\mathbf{x}_0$ and output $\mathbf{u}$.

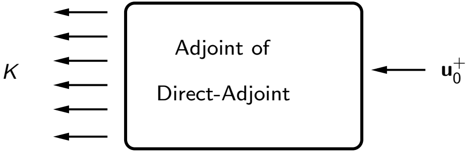


The input has a large dimension and the output a small dimension.

Riccati-less optimal control II

Such a problem is efficiently solved using the adjoint equations.

The adjoint input has a small dimension and the output a large dimension.



K is obtained from the solution of the adjoint of the direct-adjoint system.

Adjoint of the Direct-Adjoint system I

Introduce the adjoint variables $\mathbf{x}^+$ and $\mathbf{p}^+$ and multiply with the direct-adjoint equations, then integrate in time from $t = 0$ to $t = T$. Obs! here we consider that $\mathbf{u}$ has dimension $m = 1$.

$$\int_0^T \mathbf{x}^{+H} \left(\frac{\partial \mathbf{x}}{\partial t} - A\mathbf{x} + \frac{1}{l^2} B R^{-1} B^H \mathbf{p} \right) dt + \int_0^T \mathbf{p}^{+H} \left(\frac{\partial \mathbf{p}}{\partial t} + A^H \mathbf{p} + Q\mathbf{x} \right) dt = 0.$$

Adjoint of the Direct-Adjoint system II

Using integration by parts, and considering that both R and Q are symmetric, we obtain

$$-\int_0^T \mathbf{p}^H \left(\frac{\partial \mathbf{p}^+}{\partial t} - A\mathbf{p}^+ - \frac{1}{l^2} B R^{-1} B^H \mathbf{x}^+ \right) dt - \int_0^T \mathbf{x}^H \left(\frac{\partial \mathbf{x}^+}{\partial t} + A^H \mathbf{x}^+ - Q\mathbf{p}^+ \right) dt$$

$$+ \left[\mathbf{p}^H \mathbf{p}^+ \right]_0^T + \left[\mathbf{x}^H \mathbf{x}^+ \right]_0^T = 0.$$

If we now define the new adjoint equations as

Adjoint of the Direct-Adjoint system III

Using integration by parts, and considering that both R and Q are symmetric, we obtain

$$-\int_0^T \mathbf{p}^H \left(\underbrace{\frac{\partial \mathbf{p}^+}{\partial t} - A\mathbf{p}^+ - \frac{1}{l^2} BR^{-1}B^H\mathbf{x}^+}_{=0} \right) dt - \int_0^T \mathbf{x}^H \left(\underbrace{\frac{\partial \mathbf{x}^+}{\partial t} + A^H\mathbf{x}^+ - Q\mathbf{p}^+}_{=0} \right) dt$$
$$+ \left[\mathbf{p}^H \mathbf{p}^+\right]_0^T + \left[\mathbf{x}^H \mathbf{x}^+\right]_0^T = 0.$$

If we now define the new adjoint equations as

$$\frac{\partial \mathbf{p}^+}{\partial t} = A\mathbf{p}^+ + \frac{1}{l^2} BR^{-1}B^H\mathbf{x}^+,$$
$$\frac{\partial \mathbf{x}^+}{\partial t} = -A^H\mathbf{x}^+ + Q\mathbf{p}^+,$$

Adjoint of the Direct-Adjoint system IV

with $\mathbf{x}^{+}(t = T) = 0$ and $\mathbf{p}(t = T) = 0$, the remaining terms are

$$\mathbf{x}^{+H}(0)\mathbf{x}(0) + \mathbf{p}^{+H}(0)\mathbf{p}(0) = 0.$$

Recall that the original linear relation was

$$K\mathbf{x}_0 = -\frac{1}{l^2}R^{-1}B^H\mathbf{p}_0$$

- Choosing $\mathbf{p}^{+H}(t = 0)$ as one row of $-\frac{1}{l^2}R^{-1}B^H$ $(m = 1)$
- we can identify one row of K as $\mathbf{x}^{+H}(0)$. $(m = 1)$

Riccati-less optimal control: solution procedure

If we let $\mathbf{x}^+ \to -\mathbf{p}$ and $\mathbf{p}^+ \to \mathbf{x}$ we easily obtain the original (Direct-Adjoint) system. (self-adjoint)

Finally: solve the original linear system with new b.c.

$$\frac{\partial \mathbf{x}}{\partial t} = A\mathbf{x} - \frac{1}{l^2} BR^{-1}B^H\mathbf{p} \quad \text{on} \quad 0 < t < T, \quad \mathbf{x}^H(0) \quad \text{is one row of} \quad \frac{1}{l^2}R^{-1}B^H,$$
$$\frac{\partial \mathbf{p}}{\partial t} = -A^H\mathbf{p} - Q\mathbf{x} \quad \text{on} \quad 0 < t < T, \quad \text{with} \quad \mathbf{p}(T) = 0.$$

One row of K is then given by $-\mathbf{p}^H(0)$ (since $\mathbf{x}^+ = -\mathbf{p}$).

Riccati-less optimal control: solution procedure

If we let $\mathbf{x}^+ \to -\mathbf{p}$ and $\mathbf{p}^+ \to \mathbf{x}$ we easily obtain the original (Direct-Adjoint) system. (self-adjoint)

Finally: solve the original linear system with new b.c.

$$\frac{\partial \mathbf{x}}{\partial t} = A\mathbf{x} - \frac{1}{l^2} B R^{-1} B^H \mathbf{p} \quad \text{on} \quad 0 < t < T, \quad \mathbf{x}^H(0) \quad \text{is one row of} \quad \frac{1}{l^2} R^{-1} B^H,$$
$$\frac{\partial \mathbf{p}}{\partial t} = -A^H \mathbf{p} - Q\mathbf{x} \quad \text{on} \quad 0 < t < T, \quad \text{with} \quad \mathbf{p}(T) = 0.$$

One row of K is then given by $-\mathbf{p}^H(0)$ (since $\mathbf{x}^+ = -\mathbf{p}$).

IMPORTANT

Avoid solving $X_{n \times n}$

Riccati-less optimal control: solution procedure

If we let $\mathbf{x}^+ \rightarrow -\mathbf{p}$ and $\mathbf{p}^+ \rightarrow \mathbf{x}$ we easily obtain the original (Direct-Adjoint) system. (self-adjoint)

Finally: solve the original linear system with new b.c.

$$\frac{\partial \mathbf{x}}{\partial t} = A\mathbf{x} - \frac{1}{l^2} B R^{-1} B^H \mathbf{p} \quad \text{on} \quad 0 < t < T, \quad \mathbf{x}^H(0) \quad \text{is one row of} \quad \frac{1}{l^2} R^{-1} B^H,$$

$$\frac{\partial \mathbf{p}}{\partial t} = -A^H \mathbf{p} - Q\mathbf{x} \quad \text{on} \quad 0 < t < T, \quad \text{with} \quad \mathbf{p}(T) = 0.$$

One row of K is then given by $-\mathbf{p}^H(0)$ (since $\mathbf{x}^+ = -\mathbf{p}$).

IMPORTANT

Avoid solving $X_{n\times n}$ solve original system $\mathbf{x}_{n\times 1}$ m times

Applications

Applications

MCE: Minimal Control Energy

In this method the feedback matrix K is evaluated from the **unstable** open-loop solutions of the system.

Case: control of the cylinder wake (globally unstable flow)

Refs: *Carini, Pralits, Luchini, JFS, 2013*

ADA: Adjoint of the Direct-Adjoint

This method is more general and does not depend on whether the system is unstable or not.

Cases: control of the cylinder wake, boundary layer transition

Refs: *Pralits, Luchini, IUTAM Proceeding, 2010,*
Semeraro, Pralits, Rowley, Henningson, JFM, 2013

Control of the cylinder wake

Control strategy
MCE & ADA

Numerical procedure

- All equations are discretized using second-order finite-differences over a staggered, stretched, Cartesian mesh.
- An immersed-boundary technique is used to enforce the boundary conditions on the cylinder.
- The nonlinear mean-flow equations, along with their boundary conditions, are solved by a Newton-Raphson procedure.
- The linear and nonlinear evolution equations are solved using Adams-Bashforth/Crank-Nicholson.
- The eigenvalue problems are solved using an Inverse Iteration algorithm
- Discrete adjoint equations (accurate to machine precision).

Cases:
Reynolds numbers close to the first bifurcation, two-dimensional flow

MCE

The linear feedback matrix K which suppresses vortex shedding from a circular cylinder has been computed using:

MCE

The linear feedback matrix K which suppresses vortex shedding from a circular cylinder has been computed using:
Full state information,

MCE

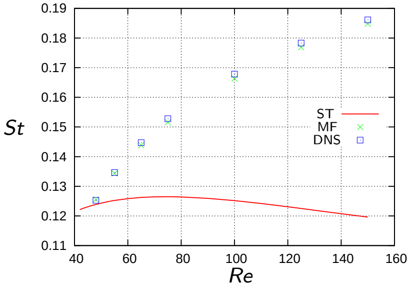
The linear feedback matrix K which suppresses vortex shedding from a circular cylinder has been computed using:
Full state information, Actuator: angular oscillation,

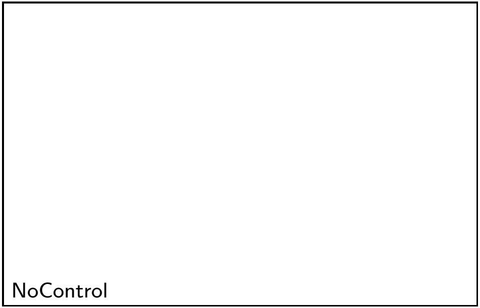
MCE

The linear feedback matrix K which suppresses vortex shedding from a circular cylinder has been computed using:
Full state information, Actuator: angular oscillation, $Re = UD/\nu$

MCE

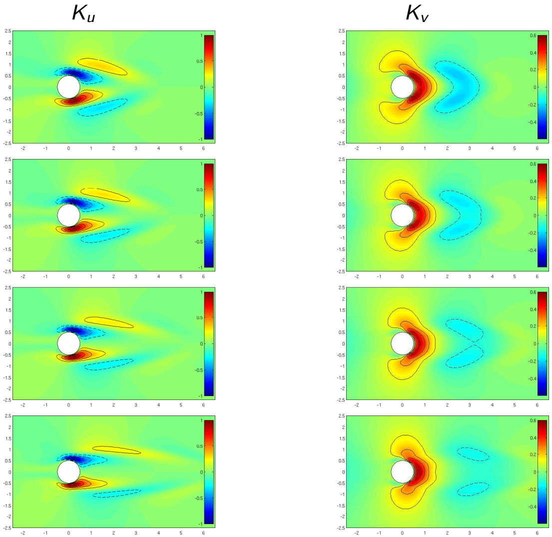
The linear feedback matrix K which suppresses vortex shedding from a circular cylinder has been computed using:
Full state information, Actuator: angular oscillation, $Re = UD/\nu$
Dimension of control $\mathbf{u}$ is $m = 1$





The feedback matrix K ($\mathbf{u} = K\mathbf{x}$)

- $Re = 55$
- $Re = 75$
- $Re = 100$
- $Re = 150$

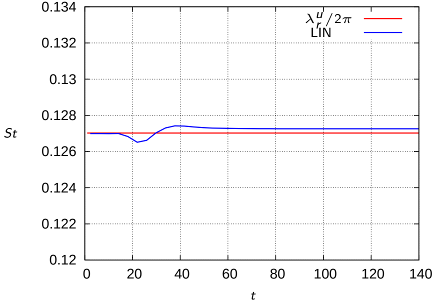


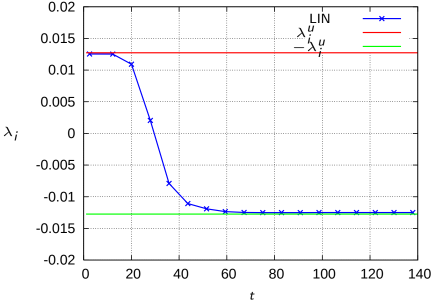
Results: linearized N-S equations

The temporal evolution of the frequency and growth rate is compared with the eigenvalue λ

- The Strouhal number: $St = fD/U$ compared to $St = \lambda_r/2\pi$
- The growth rate: $\sigma = \frac{d}{dt} log(u(t))$ compared to λ_i

Test case: $Re = 55$, control is turned on at $t = 18$





Control of vortex shedding: $Re = 55$

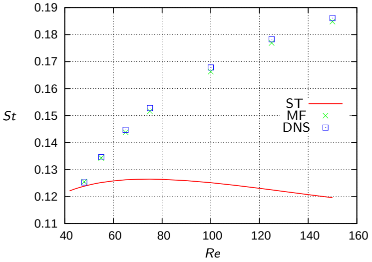
WithControl

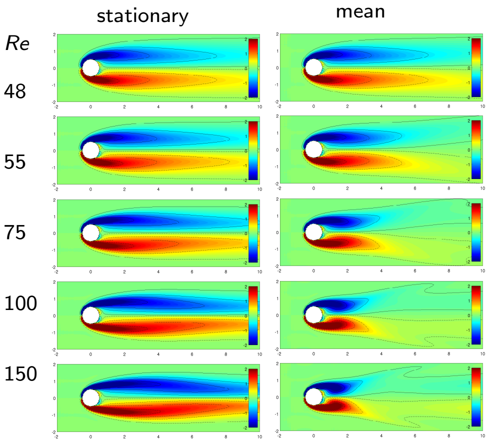
Control of vortex shedding: $Re = 55$

WithControl

Stationary vs. mean flow

St for limit cycle coincide with mean-flow eigenfrequency

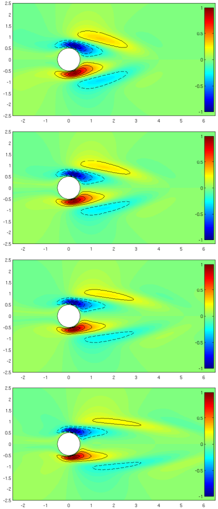




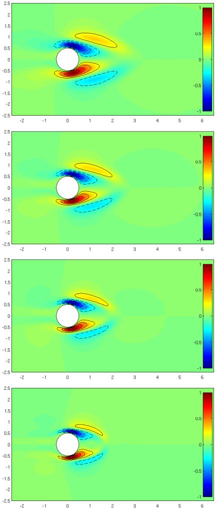
K_u stationary vs. mean

- $Re = 55$
- $Re = 75$
- $Re = 100$
- $Re = 150$

K_u stationary



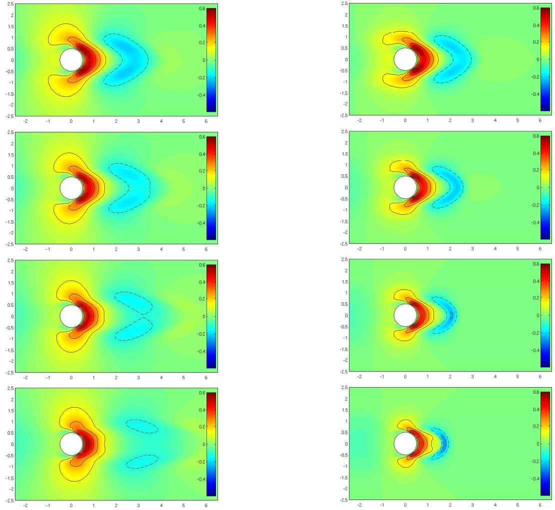
K_u mean



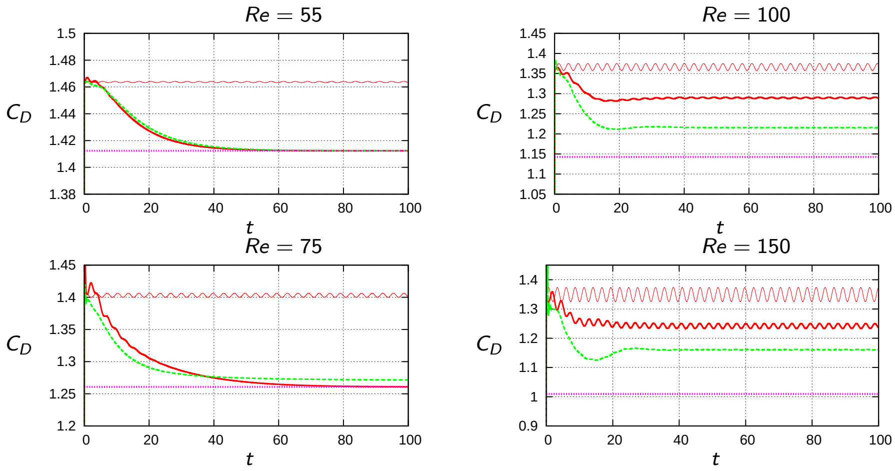
K_v stationary vs. mean

- $Re = 55$
- $Re = 75$
- $Re = 100$
- $Re = 150$

K_v stationary K_v mean



Control of vortex shedding: stationary vs. mean



Red: stationary flow, Green: mean flow

ADA

The linear feedback matrix K which suppresses vortex shedding from a circular cylinder has been computed using:

ADA

The linear feedback matrix K which suppresses vortex shedding from a circular cylinder has been computed using:
Full state information,

ADA

The linear feedback matrix K which suppresses vortex shedding from a circular cylinder has been computed using:

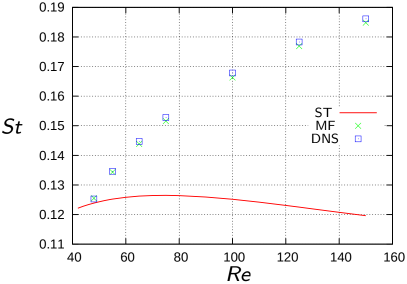
Full state information, Actuator: angular oscillation,

ADA

The linear feedback matrix K which suppresses vortex shedding from a circular cylinder has been computed using:
Full state information, Actuator: angular oscillation, $Re = UD/\nu$

ADA

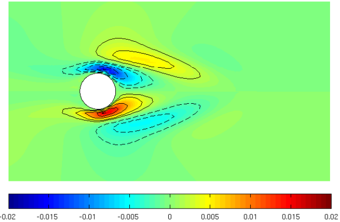
The linear feedback matrix K which suppresses vortex shedding from a circular cylinder has been computed using:
Full state information, Actuator: angular oscillation, $Re = UD/\nu$
Dimension of control $\mathbf{u}$ is $m = 1$



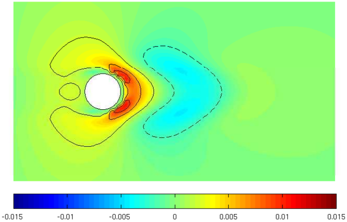
NoControl

Results: K for $Re = 55$

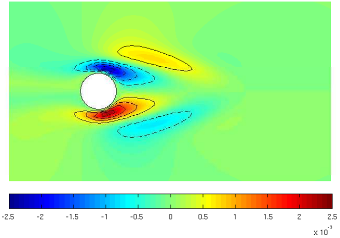
$K_u,\ l^2 = 1$



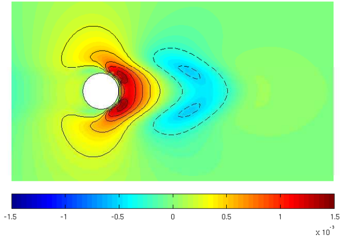
$K_v,\ l^2 = 1$



$K_u,\ l^2 \to \infty$



$K_u,\ l^2 \to \infty$

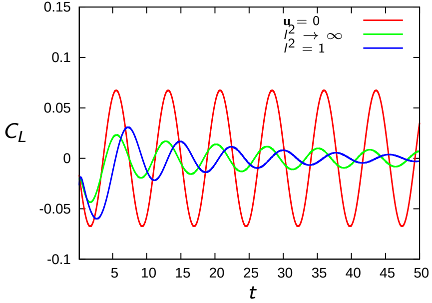


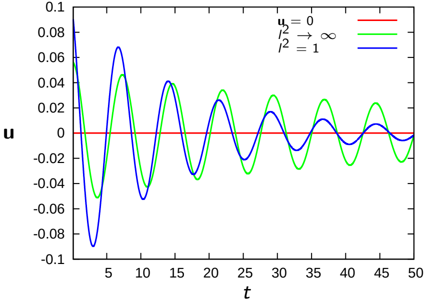
Control of vortex shedding

In the temporal evolution of the lift (C_L) and control $\mathbf{u}$:

- C_L and $\mathbf{u}$ tend to zero as the control is applied
- Control $\mathbf{u}$ strengthens as l^2 decrease

Test case: $Re = 55$, control is turned on at $t = 0$



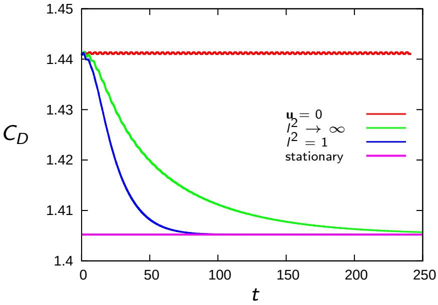


Control of vortex shedding

In the temporal evolution of drag (C_D) coefficient:

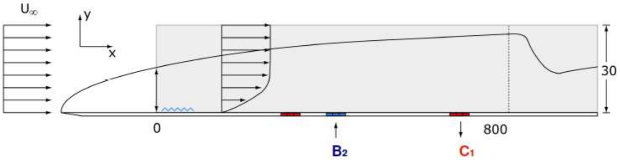
- As the control is applied C_D tends to the constant value corresponding to the steady state solution
- The control acts more quickly as l^2 is decreased

Test case: $Re = 55$, control is turned on at $t = 0$



Control of the flat plate boundary layer I

Linear quadratic controller (LQR)



$$\frac{d\mathbf{q}}{dt} = \mathbf{A}\mathbf{q} + \mathbf{B}_2\mathrm{u}$$

$$\mathrm{z} = \begin{bmatrix} \mathbf{C}_1 \\ \mathbf{0} \end{bmatrix}\mathbf{q} + \begin{bmatrix} 0 \\ l \end{bmatrix}\mathrm{u}$$

B2 – actuator

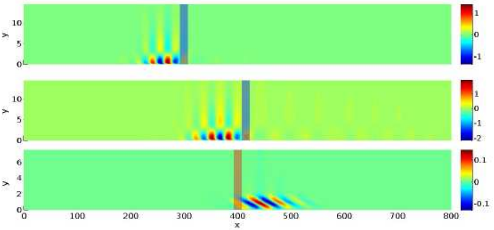
C1 – objective function

$$\mathcal{J} = \frac{1}{2}\mathrm{z}^H\mathrm{z} = \frac{1}{2}\int_0^T \left(\mathbf{q}^H\mathbf{C}_1^H\mathbf{C}_1\mathbf{q} + \mathrm{u}^H\mathbf{R}\mathrm{u}\right) dt$$

Cost function – to be minimised

Control of the flat plate boundary layer II

Full dimensional gains



Gain L, x = 300
Sensor upstream

Gain L, x = 410
Sensor downstream

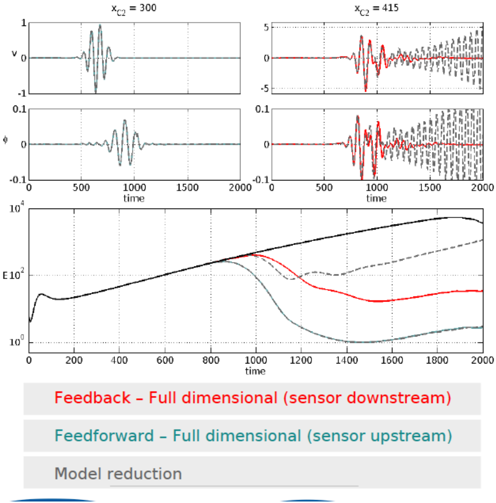
Gain K, x = 400

Streamwise component

Estimation gain located upstream of the sensor (forward solution, AAD)

Control gain located downstream of the actuator (adjoint solution, ADA)

Control of the flat plate boundary layer III



Semeraro, Pralits, Rowley, Henningson, JFM, 2013

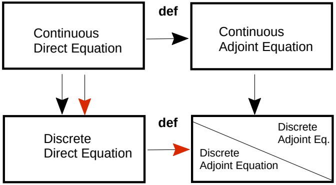
Some numerical issues

Continuous vs. Discrete Adjoint Equations

The adjoint equations can be derived using two different approaches.
Both with advantages and disadvantages.
By definition we have

$$\langle p, Lx \rangle =$$

- **Continuous approach** →: The adjoint eq
derived by definition using the continuou
equations.
  - \+ Straightforward derivation, reuse old code
programming
  - − Accuracy depends on discretization, diffic
boundary conditions
- **Discrete approach** →: The adjoint equations are derived from the discretized direct equations.
  - \+ Accuracy can be achieved close to machine precision, and can be independent of discretization !!
  - − Tricky derivation, usually requires making a new code, or larger changes of an existing code.



Here "def" means definition of the adjoint operator.
In the top row it is on continuous form while in the bottom row it is on discrete form.

Derivation of the adjoint equation I

Consider the following optimal control problem (ODE) where ϕ is the state and g the control.

$$\frac{d\phi(t)}{dt} = -A\phi(t) + Bg(t), \qquad \text{for} \qquad 0 \leq t \leq T,$$

with initial condition

$$\phi(0) = \phi_0$$

We can now define an optimization problem in which the goal is to find an optimal $g(t)$ by minimizing the following objective function

$$J = \frac{\gamma_1}{2}[\phi(T) - \Psi]^2 + \frac{\gamma_2}{2}\int_0^T g(t)^2 dt,$$

Derivation of the adjoint equation II

Continuous approach

We can solve this problem using an **adjoint identity** approach or by introducing **Lagrange multipliers**.

$$\int_0^T a[\frac{d\phi}{dt} + A\phi - Bg]\,dt = \int_0^T [-\frac{da}{dt} + A^* a]\phi\,dt - \int_0^T aBg\,dt + a(T)\phi(T) - a(0)\phi(0).$$

If we now define the adjoint equation as $-da/dt = -A^* a$ with an arbitrary initial condition $a(T)$ then the identity reduces to

$$\text{LHS} = -\int_0^T aBg\,dt + a(T)\phi(T) - a(0)\phi(0)$$

By definition the Left Hand Side is identically zero but this is exactly what must be checked numerically, i.e. error$= |\text{LHS}|$.

Derivation of the adjoint equation III

The gradient of J w.r.t. g can be derived considering the J is nonlinear in ϕ and g. We linearise by $\phi \to \phi + \delta\phi$, $g \to g + \delta g$ and then write the linearised objective function as

$$\gamma_1[\phi(T) - \Psi]\delta\phi(T) = \delta J - \gamma_2 \int_0^T g\delta g \, dt,$$

If we choose $a(T) = \gamma_1[\phi(T) - \Psi]$ then the equation for δJ can be substituted into the expression for the adjoint identity. If you further define the adjoint equations, remember that $\delta\phi(0) = 0$, then the final identity is written

$$\delta J = \int_0^T [\gamma_2 g + B^* a]\delta g \, dt$$

The adjoint equations and gradient of J w.r.t. g are written

$$-\frac{da}{dt} + A^* a, \; a(T) = \gamma_1[\phi(T) - \Psi], \quad \text{and} \quad \nabla J_g = \gamma_2 g + B^* a.$$

The so called optimality condition is given by $\nabla J_g = 0$.

Derivation of the adjoint equation IV

- The accuracy of the adjoint solution is important since it quantfies a "gradient" in the optimization problem.
- The "error" must be evaluated to quantify the accuracy the adjoint solution.
- Note that the adjoint solution depends on the resolution (Δt), and likewise the accuracy.
- **Can we do better ?**

Derivation of the adjoint equation V

Discrete approach

A discrete version of the direct equation is written

$$\frac{\phi^{i+1}-\phi^i}{\Delta t} = -A\phi^i + Bg^i, \quad \text{for} \quad i = 1, ..., N-1,$$

where N denotes the number of discrete points on the interval $[0, T]$, Δt is the constant time step, and

$$\phi^1 = \phi_0,$$

is the initial condition. This can be written as a discrete evolution equation

$$\phi^{i+1} = [I - \Delta t A]\phi^i + \Delta t B g^i, \quad \text{for} \quad i = 1, ..., N-1.$$

A discrete version of the objective function can be written

$$J = \frac{\gamma_1}{2}(\phi^N - \Phi)^2 + \frac{\gamma_2}{2}\sum_{i=1}^{N-1}\Delta t (g^i)^2.$$

An adjoint variable a^i is introduced defined on $i = 1, ..., N$ and by definition

$$a^{i+1} \cdot L\phi^i = (L^\star a^{i+1}) \cdot \phi^i, \quad \text{for} \quad i = 1, ..., N-1.$$

We then introduce the definition of the state equation on the left hand side of and impose that

$$a^i = L^\star a^{i+1} \quad \text{for} \quad i = N-1, ..., 1.$$

This is the discrete adjoint equation. Using the discrete direct and adjoint yields

$$a^{i+1} \cdot (\phi^{i+1} - \Delta t B g^i) = a^i \cdot \phi^i, \quad \text{for} \quad i = 1, ..., N-1$$

which must be valid for any ϕ and a. An error can therefore be written as

$$\text{error} = |a^N \cdot \phi^N - a^1 \cdot \phi^1 - \sum_{i=1}^{N-1} \Delta t\, a^{i+1} \cdot B g^i|.$$

Derivation of the adjoint equation VI

The discrete optimality condition is then derived. Since J is nonlinear with respect to ϕ and g we must first linearize. This can be written

$$\delta J = \gamma_1(\phi^N - \Phi) \cdot \delta\phi^N + \gamma_2 \sum_{i=1}^{N-1} \Delta t g^i \cdot \delta g^i.$$

We now choose the terminal condition of the adjoint as $a^N = \gamma_1(\phi^N - \Phi)$ and substitute this expression into the discrete adjoint identity. This is written

$$\gamma_1(\phi^N - \Phi) \cdot \delta\phi^N = a^1 \cdot \delta\phi^1 + \sum_{i=1}^{N-1} \Delta t \, a^{i+1} \cdot B\delta g^i$$

By inspection one can see that the left hand side is identical to the first term in the expression for δJ, and $\delta\phi^1 = 0$. Rearranging the terms, we get

$$\delta J = \sum_{i=1}^{N-1} \Delta t \left(\gamma_2 g^i + B^\star a^{i+1}\right) \cdot \delta g^i,$$

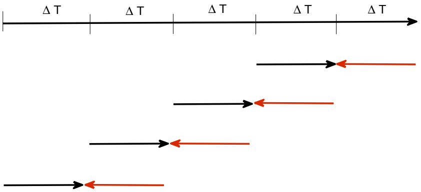
from which we get the discrete optimality condition

$$g^i = -\frac{1}{\gamma_2} B^\star a^{i+1} \quad \text{for} \quad i = 1, .., N-1.$$

Note that if B is a matrix then $B^\star = B^T$.

Checkpointing algorithm

- When the adjoint equation is forced in time by the direct solution (ex. quadratic objective function), then this poses storage requirements (hard ware). This becomes a problem for 2D and 3D problems with high resolution in space and time.
- One way to come around this is to apply Checkpointing. This consists of sampling the direct solution at given rate and then recompute the direct solution for short time intervals when needed. This means in theory that one more solution of the direct system has been added to the computational effort.
- However, since it is common to use parallel computing, and processors is becoming a smaller problem on can do something to obtain the minimal required computational time.
- This is done by rec :omputing the adjoint.



EXTRA SLIDES

Background: control using rotational oscillation

Aim: reduce C_D

Exp. Tokumaru & Dimotakis (1991), -20%, $Re = 15000$

Feedback control:

Exp. Fujisawa & Nakabayashi (2002) -16% (-70% C_L), $Re = 20000$

Exp. Fujisawa et al.(2001) "reduction", $Re = 6700$

Optimal control (using adjoints):

Num. He et al.(2000) -30 to -60% for $Re = 200 - 1000$

Num. Protas & Styczek (2002) -7% at $Re = 75$, -15% at $Re = 150$

Bergmann et al.(2005) -25% at $Re = 200$ (POD)

Aim: reduce vortex shedding

Feedback control:

Num. Protas (2004) reduction, "point vortex model", $Re = 75$

Optimal control (using adjoints):

Num. Homescu et al.(2002) reduction, $Re = 60 - 1000$

Minimal-energy control feedback

Denoting:

- $\mathbf{x}^i$ and λ^i the i-th right eigenvector and eigenvalue of A,
- $\mathbf{y}^i$ and $-\lambda^{i*}$ the i-th right eigenvector and eigenvalue of $-A^H$,
- $\mathbf{y}^{i*}$ is left eigenvector of A,

we see that the stable eigenvectors of

$$\frac{\partial \mathbf{x}}{\partial t} = A\mathbf{x} + B\mathbf{u}, \qquad \mathbf{u} = -R^{-1}B^H\mathbf{p}, \qquad \frac{\partial \mathbf{p}}{\partial t} = -A^H\mathbf{p}$$

are of two possible types:

$$\begin{array}{lll} \mathbf{p} = 0,\ \mathbf{x} = \mathbf{x}^i & \text{if} \quad \Re(\lambda^i) < 0 & \text{(stable)} \\ \mathbf{p} = \mathbf{y}^i,\ \mathbf{x} = (\lambda^{i*} + A)^{-1}BR^{-1}B^H\mathbf{y}^i & \text{if} \quad \Re(\lambda^i) > 0 & \text{(unstable)} \end{array}$$

We now project an arbitrary initial condition $\mathbf{x}_0$ onto these modes,

$$\mathbf{x}_0 = \sum_{\text{stable}} d_j \mathbf{x}^j + \sum_{\text{unstable}} f_j (\lambda^{j*} + A)^{-1} B R^{-1} B^H \mathbf{y}^j \tag{4}$$

and note that in order to reconstruct $\mathbf{p}$ we only need the f_j's, because the stable modes have $\mathbf{p} = 0$. The coefficients d_j can be eliminated from (4) by projecting the left eigenvectors:

$$\mathbf{y}^{i*}\mathbf{x}_0 = \mathbf{y}^{i*} \sum_{\text{unstable}} f_j (\lambda^{j*} + A)^{-1} B R^{-1} B^H \mathbf{y}^j = \sum_{\text{unstable}} c_{ij} f_j$$

where, since $\mathbf{y}^{i*}$ is also a left eigenvector of $(\lambda^{j*} + A)^{-1}$,

$$c_{ij} = \frac{\mathbf{y}^{i*} B R^{-1} B^H \mathbf{y}^j}{\lambda^i + \lambda^{j*}}$$

Only the unstable eigenvalues and left eigenvectors are needed.

The main theorem

Summarizing, the solution of the minimal-energy stabilizing control feedback problem can be written in terms of the unstable left eigenvectors only.

Theorem 1. *Consider a stabilizable system* $\dot{x} = A\mathbf{x} + B\mathbf{u}$ *with no pure imaginary open-loop eigenvalues. Determine the unstable eigenvalues and corresponding left eigenvectors of* A *such that* $T_u^H A = \Lambda_u T_u^H$ *(equivalently, determine the unstable eigenvalues and corresponding right eigenvectors of* A^H *such that* $A^H T_u = T_u \Lambda_u^H$*). Define* $\bar{B}_u = T_u^H B$ *and* $C = \bar{B}_u \bar{B}_u^H$*, and compute a matrix* F *with elements* $f_{ij} = c_{ij}/(\lambda_i + \lambda_j^*)$*. The minimal-energy stabilizing feedback controller is then given by* $\mathbf{u} = K\mathbf{x}$*, where* $K = -\bar{B}_u^H F^{-1} T_u^H$*.*