Controllability of linear parabolic equations and systems

Franck Boyer

Controllability of linear parabolic equations and systems

Franck Boyer
franck.boyer@math.univ-toulouse.fr

Université Paul Sabatier - Toulouse 3

February 7, 2020

Contents

Chapter I

Introduction

Disclaimer : Those lecture notes were written to support a Master course given by the author at Toulouse between 2016 and 2018. Since then, they were regularly updated but are still far from being complete and many references of the literature are lacking (I promise they will be added in the next release !).

It still contains almost surely many mistakes, inaccuracies or typos. Any reader is encouraged to send me[1] any comments or suggestions.

I.1 What is it all about ?

We shall consider a very unprecise setting for the moment : consider a (differential) dynamical system

$$\begin{cases} y' = F(t, y, v(t)), \\ y(0) = y_0, \end{cases} \tag{I.1}$$

in which the user can act on the system through the input v. Here, y (resp. v) live in a state space E (resp. a control space U) which are finite dimensional spaces (the ODE case) or in infinite dimensional spaces (the PDE case).

We assume (for simplicity) that the functional setting is such that (I.1) is globally well-posed for any initial data y_0 and any control v in a suitable functional space.

Definition I.1.1

Let $y_0 \in E$. We say that:

- (I.1) *is* **exactly controllable from** y_0 *if : for any $y_T \in E$, there exists a control $v : (0, T) \to U$ such that the corresponding solution y_{v,y_0} of* (I.1) *satisfies*

$$y_{v,y_0}(T) = y_T.$$

  If this property holds for any y_0, we simply say that the system is **exactly controllable**.

[1] franck.boyer@math.univ-toulouse.fr

- (I.1) *is* **approximately controllable from** y_0 *if : for any* $y_T \in E$*, and any* $\varepsilon > 0$*, there exists a control* $v : (0,T) \to U$ *such that the corresponding solution* y_{v,y_0} *of* (I.1) *satisfies*

$$\|y_{v,y_0}(T) - y_T\|_E \leq \varepsilon.$$

If this property holds for any y_0*, we simply say that the system is* **approximately controllable**.

- (I.1) *is* **controllable to the trajectories from** y_0 *if : for any* $\bar{y}_0 \in E$*, and any* $\bar{v} : (0,T) \to U$*, there exists a control* $v : (0,T) \to U$ *such that the corresponding solution* y_{v,y_0} *of* (I.1) *satisfies*

$$y_{v,y_0}(T) = y_{\bar{v},\bar{y}_0}(T).$$

If this property holds for any y_0*, we simply say that the system is* **controllable to trajectories**.

It is clear from the definitions that

$$\text{exact controllability} \Longrightarrow \text{approximate controllability},$$

$$\text{exact controllability} \Longrightarrow \text{controllability to trajectories}.$$

Moreover, for **linear problems** we have

$$\text{controllability to trajectories} \Longrightarrow \text{null-controllability},$$

and it can be often observed that

$$\text{controllability to trajectories} \Longrightarrow \text{approximate controllability}.$$

We will possibly also discuss about related topics like :

- Optimal control : find v such that the couple (y,v) satisfies some optimality criterion.
- Closed-loop stabilisation : Assume that 0 is an unstable fixed point of $y \mapsto F(y,0)$ (we assume here that F is autonomous), does it exist an operator K such that, if we define the control $v = Ky$, then 0 becomes an asymptotically stable fixed point of $y' = F(y,Ky)$.

I.2 Examples

Let us present a few examples.

I.2.1 The stupid example

$$\begin{cases} y' + \lambda y = v, \\ y(0) = y_0. \end{cases}$$

We want to drive y to a target y_T. Take any smooth function y that satisfy $y(0) = y_0$ and $y(T) = y_T$ and set $v = y' - \lambda y$ and we are done ... Of course there is much more to say on this example, like finding an *optimal* control in some sense.

Thanks to the Duhamel formula, we can write the solution explicitly as a function of y_0 and v

$$y(t) = e^{-\lambda t} y_0 + \int_0^t e^{-\lambda(t-s)} v(s)\, ds.$$

It follows that $y(T) = y_T$ for some v, if we have

$$\int_0^T e^{-\lambda(T-s)} v(s)\, ds = y_T - e^{-\lambda T} y_0.$$

Any function satisfying this integral condition will be a solution of our problem. It is clear that there exists plenty of such admissible functions.

- Let us try to consider a constant control $v(s) = M$ for any $s \in [0,T]$ and for some M. The equation to be solved is
$$M\frac{1-e^{-\lambda T}}{\lambda} = y_T - e^{-\lambda T}y_0.$$
It follows that
$$M = \lambda\frac{y_T - e^{-\lambda T}y_0}{1-e^{-\lambda T}}.$$
The L^2 norm on $[0,T]$ of this control is given by
$$\|v\|_{L^2(0,T)} = |M|\sqrt{T}.$$
  - If $y_T \neq 0$, we thus have
  $$\|v\|_{L^2(0,T)} \sim_{\lambda\to+\infty} \lambda\sqrt{T}|y_T|.$$
  This proves that the cost of such a control blows up as $\lambda \to \infty$.
  This is natural since the equation is more dissipative when λ is large and thus the system has more difficulties to achieve a non zero state.
  - Conversely, if $y_T = 0$, we have
  $$\|v\|_{L^2(0,T)} \sim_{\lambda\to+\infty} \lambda\sqrt{T}|y_0|e^{-\lambda T},$$
  and thus the cost of the control is asymptotically small when λ is large.
- Why do not take an exponential control ? For a given $\mu \in \mathbb{R}$, we set
$$v(t) = Me^{-\mu(T-t)},$$
the controllability condition reads
$$M\frac{1-e^{-(\lambda+\mu)T}}{\lambda+\mu} = y_T - e^{-\lambda T}y_0,$$
so that
$$M = (\lambda+\mu)\frac{y_T - e^{-\lambda T}y_0}{1-e^{-(\lambda+\mu)T}}.$$
Let us compute the L^2 norm of such a control
$$\begin{aligned}\int_0^T |v(t)|^2\,dt &= M^2\frac{1-e^{-2\mu T}}{2\mu}\\ &= \frac{(\lambda+\mu)^2}{2\mu}\frac{(y_T - e^{-\lambda T}y_0)^2}{(1-e^{-(\lambda+\mu)T})^2}(1-e^{-2\mu T}).\end{aligned}$$
We will see later that this quantity is minimal for $\mu = \lambda$ and we then obtain
$$\int_0^T |v(t)|^2\,dt = 2\lambda\frac{(y_T - e^{-\lambda T}y_0)^2}{(1-e^{-2\lambda T})^2}(1-e^{-2\lambda T}),$$
so that
$$\|v\|_{L^2(0,T)} \sim_{\lambda\to+\infty} \sqrt{2\lambda}|y_T|.$$
Observe that this cost behaves like $\sqrt{\lambda}$ for large λ compared to the constant control case which behaves like λ for large λ.

I.2.2 The rocket

We consider a rocket which is trying to land on the ground. The rocket is supposed to be a single material point (!!) and the motion is 1D (in the vertical direction). Let x be the altitude of the rocket and y its vertical velocity. The initial altitude is denoted by $x_0 > 0$ and the initial velocity is denoted by y_0 (we assume $y_0 \leq 0$ without loss of generality).

The control v is the force generated by the engines of the rocket. The equations of motion of this very simple example are

$$\begin{cases} x'(t) = y(t), \\ y'(t) = v(t) - g, \\ x(0) = x_0 > 0, \\ y(0) = y_0 \leq 0, \end{cases}$$

The goal is to land the rocket at time T : we want $x(T) = y(T) = 0$.

An explicit computation leads to

$$\begin{cases} y(t) = y_0 - gt + \int_0^t v(s)\, ds, \\ x(t) = h_0 + \int_0^t y(\tau)\, d\tau = h_0 + y_0 t - \frac{1}{2}gt^2 + \int_0^t v(s)(t-s)\, ds. \end{cases}$$

We conclude that, for a given $T > 0$, the control law v does the job if and only if it satisfies

$$\begin{cases} \int_0^T v(s)\, ds = gT + |y_0|, \\ \int_0^T v(s) s\, ds = \frac{1}{2}gT^2 + h_0. \end{cases} \tag{I.2}$$

This is our first (and not last !) contact with a *moment's problem*.

There is clearly an infinite number of solutions to the system (I.2). Let us try to build two examples:

- For some $T_0 \in (0, T)$ and some $M > 0$ to be fixed later, we look for a control of the following form

  $$v(t) = \begin{cases} M & \text{for } t < T_0, \\ 0 & \text{for } t > T_0. \end{cases}$$

  System (I.2) leads to

  $$MT_0 = gT + |y_0|,$$
  $$M\frac{T_0^2}{2} = \frac{1}{2}gT^2 + h_0.$$

  This can be solved as follows

  $$T_0 = \frac{gT^2 + 2h_0}{gT + |y_0|},$$

  and

  $$M = \frac{(gT + |y_0|)^2}{gT^2 + 2h_0}.$$

  Note that the condition $T_0 \leq T$ gives

  $$2h_0 \leq |y_0| T,$$

  which mean that such a solution is possible only for a control time T large enough.

- For some α, β to be fixed later, we set
$$v(t) = \alpha + \beta t, \ \ \forall t \in (0,T).$$
System (I.2) leads to
$$\alpha T + \beta \frac{T^2}{2} = gT + |y_0|,$$
$$\alpha \frac{T^2}{2} + \beta \frac{T^3}{3} = \frac{1}{2} g T^2 + h_0,$$
that we can solve explicitely
$$\beta \frac{T^3}{12} = h_0 - \frac{T|y_0|}{2},$$
$$\alpha \frac{T^2}{8} = \frac{h_0}{4} + \frac{1}{8} g T^2 - h_0 + \frac{T|y_0|}{2},$$
to obtain
$$v(t) = \left(g + \frac{|y_0|}{T} \right) + (t - T/2) \left(\frac{12 h_0}{T^3} - \frac{6|y_0|}{T^2} \right). \tag{I.3}$$
We observe that there is no condition on the time T for this function to be a mathematical solution of our problem. However, we have
$$\max_{[0,T]} |v(t)| \sim_{T \to 0} \frac{6h_0}{T^2},$$
which proves that, for small control times T, the magnitude of the necessary power of the engines may be infinite. This is of course not reasonable.

Similarly, for a *real* rocket, we expect v to be a non negative function. Looking at the expression above, we see that the non-negativity of v holds if and only if the following condition holds
$$|6h_0 - 3|y_0|T| \leq gT^2 + |y_0|T.$$
Here also, this condition is satisfied if T is large enough and certainly not satisfied for small values of T. It thus seems that this particular control is not physically admissible for small control times T.

The above solution defined in (I.3) is nevertheless interesting (from a modeling and mathematical point of view) since we can show that it is, for a given T, the unique solution among all possible solutions which has a minimal L^2 norm.
$$\int_0^T |v(t)|^2 \, dt = \underset{w \text{ admissible}}{\operatorname{argmin}} \int_0^T |w(t)|^2 \, dt.$$
Let us prove this in few lines : if $w : [0,T] \to \mathbb{R}$ is a control function that drives the solution at rest at time T, then it also solves the equations (I.2) and in particular we have
$$\int_0^T (v - w)(s) \, ds = 0,$$
$$\int_0^T s(v - w)(s) \, ds = 0.$$
Since v is a linear function, that is a combination of $s \mapsto 1$ and $s \mapsto s$, the above relations give
$$\int_0^T v(v - w) \, ds = 0.$$
This means that $v - w$ is orthogonal to v in L^2 and the Pythagorean theorem leads to
$$\|w\|_{L^2}^2 = \|(w - v) + v\|_{L^2}^2 = \|w - v\|_{L^2}^2 + \|v\|_{L^2}^2 \geq \|v\|_{L^2}^2,$$
with equality if and only if $v = w$.

The solution v is thus the optimal cost control with this particular definition of the cost.

Exercise I.2.2 (The *damped* rocket model)

In practice, the command of the pilot is not instantaneously transmitted to the rocket. To model this behavior, we introduce a delay time $\tau > 0$ and replace the previous model with the following one

$$\begin{cases} x'(t) = y(t), \\ y'(t) = w(t) - g, \\ w'(t) = \frac{1}{\tau}(v(t) - w(t)), \\ x(0) = x_0 > 0, \\ y(0) = y_0 \leq 0, \\ w(0) = 0. \end{cases}$$

By using the same approach as before, show that the previous system is controllable at any time $T > 0$. Compute explicitly such controls and try to find the one with minimal $L^2(0,T)$ norm.

I.2.3 Nonlinear examples

We consider a nonlinear autonomous (this is just for simplicity) ODE system of the form (I.1) and we assume that $F(0,0) = 0$ in such a way that $(y, v) = 0$ is a solution of the system. We would like to study the local controllability of the nonlinear system. To this end, we consider the linearized system

$$y' = Ay + Bv, \tag{I.4}$$

where $A = D_yF(0,0)$ and $B = D_vF(0,0)$ are the partial Jacobian matrices of F with respect to the state and the control variable respectively.

We will not discuss this point in detail but the general philosophy is the following:

- **Positive linear test:**

  If the linearized system (I.4) around $(0,0)$ is controllable, then the initial nonlinear system (I.1) is locally controllable at any time $T > 0$. More precisely, it means that for any $T > 0$, there exists $\varepsilon > 0$ such that for any $y_0, y_T \in \mathbb{R}^n$ satisfying $\|y_0\| \leq \varepsilon$ and $\|y_T\| \leq \varepsilon$, there exists a control $v \in L^\infty(0, T, \mathbb{R}^m)$ such that the solution of (I.1) starting at y_0 satisfies $y(T) = y_T$.

- **Negative linear test:**

  Unfortunately (or fortunately !) it happens that the linear test is not sufficient to determine the local controllability of a nonlinear system around an equilibrium. In other words : *nonlinearity helps !*

  There exists systems such that the linearized system is not controllable and that are nevertheless controllable.

- The nonlinear spring:

  $$y'' = -ky(1 + Cy^2) + v(t).$$

  The linearized system around the equilibrium $(y = 0, v = 0)$ is

  $$y'' = -ky + v,$$

  which is a controllable system (exercise ...). Therefore, we may prove that the nonlinear system is also controllable locally around the equilibrium $y = y' = 0$.

- The baby troller: This is an example taken from [Cor07].

  The unknowns of this system are the 2D coordinates (y_1, y_2) of the center of mass of the troller, and the direction y_3 of the troller (that is the angle with respect to any fixed direction). There are two controls v_1 and v_2 since the *pilot* can push the troller in the direction given by y_3 (with a velocity v_1) or turn the troller (with an angular veloctiy v_2). The set of equations is then

$$\begin{cases} y_1' = v_1 \cos(y_3), \\ y_2' = v_1 \sin(y_3), \\ y_3' = v_2. \end{cases}$$

Observe that any point $\bar{y} \in \mathbb{R}^3$, $\bar{v} = 0 \in \mathbb{R}^2$ is an equilibrium of the system. The linearized system around this equilibirum reads

$$\begin{cases} y_1' = v_1 \cos(\bar{y}_3), \\ y_2' = v_1 \sin(\bar{y}_3), \\ y_3' = v_2. \end{cases}$$

It is clear that this system is not controllable since the quantity

$$\sin(\bar{y}_3)y_1 - \cos(\bar{y}_3)y_2,$$

does not depend on time.

It follows that the (even local) controllability of the nonlinear system is much more difficult to prove ... and actually cannot rely on usual linearization arguments. However, it is true that the nonlinear system is locally controllable, see [Cor07].

I.2.4 PDE examples

- The transport equation : Boundary control

Let $y_0 : (0, L) \to \mathbb{R}$ and $c > 0$, we consider the following controlled problem

$$\begin{cases} \partial_t y + c\partial_x y = 0, \ \forall (t, x) \in (0, +\infty) \times (0, L), \\ y(0, x) = y_0(x), \ \ \forall x \in (0, L), \\ y(t, 0) = v(t). \end{cases} \tag{I.5}$$

When posed on the whole space $\mathbb{R}$, the exact solution of the transport problem reads

$$y(t, x) = y_0(c - xt), \ \ \forall t \geq 0, \forall x \in \mathbb{R}.$$

This can be proved by showing that the solution is constant along (backward) characteristics. In presence of an inflow boundary, the same property holds but it may happen that the characteristics touch the boundary at some positive time. In this case, the boundary condition has to be taken into account.

Therefore, for a given y_0 and v, the unique solution to Problem (I.5) is given by

$$y(t, x) = \begin{cases} y_0(x - ct), & \text{for } x \in (0, L), t < x/c, \\ v(t - x/c), & \text{for } x \in (0, L), t > x/c. \end{cases}$$

In the limit case $t = x/c$ there is an over-determination of the solution that cannot be solved in general. It follows that, even if y_0 and v are smooth, the solution is a **weak** solution which is possibly discontinuous. If, additionally, the initial condition and the boundary data satisfy the compatibility condition

$$y_0(x = 0) = v(t = 0),$$

then the exact solution is continuous.

Theorem I.2.3

- *If $T \geq L/c$ the transport problem is exactly controllable at time T, for initial data and target in $L^2(0,L)$ and with a control in $L^2(0,T)$.*

  If additionally we have $T > L/c$ and y_0,y_T are smooth, then we can find a smooth control v that produces a smooth solution y.

- *If $T < L/c$ the transport problem is not even approximately controllable at time T.*

- The heat equation : distributed internal control acting everywhere.

  Let $y_0 : (0,L) \to \mathbb{R}$, we consider the following controlled problem

$$\begin{cases} \partial_t y - \partial_x^2 y = v(t,x), \ \forall (t,x) \in (0,+\infty) \times (0,L), \\ y(0,x) = y_0(x), \ \ \forall x \in (0,L), \\ y(t,0) = y(t,L) = 0, \ \ \forall t > 0. \end{cases} \tag{I.6}$$

Take $L = \pi$ to simplify the computations. We look for y, v as a development in Fourier series

$$y(t,x) = \sqrt{2/\pi} \sum_{n \geq 1} y_n(t) \sin(nx),$$

$$v(t,x) = \sqrt{2/\pi} \sum_{n \geq 1} v_n(t) \sin(nx).$$

For each n the equation (I.6) gives

$$y_n'(t) + n^2 y_n(t) = v_n(t),$$

where $y_n(0) = y_{n,0} = \sqrt{2/\pi} \int_0^\pi y_0(x) \sin(nx)\, dx$ is the n-th Fourier coefficient of the initial data y_0. We try to achieve a state $y_T \in L^2(\Omega)$ whose Fourier coefficients are given $y_{n,T}$.

For each n we thus have to build a control v_n for a single ODE. We have seen that there are many solutions to do so. We need to take care of this choice since, at the end, we need to justify the convergence in some sense of the series that defines v.

- **Reachable set from** 0**.** We assume that $y_0 = 0$ and we would like to understand what kind of targets can be achieved and the related regularity of the control.

  * If we choose v_n to be constant in time, the computations of Section I.2.1 show that

$$v_n(t) = \frac{n^2 y_{n,T}}{1 - e^{-n^2 T}} \sim_{+\infty} n^2 y_{n,T}.$$

    Formally, we have thus found a time independent control v that reads

$$v(x) = \sqrt{2/\pi} \sum_{n \geq 1} \frac{n^2 y_{n,T}}{1 - e^{-n^2 T}} \sin(nx).$$

    The question is : what is the meaning of this series. Does it converges in $L^2(0,\pi)$ for instance ? We see that

$$\begin{aligned} v \in L^2(0,\pi) \ &\Leftrightarrow \ y_T \in H^2(0,\pi) \cap H_0^1(0,\pi), \\ v \in H^{-1}(0,\pi) \ &\Leftrightarrow \ y_T \in H_0^1(0,\pi), \\ v \in H^{-2}(0,\pi) \ &\Leftrightarrow \ y_T \in L^2(0,\pi). \end{aligned}$$

* Can we do better ? We have seen in Section I.2.1, that a better control (in the sense of a smaller L^2 norm) consists in chosing an exponential control $v_n(t) = M_n e^{-n^2(T-t)}$. In that case, we have the estimate

$$\|v_n\|_{L^2(0,T)} \sim_{+\infty} Cn|y_{n,T}|.$$

It can then be checked that the regularity of such a control is related to the regularity of y_T as follows.

$$v \in L^2(0,T,L^2(0,\pi)) \ \Leftrightarrow\ y_T \in H^1_0(0,\pi),$$

$$v \in L^2(0,T,H^{-1}(0,\pi)) \ \Leftrightarrow\ y_T \in L^2(0,\pi).$$

As a conclusion, if one wants to control to a target which is in $L^2(0,\pi)$, we can either take a time-independent control in $H^{-2}(0,\pi)$ or a time dependent control in $L^2(0,T,H^{-1}(0,\pi))$. In some sense we pay the higher regularity in space of v by a smaller regularity in time of v.

Another way to understand this analysis is that, if one wants to be able to control the equation with a control that only belongs to $L^2((0,T)\times\Omega)$, we need to impose $y_T \in H^1_0(0,\pi)$. A target y_T belonging to $L^2(0,\pi)\setminus H^1_0(0,\pi)$ (such as a indicatrix function for instance) is not achievable by controls in L^2.

– **Null-controllability :** We ask now a different question : we assume that $y_T = 0$ and that y_0 is any function. Is it possible to achieve 0 at time T starting from any y_0 ?

  * If we choose v_n to be constant in time, the computations of Section I.2.1 show that

$$v_n(t) = \frac{-n^2 e^{-n^2T} y_{n,0}}{1-e^{-n^2T}} \sim_{+\infty} -n^2 e^{-n^2T} y_{n,0}.$$

Formally, we have thus found a time independent control v that reads

$$v(x) = \sqrt{2/\pi}\sum_{n\geq 1} -\frac{n^2 e^{-n^2T} y_{n,0}}{1-e^{-n^2T}} \sin(nx).$$

and we observe that this series converges for any y_0 in a possibly very negative Sobolev space H^{-k}. This is a nice consequence of the regularizing effect of the heat equation (without source terms).
It follows immediately that the null-controllability of the heat equation is much more easy to achieve than the exact controllability to any given trajectory.

  * Just like before we could then try to find the optimal control in the L^2 sense. We will discuss this question in a more general setting later on.

In practice, we will be interested in control problems for the heat equation that are supported in a subset of the domain Ω or on the boundary. This makes the problem much more difficult as we will see in the sequel since it is no more possible to use a basic Fourier decomposition that lead to the resolution of a countable family of controlled scalar, linear, and independent ODEs.

Chapter II

Controllability of linear ordinary differential equations

In this chapter, we focus our attention on the following controlled system

$$\begin{cases} y'(t) + Ay(t) = Bv(t), \\ y(0) = y_0, \end{cases} \tag{II.1}$$

where $A \in M_n(\mathbb{R})$, $B \in M_{n,m}(\mathbb{R})$, $y(t) \in \mathbb{R}^n$ and $v(t) \in \mathbb{R}^m$. Note that A and B do not depend on time (even though some part of the following analysis can be adapted for non autonomous systems).

We shall often denote by $E = \mathbb{R}^n$ the state space and by $U = \mathbb{R}^m$ the control space.

II.1 Preliminaries

II.1.1 Exact representation formula

Given an initial data $y_0 \in \mathbb{R}^n$ and a control v, we recall that (II.1) can be explicitely solved by means of the fundamental solution of the homogeneous equation $t \mapsto e^{-tA}z$, $z \in \mathbb{R}^n$ and of the Duhamel formula. We obtain

$$y(t) = e^{-tA}y_0 + \int_0^t e^{-(t-s)A}Bv(s)\,ds, \quad \forall t \in [0,T].$$

In particular, the solution at time T (which is the object we are interested in) is given by

$$y(T) = e^{-TA}y_0 + \int_0^T e^{-(T-s)A}Bv(s)\,ds. \tag{II.2}$$

We recall that the exponential of any matrix M is defined by the series

$$e^M = \sum_{k\geq 0} \frac{M^k}{k!},$$

which is locally uniformly convergent.

The linear part (in v) of the solution will be denoted by

$$L_T v \stackrel{\text{def}}{=} \int_0^T e^{-(T-s)A}Bv(s)\,ds,$$

it corresponds to the solution of our system with the initial data $y_0 = 0$.

In the non-autonomous case, we need to use the resolvant matrix as recalled in Section A.1 of Appendix A.

II.1.2 Duality

As we will see later on, it will be very useful to adopt a dual point of view in our analysis. For the moment, we simply pick any $q_T \in \mathbb{R}^n$ and we take the Euclidean inner product of (II.2) by q_T. We get

$$\langle y(T), q_T \rangle_E = \langle e^{-TA} y_0, q_T \rangle_E + \int_0^T \langle e^{-(T-s)A} Bv(s), q_T \rangle_E \, ds,$$

that we can rewrite, using the adjoint operators (=transpose matrix in this context), as follows

$$\langle y(T), q_T \rangle_E = \langle y_0, e^{-TA^*} q_T \rangle_E + \int_0^T \langle v(s), B^* e^{-(T-s)A^*} q_T \rangle_U \, ds. \tag{II.3}$$

We can still reformulate at little bit this formula by introducing the adjoint equation of (II.1) which is the backward in time homogeneous system (i.e. without any control term)

$$-q'(t) + A^* q(t) = 0, \tag{II.4}$$

with the *final* data $q(T) = q_T$ and which can be explicitely computed

$$q(t) = e^{-(T-t)A^*} q_T.$$

We will see in Section II.5 the reason why the adjoint equation enters the game.

With this notation, (II.3) becomes

$$\langle y(T), q(T) \rangle_E = \langle y_0, q(0) \rangle_E + \int_0^T \langle v(s), B^* q(s) \rangle_U \, ds, \tag{II.5}$$

and this equation holds true for any solution q of the adjoint system (II.4)

II.1.3 Reachable states. Control spaces

The solution of our system (II.2) is well-defined as soon as $v \in L^1(0,T,\mathbb{R}^m) = L^1(0,T,U)$, see section A.2 of Appendix A and the corresponding solution map $L_T : v \mapsto y$ is continuous from $L^1(0,T,U)$ into $\mathcal{C}^0([0,T],E)$.

For any subspace $V \subset L^1(0,T,U)$ we define the set of reachable states at time T as follows

$$R_{T,V}(y_0) \stackrel{\text{def}}{=} \left\{ e^{-TA} y_0 + \int_0^T e^{-(T-s)A} Bv(s) \, ds, \ \text{ for } v \in V \right\} = e^{-TA} y_0 + L_T(V).$$

We immediately see that $R_{T,V}(y_0)$ is a (finite dimensional) affine subspace of $E = \mathbb{R}^n$. Moreover, since L_T is continuous for the $L^1(0,T,U)$ topology, we obtain that

$$R_{T,\overline{V}}(y_0) = \overline{R_{T,V}(y_0)},$$

and since this last space is finite dimensional, we finally have

$$R_{T,\overline{V}}(y_0) = R_{T,V}(y_0).$$

As a consequence, for any **dense** subspace V of $L^1(0,T,U)$, we have

$$R_{T,V}(y_0) = R_{T,L^1(0,T,U)}(y_0).$$

Therefore, in the sequel we can choose, without consequence, any dense subspace of $L^1(0,T,U)$ to study the controllability properties of our system and the corresponding reachable set will simply be denoted by $R_T(y_0)$.

As a consequence of the previous analysis, we have that if $y_T \in R_T(y_0)$ we can actually achieve this target with a control belonging to the space $\mathcal{C}_c^\infty(]0,T[)$.

II.2 Kalman criterion. Unique continuation

The first criterion we have in order to decide whether or not (II.1) is controllable is the following famous result.

Theorem II.2.1 (Kalman rank criterion)

Let $T > 0$. The following propositions are equivalent.

1. *Problem (S) is exactly controllable at time T (for any y_0, y_T ...)*
2. *Problem (S) is approximately controllable at time T (for any y_0, y_T ...)*
3. *The matrices A and B satisfy*

$$\operatorname{rank}(K) = n, \quad \text{with } K \stackrel{\text{def}}{=} \big(B|AB|\dots|A^{n-1}B\big) \in M_{n,mn}(\mathbb{R}). \tag{II.6}$$

If any of the above properties hold we say that the pair (A, B) is controllable.

The matrix K in this result is called the Kalman matrix.

Remark II.2.2

- *This result shows, in particular, that in this framework the notions of approximate and exact controllability are equivalent.*
- *It also shows that those two notions are independent of the time horizon T.*
- *It is very useful to observe that the rank condition (II.6) is equivalent to the following property*

$$\mathit{Ker}\ K^* = \{0\}.$$

Proof :

In this proof, we assume that y_0 is any fixed initial data.

1.⇔2. Since we work in a finite dimensional setting, it follows from (II.2) that

$$\begin{aligned}\text{exact controllability from } y_0 &\iff R_T(y_0) = E\\ &\iff R_T(y_0) \text{ is dense in } E\\ &\iff \text{approximate controllability from } y_0.\end{aligned}$$

1.⇒3. Assume that $\operatorname{rank}(K) < n$, or equivalently that Ker $K^* \neq \{0\}$; it follows that there exists $q_T \in \mathbb{R}^n \setminus \{0\}$ such that $K^* q_T = 0$. But we have

$$\begin{aligned}K^* q_T = 0 &\iff B^*(A^*)^p q_T = 0,\ \forall p < n\\ &\iff B^*(A^*)^p q_T = 0,\ \forall p \geq 0, \quad \text{by the Cayley-Hamilton Theorem}\\ &\iff B^* e^{-sA^*} q_T = 0,\ \forall s \in [0, T], \quad \text{by the properties of the exponential.}\end{aligned}$$

By (II.3), we deduce that such a q_T is necessarily orthogonal to the vector space $R_T(y_0) - e^{-TA} y_0$, and therefore this subspace cannot be equal to $\mathbb{R}^n$.

3.⇒1. Assume that our system is not exactly controllable at time T. It implies that, there exists a $q_T \neq 0$ which is orthogonal to $R_T(y_0) - e^{-TA} y_0$. By (II.3), we deduce that **for any control** v we have

$$\int_0^T \langle v(s), B^* e^{-(T-s)A^*} q_T \rangle_U \, ds = 0.$$

We apply this equality to the particular control $v(s) = B^* e^{-(T-s)A^*} q_T$ to deduce that we necessarily have

$$B^* e^{-sA^*} q_T = 0, \quad \forall s \in [0,T].$$

The equivalences above show that $q_T \in \text{Ker } K^*$ and thus this kernel cannot reduce to $\{0\}$.

■

Remark II.2.3

At the very beginning of the proof we have shown that

$$q_T \in \textit{Ker } K^* \iff q_T \in Q_T,$$

where Q_T is the set of the non-observable adjoint states *defined by*

$$Q_T \stackrel{\text{def}}{=} \{q_T \in \mathbb{R}^n, \ \ B^* e^{-sA^*} q_T = 0, \ \ \forall s \in [0,T]\}.$$

Thus, another formulation of the Kalman criterion is

$$(A,B) \textit{ is controllable} \iff \left(B^* e^{-sA^*} q_T = 0, \ \ \forall s \in [0,T] \Rightarrow q_T = 0 \right).$$

This last property is called the unique continuation property of the adjoint system through the observation operator B^*.
The point we want to emphasize here is that, in the infinite dimension case, it can be difficult to define a Kalman matrix *(or operator) if A is an unbounded linear operator (because we need to compute successive powers of A) but however, it seems to be affordable to define the set Q_T as soon as we have a suitable semi-group theory that gives a sense to e^{-sA^*} for $s \geq 0$ since it is not possible in general to simply set $e^{-sA^*} = \sum_{k\geq 0} \frac{1}{k!}(-sA^*)^k$ when A^* is a differential operator.*
More precisely, if we imagine for a moment that A is an unbounded linear operator in an Hilbert space (for instance the Laplace-Dirichlet operator in some Sobolev space), then it is very difficult to define a kind of Kalman operator *since it would require to consider successive powers of A, each of them being defined on different domains (that are getting smaller and smaller at each application of A).*

Example II.2.4

Without loss of generality we can assume that B is full rank $\text{rank}(B) = m$.

1. *If the pair (A,B) is controllable, then the eigenspaces of A^* (and thus also those of A) has at most dimension m. For instance if $m = 1$, a necessary condition for the controllability of the pair (A,B) is that each eigenvalue of A^* is geometrically simple.*

   Another necessary condition is that the minimal polynomial of A^* is of degree exactly n.

2. *Second order systems. With the same notations as before, the second order controlled system*

$$y'' + Ay = Bv,$$

is controllable if and only if the pair (A, B) *satisfies the Kalman criterion.*

3. *Conditions on the control: If the pair* (A, B) *is controllable then we can find controls satisfying additional properties.*

   - *For any* $v_0 \in \mathbb{R}^m$ *and* $v_T \in \mathbb{R}^m$ *we can find a control* v *from* y_0 *to* y_T *for our system such that*

   $$y(0) = y_0, \; y(T) = y_T, \; v(0) = v_0, \text{ and } v(T) = v_T.$$

   - *We can find a control* $v \in \mathcal{C}_c^\infty(0, T)$ *such that* $y(0) = y_0$ *and* $y(T) = y_T$.

In view of the techniques we will present later on on the controllability of parabolic PDEs, we shall now present another proof of the previous theorem.

Proof (of Theorem II.2.1 - direct proof):

We shall actually prove that, if the Kalman condition is satisfied then our system is indeed controllable. Moreover, we shall give a **constructive** proof of the control.

For simplicity (and since we are mainly interested in presenting the method and not in the general result that we have already proved before), we shall assume that $m = 1$. We also assume that $y_T = 0$ (which is always possible for a linear system).

By assumption the square (since $m = 1$) matrix K is invertible and thus we shall use the change of variable $y = Kz$ in order to transform our control system. A simple computation shows that

$$B = K \underbrace{\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{=\bar{B}}, \text{ and } AK = K \underbrace{\begin{pmatrix} 0 & \cdots & \cdots & 0 & a_{1,n} \\ 1 & 0 & \cdots & \vdots & a_{2,n} \\ 0 & & \ddots & \vdots & a_{3,n} \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & a_{n,n} \end{pmatrix}}_{=\bar{A}}.$$

It follows that the equation for z

$$Kz' + AKz = Bv,$$

becomes

$$K(z' + \bar{A}z) = K\bar{B}v,$$

and since K is invertible

$$z' + \bar{A}z = \bar{B}v \tag{II.7}$$

With the Kalman matrix, we thus have been able to put our system into a canonical form where $\bar{A}$ has a companion structure (it looks pretty much like a Jordan block) and $\bar{B}$ is the first vector of the canonical basis of $\mathbb{R}^n$.

This structure if often called **cascade systems** in control theory. The important feature of $\bar{A}$ is that its under diagonal terms do not vanish. It reveals the particular way by which the control v acts on the system. Indeed, v directly appears in the first equation and then tries to drive z_1 to the target at time T (observe however that the dynamics is also coupled with the rest of the system by the term $a_{1,n}z_n$)

$$z_1'(t) + a_{1,n}z_n(t) = v(t).$$

The control v does not appear in the second equation

$$z_2'(t) + z_1(t) + a_{2,n}z_n(t) = 0,$$

but this equation contains a term z_1 that plays the role of an indirect control of z_2, and so on ...

Let us now give the construction of the control v:

- We start by defining $(\bar{z}_i)_{1\leq i\leq n}$ to be the free solution of the system (the one with $v=0$).
- We choose a truncature function $\eta : [0,T] \to \mathbb{R}$ such that $\eta = 1$ on $[0,T/3]$ and $\eta = 0$ on $[2T/3,T]$.
- We start by choosing

  $$z_n(t) \stackrel{\text{def}}{=} \eta(t)\bar{z}_n(t),$$

  then, by using the last equation of the system (II.7), we need to define

  $$z_{n-1}(t) \stackrel{\text{def}}{=} z_n'(t) - a_{n-1,n}z_n(t).$$

  Similarly, by using the equation number $n-1$ of (II.7), we set

  $$z_{n-2}(t) \stackrel{\text{def}}{=} z_{n-1}'(t) - a_{n-2,n}z_n(t).$$

  by induction, we define $z_{n-3}, ..., z_2$ in the same way.

  Finally, the first equation of the system (II.7) gives us the control we need

  $$v(t) = z_1'(t) + a_{1,n}z_n(t).$$

  By such a construction, the functions $(z_i)_i$ satisfy the controlled system with the control v we just defined.
- Let us prove, by reverse induction that, for any k we have

  $$\begin{cases} z_k = \bar{z}_k, \text{ in } [0,T/3], \\ z_k = 0, \text{ in } [2T/3,T]. \end{cases} \tag{II.8}$$

  This will in particular prove that $z(T) = 0$ and that $z(0) = \tilde{z}(0) = \bar{z}(0) = z_0$.

  - For $k = n$, the properties (II.8) simply comes from the choice of the truncature function.
  - For $k = n-1$, we observe that, by construction and induction, for any $t \in [0,T/3]$,

    $$z_{n-1}(t) = z_n'(t) - a_{n-1,n}z_n(t) = \bar{z}_n{}'(t) - a_{n-1,n}\bar{z}_n(t) = \bar{z}_{n-1}(t),$$

    the last equality coming from the fact that $\bar{z}$ solves the free equation.
  - And so on up to $k = 1$, ...

■

Exercise II.2.5

Propose a similar proof to deal with the case $m = 2$ and $\operatorname{rank}(B) = m = 2$.

Exercise II.2.6

Assume that A, B are such that the rank r of the Kalman matrix K satisfies $r < n$. Then there exists a $P \in GL_n(\mathbb{R})$ such that

$$A = P\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22}\end{pmatrix}P^{-1}, \text{ and } B = P\begin{pmatrix} B_1 \\ 0\end{pmatrix},$$

and moreover the pair (A_{11}, B_1) is controllable.

What are the consequences of this result for the controllability of the initial system ?

Exercise II.2.7 (Partial controllability)

We assume given $p \leq n$ and a matrix $P \in M_{p,n}(\mathbb{R})$. We say that (II.1) is partially controllable relatively to P if and only if for any $y_0 \in \mathbb{R}^n$ and any $\bar{y}_T \in \mathbb{R}^p$ there exists a control $v \in L^2(0,T;U)$ such that the associated solution to (II.1) satisfies

$$Py(T) = \bar{y}_T.$$

Show that (II.1) is partially conntrollable relatively to P if and only if

$$\text{rank}(K_P) = p,$$

where

$$K_P \stackrel{\text{def}}{=} \left(PB|PAB|\dots|PA^{n-1}B\right) \in M_{p,mn}(\mathbb{R}).$$

II.3 Fattorini-Hautus test

We are going to establish another criterion for the controllability of autonomous linear ODE systems. This one will only be concerned with the eigenspaces of the matrix A^*, and we know that there are plenty of unbounded operators for which we can define a suitable spectral theory. It is then easy to imagine that we will be able, at least, to formulate a similar result in the infinite dimension case.

Theorem II.3.8 (Fattorini-Hautus test)

The pair (A,B) is controllable if and only if we have

$$\textit{Ker}\,(B^*) \cap \textit{Ker}\,(A^* - \lambda I) = \{0\}, \quad \forall \lambda \in \mathbb{C}. \tag{II.9}$$

In other words : (A,B) is controllable if and only if

$$B^*\phi \neq 0, \quad \textit{for any eigenvector } \phi \textit{ of } A^*.$$

Let us start with the following straightforward lemma (in which the space Q_T is considered as a subspace of $\mathbb{C}^n$).

Lemma II.3.9

For any polynomial $P \in \mathbb{C}[X]$ we have

$$P(A^*)Q_T \subset Q_T.$$

Proof :

Let $q_T \in Q_T$. By definition, we have

$$B^* e^{sA^*} q_T = 0, \ \forall s \in \mathbb{R},$$

so that by differentiating k times with respect to s, we get

$$B^* e^{sA^*} (A^*)^k q_T = 0, \ \forall s \in \mathbb{R}.$$

It means that $(A^*)^k q_T \in Q_T$. The proof is complete. ■

Proof (of Theorem II.3.8):

The Kalman criterion says that (A,B) is controllable if and only if we have Ker $K^* = \{0\}$. Moreover, we saw at the end of Section II.2 that this condition is equivalent to saying that there is no non-observable adjoint states excepted 0, that is

$$Q_T = \{0\}.$$

- Assume first that (II.9) is not true. There exists a $\lambda \in \mathbb{C}$ and a $\phi \neq 0$ such that
$$A^*\phi = \lambda\phi, \quad \text{and} \quad B^*\phi = 0.$$
Note that, in particular, λ is an eigenvalue of A^*. A straighforward computation shows that
$$B^* e^{-sA^*}\phi = B^*\left(e^{-s\lambda}\phi\right) = e^{-s\lambda}B^*\phi = 0.$$
This proves that $\phi \in Q_T$ so that $Q_T \neq \{0\}$. Therefore the system does not fulfill the Kalman criterion. We have proved the non controllability of the system.

- Assume that (II.9) holds and let $\phi \in Q_T$. We shall prove that $\phi = 0$. To begin with we take $\lambda \in \mathbb{C}$ an eigenvalue of A^* and we introduce E_λ the generalized eigenspace associated with λ, that is
$$E_\lambda = \text{Ker}_{\,\mathbb{C}^n}(A^* - \lambda I)^n.$$
Linear algebra says that we can write the direct sum
$$\mathbb{C}^n = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p},$$
with distinct values of $(\lambda_i)_i$.
We recall that the projector on E_λ associated with such a direct sum can be expressed as a polynomial in A^* : there exists polynomials $P_\lambda \in \mathbb{C}[X]$ such that
$$\phi = \sum_{i=1}^{p} P_{\lambda_i}(A^*)\phi, \text{ with } P_{\lambda_i}(A^*)\phi \in E_{\lambda_i}, \quad \forall i \in \{1, ..., p\}. \tag{II.10}$$
By Lemma II.3.9, we have $\phi_\lambda \stackrel{\text{def}}{=} P_\lambda(A^*)\phi \in Q_T$. We want to show that $\phi_\lambda = 0$. If it is not the case, there exists $k \geq 1$ such that
$$(A^* - \lambda I)^k\phi_\lambda = 0, \text{ and } (A^* - \lambda I)^{k-1}\phi_\lambda \neq 0.$$
This proves that $(A^* - \lambda I)^{k-1}\phi_\lambda$ is an eigenvector of A^* and, by Lemma II.3.9 it belongs to Q_T. Since by definition we have $Q_T \subset \text{Ker}\, B^*$, we have proved that
$$(A^* - \lambda I)^{k-1}\phi_\lambda \in \text{Ker}\,(B^*) \cap \text{Ker}\,(A^* - \lambda I),$$
which is a contradiction with (II.9).
Therefore, $\phi_\lambda = 0$ for any eigenvalue λ and, by (II.10), we eventually get $\phi = 0$.

■

Remark II.3.10

The above proof of the Fattorini-Hautus test is not necessarily the simplest one in the finite dimension case but it has the advantage to be generalizable to the infinite dimensional setting, see Theorem III.3.7.

Exercise II.3.11 (Simultaneous control)

Let us assume that $m = 1$ and we are given two pairs (A_1, B_1) (dimension n_1) and (A_2, B_2) (of dimension n_2). We assume that both pairs are controllable and we ask the question of whether they are simultaneously controllable (that is we can drive the two systems from one point to another by using the same control for both systems).
Show that the two systems are simultaneously controllable if and only if $\text{Sp}(A_1) \cap \text{Sp}(A_2) = \emptyset$.

II.4 The moments method

We shall now describe, still in the simplest case of an autonomous linear controlled system of ODEs, one of the methods that can be used to construct a control and that will appear to be powerful in the analysis of the control of evolution PDEs in the next chapters. We will assume that the Fattorini-Hautus condition (II.9) holds and we fix the target to be $y_T = 0$ to simplify a little the computations.

This method relies on more or less explicit formulas for the exponential matrices e^{-sA^*} using eigenelements of A^*.

We present the method in the case $m = 1$ (B is thus a single column vector) even though it can be adapted to more general settings. Let us denote by $\Lambda = \mathrm{Sp}(A^*)$ the complex spectrum of A^*. Since $m = 1$, we known by the Hautus test (or by Example II.2.4) that all the eigenspaces are one dimensional.

For each $\lambda \in \Lambda$, we can then choose one eigenvector $\Phi^0_\lambda \in \mathbb{C}^n$. Let $\alpha_\lambda \in \mathbb{N}^*$ be the algebraic multiplicity of the eigenvalue λ and Φ^j_λ, $1 \leq j \leq \alpha_\lambda - 1$ be an associated Jordan chain, that is a sequence of generalized eigenvectors that satisfy

$$(A^* - \lambda)\Phi^l_\lambda = \Phi^{l-1}_\lambda, \quad l \in \{1, ..., \alpha_\lambda\}.$$

Those vectors are defined up to the addition of any multiple of the eigenvector Φ^0_λ. Since $B^*\Phi^0_\lambda \neq 0$ by (II.9) we can impose, in addition, the condition

$$B^*\Phi^l_\lambda = 0, \quad \forall 1 \leq l \leq \alpha_\lambda - 1. \tag{II.11}$$

With those notations, we can compute for any $s \in \mathbb{R}$, the action of the exponential on the Jordan chain as follows

$$e^{-sA^*}\Phi^l_\lambda = \sum_{j=0}^{l} e_s[(\lambda)^{j+1}]\Phi^{l-j}_\lambda,$$

or with the Leibniz formula

$$e^{-sA^*}\Phi^l_\lambda = (e_s\Phi)[(\lambda)^{l+1}].$$

Using (II.3), we see that a function v is a control (with target $y_T = 0$) if and only if we have (here $U = \mathbb{R}$)

$$\int_0^T v(s)B^*e^{-(T-s)A^*}q_T\,ds = -\langle y_0, e^{-TA^*}q_T\rangle_E = -\langle e^{-TA}y_0, q_T\rangle_E, \ \forall q_T \in \mathbb{R}^n.$$

Note that we can also test this equality with complex adjoint states $q_T \in \mathbb{C}^n$.

By linearity, it is enough to test this equality on a basis of $\mathbb{C}^n$. In particular, we can use the basis $(\Phi^l_\lambda)_{\substack{\lambda\in\Lambda\\ 0\leq l\leq\alpha_\lambda-1}}$ and we obtain that v is a control if and only if we have

$$\int_0^T v(s)(e_{T-s}B^*\Phi)[(\lambda)^{l+1}]\,ds = -\langle e^{-TA}y_0, \Phi^l_\lambda\rangle, \ \forall\lambda \in \Lambda, \ \forall 0 \leq l \leq \alpha_\lambda - 1.$$

Using (II.11), we get that this set of equations simplifies as follows

$$(B^*\Phi^0_\lambda)\int_0^T v(s)e_{T-s}[(\lambda)^{l+1}]\,ds = -\langle e^{-TA}y_0, \Phi^l_\lambda\rangle, \ \forall\lambda \in \Lambda, \ \forall 0 \leq l \leq \alpha_\lambda - 1.$$

Defining

$$\omega^l_\lambda \stackrel{\text{def}}{=} -\frac{\langle e^{-TA}y_0, \Phi^l_\lambda\rangle}{B^*\Phi^0_\lambda},$$

we see that v is control for our problem if and only if the function $u(t) = v(T - t)$ (introduced to simplify the formulas) satisfies

$$\int_0^T u(s)e_s[(\lambda)^{l+1}]\,ds = \omega^l_\lambda, \ \forall\lambda \in \Lambda, \ \forall 0 \leq l \leq \alpha_\lambda - 1. \tag{II.12}$$

This kind of problem is called a moments problem : we need to find a function u whose integrals against a given family of functions is prescribed, or in other words, to find a function u whose $L^2(0,T)$ inner products against a family of functions in L^2 is prescribed. If this family was orthogonal in L^2 the solution will be straightforward but unfortunately it is clearly not the case here.

However it can easily be seen that

$$E = \{e[(\lambda)^{l+1}], \ \ \lambda \in \Lambda, 0 \leq l \leq \alpha_\lambda\},$$

is a linearly independent family in $L^2(0,T)$.

By Proposition A.4.10, we know that there exists a biorthogonal family in $L^2(0,T)$ to E that we denote by

$$F = \{f_\lambda^l, \ \ \lambda \in \Lambda, 0 \leq l \leq \alpha_\lambda\}.$$

This means that we have

$$\int_0^T e_s[(\lambda)^{l+1}] f_\mu^k(s)\, ds = \delta_{\lambda,\mu}\delta_{l,k}.$$

It is then clear that the function

$$u(t) = \sum_{\lambda\in\Lambda} \sum_{l=0}^{\alpha_\lambda - 1} \omega_\lambda^l f_\lambda^l(t),$$

is a solution to (II.12). Therefore $v(t) = u(T-t)$ is a control that drives the solution to our system to $y_T = 0$ at time T.

Remark II.4.12

The argument above is actually an alternative proof that the Fattorini-Hautus criterion is a sufficient controllability condition for our system (indeed we managed to build a control by simply using the fact that $B^*\phi \neq 0$ for any ϕ which is an eigenvector of A^*).

Remark II.4.13 (Optimal $L^2(0,T)$ control)

The construction above strongly depends on the choice of the biorthogonal family F since there are infinitely many such families. However, choosing the unique such family that satisfy

$$F \subset \mathrm{Span}(E), \tag{II.13}$$

as mentioned in Proposition (A.4.10), then we can prove that the associated control, that we call v_0, is the one of minimal $L^2(0,T)$-norm.

Indeed, assume that $v \in L^2(0,T)$ is any other admissible control for our problem and let $u_0(t) = v_0(T-t)$ and $u(t) = v(T-t)$. Since u and u_0 both satisfy the same system of linear equations (II.12), we first deduce that

$$\int_0^T (u(s) - u_0(s)) e_s[(\lambda)^{l+1}]\, ds = 0, \ \ \forall \lambda \in \Lambda, \forall 0 \leq l \leq \alpha_\lambda - 1.$$

Using now the fact that u_0 is a combination of the elements in F and by the assumption (II.13), we conclude that

$$\int_0^T (u(s) - u_0(s)) u_0(s)\, ds = 0.$$

This naturally implies that

$$\|u\|_{L^2}^2 = \|u_0\|_{L^2}^2 + \|u - u_0\|_{L^2}^2,$$

and of course that

$$\|v\|_{L^2}^2 = \|v_0\|_{L^2}^2 + \|v - v_0\|_{L^2}^2.$$

This actually proves that v_0 is the **unique** *admissible control with minimal L^2 norm.*

II.5 Linear-Quadratic optimal control problems

In this section, we will discuss a class of problems which is slightly different from the controllability issues that we discussed previously. However, some of those results will be useful later on and are interesting by themselves (in particular in applications).

II.5.1 Framework

Since it does not change anything to the forthcoming analysis we do not assume in this section that the linear ODE we are studying is autonomous. More precisely, we suppose given continuous maps $t \mapsto A(t) \in M_n(\mathbb{R})$ and $t \mapsto B(t) \in M_{n,m}(\mathbb{R})$ and an initial data y_0 and we consider the following controlled ODE

$$\begin{cases} y'(t) + A(t)y(t) = B(t)v(t), \\ y(0) = y_0. \end{cases} \tag{II.14}$$

Following Sections A.1 and A.2 of the Appendix A, this problem is well-posed for $v \in L^1(0,T,\mathbb{R}^m)$, in which case the solution satisfies $y \in \mathcal{C}^0([0,T],\mathbb{R}^n)$ and the solution map $v \in L^1 \mapsto y \in \mathcal{C}^0$ is continuous.

Let now $t \mapsto M_y(t) \in S_n^+(\mathbb{R})$, $t \mapsto M_v(t) \in S_m^+(\mathbb{R})$ be two continuous maps with values in the set of symmetric semi-definite positive matrices $S_n^+(\mathbb{R})$ and $M_T \in S_n^+$ be a symmetric semi-definite positive matrix. We assume that M_v is uniformly definite positive :

$$\exists \alpha > 0, \;\; \langle M_v(t)\xi, \xi\rangle_U \geq \alpha \|\xi\|^2, \;\; \forall \xi \in \mathbb{R}^m, \forall t \in [0,T]. \tag{II.15}$$

For any given control function $v \in L^2(0,T,\mathbb{R}^m)$, we can now define the cost functional

$$F(v) \overset{\text{def}}{=} \frac{1}{2}\int_0^T \langle M_y(t)y(t), y(t)\rangle_E \, dt + \frac{1}{2}\int_0^T \langle M_v(t)v(t), v(t)\rangle_U \, dt + \frac{1}{2}\langle M_T y(T), y(T)\rangle_E,$$

where, in this formula, y is the unique solution to (II.14) associated with the given control v. Since y depends linearly on the couple (y_0, v), we see that the functional F is quadratic and convex. Moreover, it is strictly convex thanks to the assumption (II.15).

II.5.2 Main result. Adjoint state

Theorem II.5.14

Under the assumptions above, there exists a unique minimiser $\bar{v} \in L^2(0,T,\mathbb{R}^m)$, of the functional F on the set $L^2(0,T,\mathbb{R}^m)$.
Moreover, $\bar{v}$ is the unique function in $L^2(0,T,\mathbb{R}^m)$ such that there exists $q \in \mathcal{C}^1([0,T],\mathbb{R}^n)$ satisfying the set of equations

$$\begin{cases} y'(t) + A(t)y(t) = B(t)\bar{v}(t), \\ y(0) = y_0, \\ -q'(t) + A^*(t)q(t) + M_y(t)y(t) = 0, \\ q(T) = -M_T y(T), \\ \bar{v}(t) = M_v(t)^{-1}B^*(t)q(t). \end{cases} \tag{II.16}$$

Moreover, the optimal energy is given by

$$\inf_{L^2(0,T,\mathbb{R}^m)} F = F(\bar{v}) = -\frac{1}{2}\langle q(0), y_0\rangle_E.$$

Such a function q is called **adjoint state** associated with our optimization problem.

Observe that there is no assumption on A and B for such an optimization problem to have a solution.

Remark II.5.15

One of the consequence of the previous theorem is that the optimal control $\bar{v}$ is at least continuous in time and, if all the matrix-valued functions in the problem are $\mathcal{C}^k$ then the solution $\bar{v}$ is itself $\mathcal{C}^k$.

Before proving the theorem we can make the following computation.

Proposition II.5.16

Assume that (y, q, v) is a solution to system (II.16), then we define $\phi(t) = \langle y(t), q(t)\rangle$ and we have

$$\phi'(t) = \langle M_y(t)y(t), y(t)\rangle_E + \langle M_v(t)v(t), v(t)\rangle_U.$$

In particular, the solution of (II.16) (if it exists) is unique.

Proof :

We just compute the derivative of ϕ to get

$$\begin{aligned}\phi'(t) &= \langle q'(t), y(t)\rangle_E + \langle q(t), y'(t)\rangle_E \\ &= \langle A^*(t)q(t) + M_y(t)y(t), y(t)\rangle_E - \langle q(t), A(t)y(t) - B(t)v(t)\rangle_E \\ &= \langle M_y(t)y(t), y(t)\rangle_E + \langle B^*(t)q(t), v(t)\rangle_U \\ &= \langle M_y(t)y(t), y(t)\rangle_E + \langle M_v(t)v(t), v(t)\rangle_U.\end{aligned}$$

In particular, ϕ is non-decreasing. If $y_0 = 0$, then $\phi(0) = 0$ and thus $\phi(T) \geq 0$ and by construction we have

$$\phi(T) = -\langle M_T y(T), y(T)\rangle_E \geq 0.$$

By assumption on M_T, we deduce that $M_T y(T) = 0$ (notice that M_T is not assumed to be definite positive) and using the equation relating $q(T)$ to $y(T)$, we deduce that $q(T) = 0$ and that $\phi(T) = 0$.

It follows, by integration over the time interval $(0, T)$, that

$$\int_0^T \langle M_y y, y\rangle_E + \langle M_v v, v\rangle_U \, dt = \int_0^T \phi'(t)\, dt = \phi(T) - \phi(0) = 0.$$

By assumption on M_v, we deduce that $v = 0$. The equation for y leads to $y = 0$ and finally the equation on q gives $q = 0$.

■

Let us now prove the main result.

Proof (of Theorem II.5.14):

Uniqueness of the minimizer comes from the strict convexity of F. Moreover, F is non-negative and therefore has a finite infimum. In order to prove existence of the minimizer, we consider a minimizing sequence $(v_n)_n \subset L^2(0, T, \mathbb{R}^m)$:

$$F(v_n) \xrightarrow[n\to\infty]{} \inf F.$$

We want to prove that $(v_n)_n$ is convergent. We may proceed by weak convergence arguments (that are more general) but in the present case we can simply use the fact that F is quadratic and that the dependence of y with respect to v is affine. In particular, we have

$$\begin{aligned}8F\left(\frac{v_1+v_2}{2}\right) &= \int_0^T \langle M_y(y_1+y_2)(t), (y_1+y_2)(t)\rangle_E \, dt \\ &\qquad + \int_0^T \langle M_v(v_1+v_2)(t), (v_1+v_2)(t)\rangle_U \, dt + \langle M_T(y_1+y_2)(T), (y_1+y_2)(T)\rangle_E,\end{aligned}$$

and by the parallelogram formula we have

$$8F\left(\frac{v_1+v_2}{2}\right) = 4F(v_1)+4F(v_2) \\ -8\bigg(\int_0^T \langle M_y(y_1-y_2)(t),(y_1-y_2)(t)\rangle_E\,dt + \int_0^T \langle M_v(v_1-v_2)(t),(v_1-v_2)(t)\rangle_U\,dt \\ + \langle M_T(y_1-y_2)(T),(y_1-y_2)(T)\rangle_E\bigg).$$

By (II.15), we deduce that

$$2F\left(\frac{v_1+v_2}{2}\right) \leq F(v_1)+F(v_2) - \alpha\|v_1-v_2\|^2_{L^2}.$$

Applying this inequality to two elements of the minimizing sequence v_n and v_{n+p}, we get

$$2\inf F \leq 2F\left(\frac{v_n+v_{n+p}}{2}\right) \leq F(v_n)+F(v_{n+p}) - \alpha\|v_n-v_{n+p}\|^2_{L^2},$$

from which we deduce that

$$\lim_{n\to\infty}\left(\sup_{p\geq 0}\|v_n-v_{n+p}\|_{L^2}\right)=0.$$

This proves that $(v_n)_n$ is a Cauchy sequence in $L^2(0,T,\mathbb{R}^m)$. Since this space is complete, we deduce that $(v_n)_n$ converges towards some limit $\bar{v}$ in this space. Let y_n be the solution of (II.14) associated with v_n and $\bar{y}$ the solution associated with $\bar{v}$. The continuity of the solution operator $v\mapsto y$ (see Appendix A.2) gives that y_n converges towards $\bar{y}$ in $\mathcal{C}^0([0,T],\mathbb{R}^n)$.

It is thus a simple exercice to pass to the limit in the definition of $F(v_n)$ and to prove that it actually converges towards $F(\bar{v})$. The proof of the first part is complete.

Let us compute the differential of F at the equilibrium $\bar{v}$ in the direction $h\in L^2(0,T,\mathbb{R}^m)$. We have

$$dF(\bar{v}).h = \int_0^T \langle M_y(t)y(t),\delta(t)\rangle_E\,dt + \int_0^T \langle M_v(t)\bar{v}(t),h(t)\rangle_U\,dt + \langle M_T y(T),\delta(T)\rangle_E,$$

where δ is the solution of the problem

$$\begin{cases}\delta'(t)+A(t)\delta(t)=B(t)h(t),\\ \delta(0)=0.\end{cases}$$

Let q be the unique solution to the adjoint problem

$$\begin{cases}-q'(t)+A^*(t)q(t)+M_y y(t)=0,\\ q(T)=-M_T y(T),\end{cases}$$

We deduce that

$$\begin{aligned}\int_0^T \langle M_y(t)y(t),\delta(t)\rangle_E\,dt &= -\int_0^T \langle -q'(t)+A^*(t)q(t),\delta(t)\rangle_E\,dt\\ &= -\int_0^T \langle q(t),\delta'(t)+A(t)\delta(t)\rangle_E\,dt + \langle q(T),\delta(T)\rangle_E - \langle q(0),\delta(0)\rangle_E\\ &= -\int_0^T \langle q(t),B(t)h(t)\rangle_E\,dt - \langle M_T y(T),\delta(T)\rangle_E\\ &= -\int_0^T \langle B^*(t)q(t),h(t)\rangle_U\,dt - \langle M_T y(T),\delta(T)\rangle_E.\end{aligned}$$

It follows that

$$dF(\bar{v}).h = \int_0^T \langle M_v(t)\bar{v}(t) - B^*(t)q(t), h(t)\rangle_U \, dt.$$

The Euler-Lagrange equation for the minimization problem for F gives $dF(\bar{v}) = 0$ so that we finally find that

$$M_v(t)\bar{v}(t) = B^*(t)q(t), \quad \forall t \in [0, T].$$

This is the expected condition between the optimal control $\bar{v}$ and the adjoint state q. The first part of the proof is complete.

We introduce the function $\phi(t) = \langle q(t), y(t)\rangle_E$, we have $\phi(T) = -\langle M_T y(T), y(T)\rangle_E$, and by Proposition II.5.16 we conclude that

$$\inf_{L^2(0,T,\mathbb{R}^m)} F = F(\bar{v}) = -\frac{1}{2}\phi(T) + \frac{1}{2}\int_0^T \phi'(t)\, dt = -\frac{1}{2}\phi(0) = -\frac{1}{2}\langle y_0, q(0)\rangle_E.$$

■

II.5.3 Justification of the gradient computation

It remains to explain how we obtain in general the equations for the adjoint state. The formal computation (that may be fully justified in many cases) makes use of the notion of Lagragian.

Let us set $J(v, y)$ to be the same definition as F but with independent unknowns v and y. Minimizing F amounts at minimizing J with the additional constraints that $y(0) = y_0$ and $y'(t) + A(t)y(t) = B(t)v(t)$.

To take into account those constraints, we introduce two dual variables $q : [0, T] \to \mathbb{R}^n$ and $q_0 \in \mathbb{R}^n$. The Lagrangian functional is thus defined by

$$L(v, y, q, q_0) = J(v, y) + \int_0^T \langle q(t), y'(t) + A(t)y(t) - B(t)v(t)\rangle_E \, dt + \langle q_0, y(0) - y_0\rangle_E.$$

A simple integration by parts leads to

$$\begin{aligned} L(v, y, q, q_0) = J(v, y) + \int_0^T \langle -q'(t) + A^*(t)q(t), y(t)\rangle_E \, dt - \int_0^T \langle B^*(t)q(t), v(t)\rangle_U \, dt \\ + \langle q(T), y(T)\rangle_E - \langle q(0), y(0)\rangle_E + \langle q_0, y(0) - y_0\rangle_E. \end{aligned}$$

And finally, the initial functional F satisfies

$$F(v) = L(v, y[v], q[v], q_0[v]),$$

for any choice of $q[v]$ and $q_0[v]$ since $y[v]$ satisfies both constraints. It follows that the differential of F satisfies

$$\begin{aligned} dF(v).h &= \partial_v L.h + \partial_y L.(dy[v].h) + \partial_q L.(dq[v].h) + \partial_{q_0} L.(dq_0[v].h), \\ &= \partial_v L.h + \partial_y L.(dy[v].h), \end{aligned}$$

since $\partial_q L$ and $\partial_{q_0} L$ are precisely the two constraints satisfied by $y[v]$. The idea is now to choose $q[v]$ and $q_0[v]$ so as to eliminate the term in $\partial_y L$.

For any $\delta : [0, T] \to \mathbb{R}^n$, we have

$$\partial_y L.\delta = \int_0^T \langle M_y y(t) - q'(t) + A^*(t)q(t), \delta(t)\rangle_E \, dt + \langle M_T y(T), \delta(T)\rangle_E + \langle q(T), \delta(T)\rangle_E - \langle q(0) - q_0, \delta(0)\rangle_E.$$

This quantity vanishes for any δ if and only if we have the relations

$$\begin{cases} q_0 = q(0), \\ q(T) = -M_T y(T), \\ -q'(t) + A^*(t)q(t) = -M_y y(t). \end{cases}$$

This defines the dual variables q and q_0 in a unique way for a given v (and thus a given y). Those are the Lagrange multipliers of the constrained optimization problem.

Once we have defined those values, the computation of the differential of F leads to

$$dF(v).h = \partial_v L(v, y[v], q[v], q_0[v]).h = \int_0^T \langle M_v(t)v(t), h(t)\rangle_U \, dt - \int_0^T \langle B^* q(t), h(t)\rangle_U \, dt,$$

which is of course the same expression as above.

II.5.4 Ricatti equation

The set of optimality equations (II.16) is in general a complicated system of coupled ODEs that is **not** a Cauchy problem. It is remarkable that its solution can be obtained through the resolution of a Cauchy problem for a nonlinear matrix-valued ordinary differential equation. It has in particular some important applications to the closed-loop stabilization of the initial problem.

Theorem II.5.17 (Adjoint state and Ricatti equation)

Under the previous assumptions, there exists a matrix-valued map $t \in [0,T] \mapsto E(t)$ that only depends on A, B, M_y, M_v, M_T, and T, such that the adjoint state q in the previous theorem satisfies

$$q(t) = -E(t)y(t), \quad \forall t \in [0,T].$$

In other words, the optimal control $\bar{v}$ can be realized, whatever the initial data y_0 is, as a closed-loop control

$$\bar{v}(t) = -M_v(t)^{-1}B^*(t)E(t)y(t).$$

Moreover, the function E is the unique solution in $[0,T]$ to the following (backward in time) Cauchy problem associated with a Ricatti differential equation

$$\begin{cases} E'(t) = -M_y(t) + A^*(t)E(t) + E(t)A(t) + E(t)B(t)M_v(t)^{-1}B^*(t)E(t), \\ E(T) = M_T. \end{cases} \tag{II.17}$$

Finally, $E(t)$ is symmetric semi-definite positive for any t and even definite positive if M_T is definite positive, and we have

$$\inf_{L^2(0,T,\mathbb{R}^m)} F(v) = \frac{1}{2}\langle E(0)y_0, y_0\rangle_E.$$

Observe that the Ricatti equation is a matrix-valued nonlinear differential equation which is not necessarily easy to solve. Actually, it is not even clear that the solution exists on the whole time interval $[0,T]$; this will be a consequence of the proof.

Proof :

The Cauchy-Lipschitz theorem ensures that (II.17) has a unique solution **locally** around $t = T$.

We start by assuming that this solution is defined on the whole time interval $[0,T]$. It is clear that E^* satisfies the same Cauchy problem as E and thus, by uniqueness, $E = E^*$.

Then we define y to be the unique solution of the Cauchy problem

$$\begin{cases} y'(t) + A(t)y(t) = -B(t)M_v(t)^{-1}B^*(t)E(t)y(t), \\ y(0) = y_0. \end{cases}$$

Then we set

$$q(t) \stackrel{\text{def}}{=} -E(t)y(t),$$

and

$$v(t) \stackrel{\text{def}}{=} -M_v(t)^{-1}B^*(t)E(t)y(t).$$

In order to show that such a v is the optimal control, we need to check all the equations in (II.16). The first two equations and the last two are satisfied by construction, it remains to check the third equation. This is a simple computation

$$\begin{aligned}-q'(t) + A^*(t)q(t) =& E'(t)y(t) + E(t)y'(t) - A^*(t)E(t)y(t)\\ =& - M_y(t)y(t) + E(t)y'(t)\\ &+ E(t)\big[A(t)y(t) + B(t)M_v(t)^{-1}B^*(t)E(t)y(t)\big]\\ =& - M_y(t)y(t).\end{aligned}$$

This proves the fact that, provided that E exists, the triple (y, v, q) is the unique solution of our optimality condition equations.

The fact that the optimal energy is given by $\frac{1}{2}\langle E(0)y_0, y_0\rangle_E$ is a simple consequence of Proposition II.5.16 and of the fact that $\phi(T) = -\langle M_T y(T), y(T)\rangle_E$, so that

$$\inf_{L^2(0,T,\mathbb{R}^m)} F = F(v) = -\frac{1}{2}\phi(T) + \frac{1}{2}\int_0^T \phi'(t)\,dt = -\frac{1}{2}\phi(0).$$

As a consequence, $\phi(0)$ is non-positive for any y_0, which proves that E is semi-definite positive.

Moreover, we deduce that $\frac{1}{2}\langle E(0)y_0, y_0\rangle_E$ is not larger than the value of the cost functional F when computed on the control $v = 0$. A simple computation of the solution of the ODE without control gives that the following bound holds

$$\langle E(0)y_0, y_0\rangle_E \leq \left(\|M_T\| + \int_0^T \|M_y\|\right) e^{2\int_0^T \|A\|} \|y_0\|^2, \;\; \forall y_0 \in \mathbb{R}^n.$$

This gives a bound on $\|E(0)\|$.

We can now prove the global existence of E on $[0, T]$. Indeed, if we assume that E is defined on $[t^*, T]$ for some $0 \leq t^* < T$, the previous computation (with the initial time t^* instead of 0) shows that

$$\begin{aligned}\|E(t^*)\| &\leq \left(\|M_T\| + \int_{t^*}^T \|M_y\|\right) e^{2\int_{t^*}^T \|A\|}\\ &\leq \left(\|M_T\| + \int_0^T \|M_y\|\right) e^{2\int_0^T \|A\|}.\end{aligned}$$

It follows that E is bounded independently of t^* and therefore can not blow up in finite time. Therefore the existence and uniqueness of E over the whole time interval $[0, T]$ is proved. ∎

II.6 The HUM control

Let us come back now to the controllability question (and we assume again that A and B are time-independent).

We would like to address the question of the characterisation of a **best** control among all the possible controls, if such controls exist. Of course, this notion will depend on some criterion that we would like to choose as a measure of the "quality" or the "cost" of the control.

The HUM formulation Assume that y_0, y_T are such that $y_T \in R_T(y_0)$. We can easily prove that the set of admissible controls

$$\mathrm{adm}(y_0, y_T) \stackrel{\text{def}}{=} \{v \in L^2(0, T; U), \;\; y_v(T) = y_T\},$$

is a non-empty convex set which is closed in $L^2(0, T; U)$. Therefore, there exists a unique control of minimal L^2-norm, that we denote by v_0. It satisfies the optimization problem

$$F(v_0) = \inf_{v \in \mathrm{adm}(y_0, y_T)} F(v), \tag{II.18}$$

where we have introduced
$$F(v) \stackrel{\text{def}}{=} \frac{1}{2}\int_0^T \|v(t)\|_U^2\,dt, \quad \forall v \in L^2(0,T;U).$$
We recall the definition of the solution operator (without initial data)
$$L_T : v \in L^2(0,T;U) \mapsto \int_0^T e^{-(T-s)A}Bv(s)\,ds \in E,$$
in such a way that the (affine) constraint set reads
$$\mathrm{adm}(y_0,y_T) = \{v \in L^2(0,T;U), \;\; L_T(v) = y_T - e^{-TA}y_0\}.$$
Since v_0 is a solution of the constrained optimisation problem, we can use the Lagrange multiplier theorem to affirm that there exists a vector $q_T \in E$ such that
$$dF(v_0).w = \langle q_T, dL_T(v_0).w\rangle_E, \quad \forall w \in L^2(0,T;U).$$
Since L_T is linear, we have $dL_T(v_0).w = L_T(w)$ and the differential of the quadratic functional F is given by
$$dF(v_0).w = \int_0^T \langle v_0(s), w(s)\rangle_U\,ds, \quad \forall w \in L^2(0,T;U).$$
It follows that v_0 satisfies, for some $q_T \in E$ and for any $w \in L^2(0,T;U)$ the equation
$$\int_0^T \langle v_0(s), w(s)\rangle_U\,ds = \int_0^T \langle q_T, e^{-(T-s)A}Bw(s)\rangle_E\,ds,$$
which gives
$$v_0(s) = B^*e^{-(T-s)A^*}q_T. \tag{II.19}$$
This proves that the HUM control v_0 has a special form as shown above. In particular if one wants to compute v_0 we only have to determine the Lagrange multiplier q_T. To this end, we plug the form (II.19) into the equation that v_0 has to fulfill
$$y_T = e^{-TA}y_0 + \left(\int_0^T e^{-(T-s)A}BB^*e^{-(T-s)A^*}\,ds\right)q_T,$$
which is a linear system in q_T that we write
$$\Lambda q_T = y_T - e^{-TA}y_0, \tag{II.20}$$
where we have introduced the Gramian matrix
$$\Lambda \stackrel{\text{def}}{=} \int_0^T e^{-(T-s)A}BB^*e^{-(T-s)A^*}\,ds.$$
We observe that Λ is a symmetric positive semi-definite matrix and that is definite if and only if the Kalman criterion is satisfied.

Finally, the HUM control v_0 can be computed by solving first the linear system (II.20), whose unique solution is denoted by $q_{T,opt}$ and then by using (II.19).

It is also of interest to observe that the optimal $q_{T,opt} \in E$ is the unique solution of the optimization problem
$$J(q_{T,opt}) = \inf_{q_T \in E} J(q_T), \tag{II.21}$$
where we have introduced the functional
$$J(q_T) \stackrel{\text{def}}{=} \frac{1}{2}\int_0^T \left\|B^*e^{-(T-s)A^*}q_T\right\|_U^2\,ds + \langle y_0, e^{-TA^*}q_T\rangle_E - \langle y_T, q_T\rangle_E.$$

One can prove, by the Fenchel-Rockafellar duality theorem, that J is the adjoint problem associated with the initial optimisation problem (II.18).

Observe that (II.21) is an unconstrained finite dimensional optimization problem whereas (II.18) is a constrained infinite dimensional optimization problem. This is one of the reason why it is often more suitable to solve (II.21) instead of (II.18).

Actually, the explicit computation of the matrix Λ and its inversion can be quite heavy (in large dimension) and, in practice, we may prefer to solve the linear system (II.20) by using an iterative method (like the conjugate gradient for instance) that only necessitates to compute matrix-vector products. For any given q_T, the product Λq_T, can be obtained with the following general procedure:

- Solve the adjoint (backward) equation $-q'(t) + A^* q(t) = 0$ with the final data $q(T) = q_T$, in the present case, it gives
$$q(t) = e^{-(T-t)A^*} q_T.$$
- Define the control v by $v(t) = B^* q(t)$.
- Solve the primal (forward) problem $y'(t) + Ay(t) = Bv(t)$, with initial data $y(0) = 0$. In the present case it gives
$$y(t) = \int_0^t e^{-(t-s)A} Bv(s)\, ds.$$
- The value of Λq_T is then given by
$$\Lambda q_T = y(T),$$
since we have
$$\begin{aligned} y(T) &= \int_0^T e^{-(T-s)A} Bv(s)\, ds \\ &= \int_0^T e^{-(T-s)A} BB^* q(s)\, ds \\ &= \int_0^T e^{-(T-s)A} BB^* e^{-(T-s)A^*} q_T\, ds \end{aligned}$$

Remark II.6.18

At the end of this analysis, we have actually proved that the optimal control in $L^2(0,T;U)$ (the HUM control) has the particular form (II.19), which proves in particular that v_0 is smooth and thus the ODE system is satisfied in the usual sense for this control.

Remark II.6.19

Our analysis shows, as a side effect, that v_0 is the unique possible control for our system that we can write under the form (II.19).

Exercise II.6.20

Assume that the pair (A, B) is controllable, and let $T > 0$ given. Show that there exists $\varepsilon > 0$ such that for any $y_0, y_T \in E$, there exists a control for our problem that belongs to $\mathcal{C}^\infty([0, T])$ and such that $\operatorname{Supp} v \subset [\varepsilon, T-\varepsilon]$.

II.7 How much it costs ? Observability inequalities

We can now ask the question of computing the cost of the control. We suppose given A, B, the initial data y_0 and the target y_T.

The *best* control v_0 (the so-called HUM control) is given as a solution of the optimization problem described above and we have the following result.

Proposition II.7.21

Assume that the Kalman rank condition is satisfied for the pair (A, B), then the optimal cost of control from y_0 to y_T for our system is given by

$$\int_0^T \|v_0(t)\|_U^2 \, dt = \sup_{q_T \in E} \frac{|\langle y_T, q_T\rangle_E - \langle y_0, e^{-TA^*} q_T\rangle_E|^2}{\langle \Lambda q_T, q_T\rangle_E},$$

where Λ is the Gramiam operator that we built in the previous section.

Proof :

Let C be the value of the supremum in the right-hand side (this supremum is finite since the quantity is homogeneous in q_T and, by the Kalman condition, we know that $\langle \Lambda q_T, q_T\rangle_E \neq 0$ as soon as $q_T \neq 0$).

Let $q_{T,opt}$ be the unique solution to (II.20), in such a way that $v_0(s) = B^* e^{-sA^*} q_{T,opt}$. We observe first that

$$\langle \Lambda q_{T,opt}, q_{T,opt}\rangle_E = \int_0^T \|B^* e^{-sA^*} q_{T,opt}\|_U^2 \, ds = \int_0^T \|v_0(s)\|_U^2 \, ds,$$

and second, by (II.20), we have

$$\langle \Lambda q_{T,opt}, q_{T,opt}\rangle_E = \langle y_T, q_{T,opt}\rangle_E - \langle y_0, e^{-TA^*} q_{T,opt}\rangle_E.$$

It follows that

$$C \geq \frac{|\langle y_T, q_{T,opt}\rangle_E - \langle y_0, e^{-TA^*} q_{T,opt}\rangle_E|^2}{\langle \Lambda q_{T,opt}, q_{T,opt}\rangle_E} = \langle \Lambda q_{T,opt}, q_{T,opt}\rangle_E = \int_0^T \|v_0(s)\|_U^2 \, ds.$$

Conversely, if v is any control that drives the solution from y_0 to y_T we see from (II.5) and the Cauchy-Schwarz inequality that

$$|\langle y_T, q_T\rangle_E - \langle y_0, e^{-TA^*} q_T\rangle_E| \leq \left(\int_0^T \|v(s)\|_U^2 \, ds\right)^{\frac{1}{2}} \langle \Lambda q_T, q_T\rangle_E^{\frac{1}{2}}.$$

Taking the square of this inequality and then the supremum over all the possible q_T gives that

$$C \leq \int_0^T \|v(s)\|_U^2 \, ds,$$

and since this is true for all possible controls, this is in particular true for the optimal control v_0 and we get

$$C \leq \int_0^T \|v_0(s)\|_U^2 \, ds.$$

■

The previous result gives an estimate of the control cost, in the case where the pair (A, B) is controllable. We can actually be a little bit more precise: we shall prove that the boundedness of the supremum in the previous condition is a necessary and sufficient condition for the system to be controllable from y_0 to y_T.

Theorem II.7.22

Let A, B be any pair of matrices (we do not assume that the Kalman condition holds). Then, System (II.1) is controllable from y_0 to y_T if and only if, for some $C \geq 0$, the following inequality holds

$$|\langle y_T, q_T\rangle_E - \langle y_0, e^{-TA^*}q_T\rangle_E|^2 \leq C^2 \int_0^T \|B^* e^{-(T-s)A^*} q_T\|_U^2 \, ds, \quad \forall q_T \in E. \tag{II.22}$$

Moreover, the best constant C in this inequality is exactly equal the $L^2(0,T;U)$ norm of the HUM control v_0 from y_0 to y_T.

The above inequality is called an **observability inequality** on the adjoint equation. It amounts to control some information on any solution of the problem (in the left-hand side of the inequality) by the **observation** (which is the right-hand side term of the inequality). The operator B^* is called the observation operator.

We also note that, by definition of the Gramiam Λ, the right-hand side of the required observability inequality can also be written as follows

$$C^2\langle \Lambda q_T, q_T\rangle_E.$$

Proof :

Since e^{-TA} is invertible [1] we can always write

$$y_T = e^{-TA}\Big(e^{TA}y_T\Big).$$

So that the control problem is the same if we replace y_T by 0 and y_0 by $y_0 - e^{TA}y_T$ and we see that the left-hand side in the inequality is changed accordingly.

From now on, we will thus assume without loss of generality that $y_T = 0$ and that y_0 is any element in E.

- We first assume that there exists a control $v \in L^2(0,T)$ that drives y_0 to 0 at time T. Hence the set $\mathrm{adm}(y_0, 0)$ is not empty. We define v_0 to be the unique minimal L^2-norm element in $\mathrm{adm}(y_0,0)$. The same argument as in the previous proposition shows that for any q_T we have

$$|\langle y_0, e^{-TA^*}q_T\rangle_E|^2 \leq \left(\int_0^T \|v_0(s)\|_U^2 \, ds\right)\left(\int_0^T \|B^* e^{-(T-s)A^*}q_T\|^2 \, ds\right).$$

  This proves (II.22) with $C = \|v_0\|_{L^2(0,T;U)}$.

- Assume now that (II.22) holds for some $C > 0$. We would like to prove that $\mathrm{adm}(y_0,0)$ is not empty. The idea is to replace the constraint $v \in \mathrm{adm}(y_0,0)$ (that is $y(T)=0$) in the optimization problem (II.18) by a penalty term.

  For any $\varepsilon > 0$, we set

$$F_\varepsilon(v) = \frac{1}{2}\int_0^T \|v(s)\|_U^2 \, ds + \frac{1}{2\varepsilon}\|y(T)\|_E^2,$$

  where in this expression, y is the unique solution of (II.1) starting from the initial data y_0.

  The last term penalizes the fact that we would like $y(T) = 0$. Formally, we expect that, as $\varepsilon \to 0$, this term will impose $y(T)$ to get close from y_T.

  We consider now the following optimization problem: to find $v_\varepsilon \in L^2(0,T;U)$ such that

$$F_\varepsilon(v_\varepsilon) = \inf_{v \in L^2(0,T;U)} F_\varepsilon(v). \tag{II.23}$$

[1] this will not be true anymore for infinite dimensional problems when the underlying equation is not time reversible, which is precisely the case of parabolic equations

This functional exactly falls into the framework of the LQ optimal control problems that we studied in Section II.5, in the particular case where

$$M_v(t) = \mathrm{Id},\ M_y(t) = 0,\ \ \forall t \in [0,T], \quad \text{and} \quad M_T = \frac{1}{\varepsilon}\,\mathrm{Id}.$$

The characterisation theorem II.5.14 implies that this functional F_ε has a unique minimiser v_ε which is characterised by the following set of equations

$$\begin{cases} y_\varepsilon'(t) + Ay_\varepsilon(t) = Bv_\varepsilon(t), \\ y_\varepsilon(0) = y_0, \\ -q_\varepsilon'(t) + A^*q_\varepsilon(t) = 0, \\ q_\varepsilon(T) = -\frac{1}{\varepsilon}y_\varepsilon(T), \\ v_\varepsilon(t) = B^*q_\varepsilon(t). \end{cases}$$

Our goal is to study the behavior of $(v_\varepsilon, y_\varepsilon, q_\varepsilon)$ when $\varepsilon \to 0$. To this end, we try to obtain uniform bounds on those quantities.

To this end, we multiply (in the sense of the euclidean inner product of E) the state equation (the first one) by $q_\varepsilon(t)$ and we integrate the result over $(0,T)$. Using integration by parts and the other equations in the optimality system above, we obtain

$$\begin{aligned} \int_0^T \|v_\varepsilon\|^2\,dt &= \int_0^T \langle v_\varepsilon, B^*q_\varepsilon\rangle_U\,dt \\ &= \int_0^T \langle Bv_\varepsilon, q_\varepsilon\rangle_E\,dt \\ &= \int_0^T \langle y_\varepsilon' + Ay_\varepsilon, q_\varepsilon\rangle_E\,dt \\ &= \langle y_\varepsilon(T), q_\varepsilon(T)\rangle_E - \langle y_0, q_\varepsilon(0)\rangle_E + \int_0^T \langle y_\varepsilon, -q_\varepsilon' + A^*q_\varepsilon\rangle_E\,,dt \\ &= -\frac{1}{\varepsilon}\|y_\varepsilon(T)\|^2 - \langle y_0, q_\varepsilon(0)\rangle_E. \end{aligned}$$

It follows that

$$\|v_\varepsilon\|^2_{L^2(0,T,U)} + \frac{1}{\varepsilon}\|y_\varepsilon(T)\|^2 = -\langle y_0, q_\varepsilon(0)\rangle_E.$$

And, if we set $q_{T,\varepsilon} = q_\varepsilon(T)$, we can write this formula by using only the adjoint variable

$$\int_0^T \|B^*e^{-(T-t)A^*}q_{T,\varepsilon}\|^2\,dt + \varepsilon\|q_{T,\varepsilon}\|^2 = -\langle y_0, e^{-TA^*}q_{T,\varepsilon}\rangle_E. \tag{II.24}$$

We use now the observability inequality (II.22) (where we recall that y_T was taken to be 0 here). This inequality exactly gives us a bound on the right-hand side term

$$-\langle y_0, e^{-TA^*}q_{T,\varepsilon}\rangle_E \leq C\left(\int_0^T \|B^*e^{-(T-t)A^*}q_{T,\varepsilon}\|^2\,dt\right)^{\frac{1}{2}}.$$

We deduce that

$$\|v_\varepsilon\|^2_{L^2} = \int_0^T \|B^*e^{-(T-t)A^*}q_{T,\varepsilon}\|^2\,dt \leq C^2,$$

$$\varepsilon\|q_{T,\varepsilon}\|^2 \leq C^2.$$

From those estimates we obtain that $(v_\varepsilon)_\varepsilon$ is bounded in $L^2(0,T;U)$ and therefore we can extract a subsequence $(v_{\varepsilon_k})_k$ that weakly converges towards some $v \in L^2(0,T;U)$. Let y be the solution of (II.1) associated with this control v and the initial data y_0. Since the solution operator L_T is continuous from $L^2(0,T;U)$ into E, we deduce that $(L_T(v_{\varepsilon_k}))_k$ weakly converges towards $L_T(v)$ as $k \to \infty$ (note however that E is finite dimensional so that this convergence is also strong). It follows that $y_\varepsilon(T) \to y(T)$ as $\varepsilon \to 0$

Moreover, by definition of $q_{T,\varepsilon}$, we have the relation

$$y_\varepsilon(T) = -\varepsilon q_{T,\varepsilon},$$

and from the bound below we deduce that

$$\|y_\varepsilon(T)\|_E \leq \varepsilon \|q_{T,\varepsilon}\|_E \leq C\sqrt{\varepsilon} \xrightarrow[\varepsilon\to 0]{} 0.$$

Gathering all the above properties, we have shown that the weak limit v is such that the solution y satisfies

$$y(T) = 0,$$

which exactly means that the control v drives the solution of our system from 0 to y_T, or in other words $v \in \mathrm{adm}(y_0, 0)$.

This set being non empty we can consider the miminal L^2 norm control v_0 and, from the first part of the proof we know that necessarily we have

$$C \leq \|v_0\|_{L^2(0,T;U)} \leq \|v\|_{L^2(0,T;U)}.$$

Coming back to the bound on v_ε obtained above we see that

$$\limsup_{k\to\infty} \|v_{\varepsilon_k}\|_{L^2(0,T;U)} \leq C,$$

and since v is the weak limit of $(v_{\varepsilon_k})_k$ we conclude by usual properties of weak convergence in an Hilbert space that the convergence is actually strong and that we have the equality $\|v\|_{L^2(0,T;U)} = C$.

This implies in particular that $\|v\|_{L^2(0,T;U)} \leq \|v_0\|_{L^2(0,T;U)}$ and since v_0 is the unique minimal L^2-norm control, we deduce that $v = v_0$. In particular $C = \|v_0\|_{L^2(0,T,U)}$.

The standard uniqueness argument finally shows that the whole family $(v_\varepsilon)_\varepsilon$ strongly converges towards the HUM control v_0.

Observe that the family of the optimal adjoint states for the penalized problems $(q_{T,\varepsilon})_\varepsilon$ may not converge in this setting (except in the case where the Kalman rank condition is satisfied).

■

Remark II.7.23

If we have no other information on the matrices A, B or on the initial data y_0, the only hope to bound the right-hand side of (II.24) is to write

$$-\langle y_0, e^{-TA^*} q_{T,\varepsilon}\rangle_E \leq \|y_0\| \|e^{-TA^*}\| \|q_{T,\varepsilon}\|,$$

and to use the Young inequality to absorb the norm of $q_{T,\varepsilon}$ by the second term in the left-hand side to obtain

$$\int_0^T \|B^* e^{-(T-t)A^*} q_{T,\varepsilon}\|^2 \, dt + \varepsilon \|q_{T,\varepsilon}\|^2 \leq \frac{1}{\varepsilon} \|y_0\|^2 \|e^{-TA^*}\|^2.$$

This estimate is clearly useless since it does not provide a uniform bound on the control v_ε (and this is of course what is expected !).

As a conclusion of this analysis, we have converted a controllability question (which is a problem of proving the existence of some mathematical object satisfying some requirements) into an *observability* question which is : can we prove an *a priori* inequality like (II.22) that concerns solutions to an uncontrolled equation (the adjoint problem).

Remark II.7.24

If, for any q_T, we introduce $t \mapsto q(t)$ the solution of the adjoint equation

$$-q'(t) + A^* q(t) = 0, \;\; q(T) = q_T,$$

the observability inequality can be written as follows

$$|\langle y_T, q_T \rangle_E - \langle y_0, q(0) \rangle_E|^2 \leq C^2 \int_0^T \|B^* q(s)\|_U^2 \, ds, \quad \forall q_T \in \mathbb{R}^n,$$

which is slightly more general since it does not require any semi-group theory (and in particular can be generalised to non-autonomous equations).

Let us consider two particular cases of interest:

- **Exact controllability :** we assume that $y_0 = 0$ and $y_T \in \mathbb{R}^n$ is any target. The control cost is denoted by $C(0, y_T)$ and is the best constant in the inequality
$$|\langle y_T, q_T \rangle_E|^2 \leq C(0, y_T)^2 \int_0^T \|B^* e^{-(T-s)A^*} q_T\|_U^2 \, ds, \quad \forall q_T \in E. \tag{II.25}$$
- **Null-controllability :** we assume that $y_T = 0$ and $y_0 \in E$ is any initial data. The control cost is denoted by $C(y_0, 0)$ and is the best constant in the inequality
$$|\langle y_0, e^{-TA^*} q_T \rangle_E|^2 \leq C(y_0, 0)^2 \int_0^T \|B^* e^{-(T-s)A^*} q_T\|_U^2 \, ds, \quad \forall q_T \in E. \tag{II.26}$$

In the finite dimensional setting those two cases are very similar but it will make some difference when we will study parabolic PDEs.

Let ϕ be a normalized eigenvector of A^* associated with the eigenvalue λ and we assume that $\mathcal{R}e\,(\lambda) > 0$ (we mimick here the expected behavior of a parabolic PDE). Let us evaluate the costs $C(\phi, 0)$ and $C(0, \phi)$.

- We first take $q_T = \phi$ in (II.25) (with $y_T = \phi$) to get
$$C(0, \phi)^2 \geq \frac{2\mathcal{R}e\,(\lambda)}{\|B^*\phi\|_U^2 (1 - e^{-2T\mathcal{R}e\,(\lambda)})},$$
and we can obtain a rough bound from below
$$C(0, \phi)^2 \geq \frac{2\mathcal{R}e\,(\lambda)}{\|B^*\phi\|_U^2}.$$
This illustrates the fact that, if B^* is a given bounded operator, the cost of the exact controllability for a given eigenmode increases at least with the dissipation rate $\mathcal{R}e\,(\lambda)$. In the limit $\mathcal{R}e\,(\lambda) \to \infty$, this cost is therefore blowing up.

  This is not a good news if one imagines that we eventually want to control parabolic PDEs which are typically based on operators with sequences of eigenvalues that tends to infinity.

  The *physical* interpretation of this phenomenon is clear : the natural behavior of such a system for large values of $\mathcal{R}e\,(\lambda)$ is to strongly dissipate the solution with time which is exactly the converse of the fact that we require the solution to be driven to a constant normalized state ϕ at time T.

  This is the first appearance of the fact that, for dissipative systems (i.e. parabolic PDEs), the exact controllability property is not a good notion.

- Let us do the same computation in (II.26) by taking $y_0 = \phi$ and $q_T = \phi$, we get

$$C(\phi, 0)^2 \geq \frac{2\mathcal{R}e\,(\lambda)e^{-2\mathcal{R}e\,(\lambda)T}}{\|B^*\phi\|_U^2}.$$

  This is a much better behavior : if $B^*\phi$ remains away from zero, the lower bound of the cost exponentially decreases when $\mathcal{R}e\,(\lambda)$ increases. Of course, this is only a lower bound and thus it does not give any information on the boundedness of $C(\phi, 0)$ itself but it seems to be reasonable to expect null controllability for a dissipative system, and bounds that are in some sense, uniform in λ.

Observe that, in both cases, the observability cost for one single mode ϕ depends on the size of $\|B^*\phi\|_U$. The smaller this quantity is, the larger is the observability cost.

Global notions If we want to come back to more global properties (namely that are independent of the initial data and of the target) we have the following characterisations.

Theorem II.7.25

1. *System (II.1) is exactly controllable at time T if and only if for some $C_{obs,exact} \geq 0$ we have*

$$\|q_T\|_E^2 \leq C_{obs,exact}^2 \int_0^T \|B^* e^{-sA^*} q_T\|_U^2\, ds, \quad \forall q_T \in \mathbb{R}^n.$$

   If this inequality holds, then for any y_0, y_T there exists a control $v \in \mathrm{adm}(y_0, y_T)$ such that

$$\|v\|_{L^2(0,T;U)} \leq C_{obs,exact}\|y_T - e^{-TA}y_0\|_E.$$

2. *System (II.1) is null-controllable at time T if and only if for some $C_{obs,null} \geq 0$ we have*

$$\|e^{-TA^*} q_T\|_E^2 \leq C_{obs,null}^2 \int_0^T \|B^* e^{-sA^*} q_T\|_U^2\, ds, \quad \forall q_T \in \mathbb{R}^n.$$

   If this inequality holds, then for any y_0 there exists a control $v \in \mathrm{adm}(y_0, 0)$ such that

$$\|v\|_{L^2(0,T;U)} \leq C_{obs,null}\|y_0\|_E.$$

Of course, in the finite dimensional setting the two notions are equivalent but the values of the constants $C_{obs,exact}$ and $C_{obs,null}$ may not be the same.

Exercise II.7.26 (Asymptotics of the observability constants, see [Sei88])

The above observability constants actually depend on the control time T and it is clear that this cost should blow up when T gets smaller.
More precisely, we can show (by mentioning explicitly the dependence in T of the consant) that

$$C_{obs,exact,T} \underset{T\to 0}{\sim} \frac{\gamma}{T^{K+\frac{1}{2}}},$$

where K is the smallest integer such that

$$\mathrm{rank}(B|AB|\dots|A^K B) = n,$$

and $\gamma > 0$ is a computable constant depending only on A and B.

Chapter III

Controllability of abstract parabolic PDEs

III.1 General setting

Let us consider now an abstract setting : E and U are two Hilbert spaces

- $\mathcal{A} : D(\mathcal{A}) \subset E \to E$ is some unbounded operator[1] such that $-\mathcal{A}$ generates a strongly continuous semi-group in E. The semi-group will be denoted by $t \mapsto e^{-t\mathcal{A}} \in L(E)$. We refer to usual textbooks in functional analysis for precise definition of those concepts (see for instance [Bre83], [Cor07, Appendix A], [TW09], [EN00]. We will also give a simple contruction of the heat semi-group at the beginning of Chapter IV.

  We recall that a necessary and sufficient condition for the existence of this semigroup is (Hille-Yosida theorem) that $D(\mathcal{A})$ is dense in E and

  $$\exists \omega \in \mathbb{R}, M \geq 1, \text{ s.t. } (\lambda I + \mathcal{A}) \text{ is invertible for any } \lambda > \omega \text{ and } \|(\lambda I + \mathcal{A})^{-m}\| \leq M(\lambda - \omega)^{-m}, \forall m \geq 0.$$

  We will sometimes need to assume that the semi-group is analytic which means that there exists an analytic extension $z \mapsto e^{-z\mathcal{A}}$ in a sector of $\mathbb{C}$ of the form

  $$S_\alpha = \{z \in \mathbb{C},\ \mathcal{R}e\,(z) \geq 0, \text{ and } |\mathcal{I}m\, z| \leq \alpha \mathcal{R}e\, z\},$$

  for some $\alpha > 0$. This property always holds in the case for parabolic equations.

  The adjoint semi-group will be denoted by $t \mapsto e^{-t\mathcal{A}^*}$.

- $\mathcal{B} : U \to D(\mathcal{A}^*)'$ the control operator. It is actually more easy to work with the adjoint $\mathcal{B}^*$ of $\mathcal{B}$, which is, by definition an operator from $D(\mathcal{A}^*)$ into U (since we identify U with its dual space).

- We assume that $\mathcal{B}$ is admissible in the following sense

  $$\left(s \mapsto \mathcal{B}^* e^{-s\mathcal{A}^*} q_T\right) \in L^2(0, T; U),\ \ \forall q_T \in E,$$

  and moreover, there exists a $C > 0$ such that

  $$\int_0^T \|\mathcal{B}^* e^{-(T-t)\mathcal{A}^*} q_T\|_U^2 \, dt \leq C^2 \|q_T\|_E^2,\ \ \forall q_T \in E.$$

  In practice, it is enough to check the above inequality for $q_T \in D(\mathcal{A}^*)$ since $D(\mathcal{A}^*)$ is dense in E.

The (formal) control problem we are looking at is the following

$$\begin{cases} \partial_t y + \mathcal{A} y = \mathcal{B} v \text{ in }]0, T[, \\ \quad y(0) = y_0. \end{cases} \tag{III.1}$$

[1] let say self-adjoint with compact resolvent, if you want to simplify

The suitable meaning we give to this problem is by duality.

Theorem III.1.1 (Well-posedness in a dual sense)

For any $y_0 \in E$ and $v \in L^2(0,T;U)$, there exists a unique $y = y_{v,y_0} \in \mathcal{C}^0([0,T],E)$ such that

$$\langle y(t), q_t\rangle_E - \langle y_0, e^{-t\mathcal{A}^*}q_t\rangle_E = \int_0^t \langle v(s), \mathcal{B}^* e^{-(t-s)\mathcal{A}^*}q_t\rangle_U\, ds, \ \ \forall t \in [0,T], \forall q_t \in E.$$

Moreover, there exists $C > 0$ such that

$$\sup_{t\in[0,T]} \|y(t)\|_E \leq C(\|y_0\|_E + \|v\|_{L^2(0,T;U)}).$$

Proof :

This is a consequence of the admissibility assumption for $\mathcal{B}$ and of the Riesz representation theorem.

- Let us fix a $t \in [0,T]$. We consider the linear map

$$q_t \in E \longmapsto \langle y_0, e^{-t\mathcal{A}^*}q_t\rangle_E + \int_0^t \langle v(s), \mathcal{B}^* e^{-(t-s)\mathcal{A}^*}q_t\rangle_U\, ds.$$

  Thanks to the admissibility condition for $\mathcal{B}$, we see that this linear map is continuous on E. Thanks to the Riesz representation theorem, we deduce that there exists a unique element $y_t \in E$ satisfying the equality

$$\langle y_t, q_t\rangle_E = \langle y_0, e^{-t\mathcal{A}^*}q_t\rangle_E + \int_0^t \langle v(s), \mathcal{B}^* e^{-(t-s)\mathcal{A}^*}q_t\rangle_U\, ds, \ \ \forall q_t \in E.$$

  Additionally, we have the bound

$$\|y_t\|_E \leq C(\|y_0\|_E + \|v\|_{L^2(0,T;U)}),$$

  for some constant $C > 0$.

- We set $y(t) = y_t$ for any t. It is clear, by definition, that $y(0) = y_0$. It remains to check that the map y is strongly continuous in time.

  Let $(t_n)_n \subset [0,T]$ a sequence that converges towards some $t \in [0,T]$, we need to prove that $y(t_n) \to y(t)$ in E. To this end, we consider $(q_{t_n})_n \subset E$ a sequence that weakly converges towards some $q_t \in E$ and we want to show that

$$\langle y(t_n), q_{t_n}\rangle_E \xrightarrow[n\to\infty]{} \langle y(t), q_t\rangle_E.$$

  We consider $\bar{v} \in L^2(\mathbb{R})$ the extension of v by zero outside the interval $(0,T)$. We can write

$$\begin{aligned}\langle y(t_n), q_{t_n}\rangle_E &= \langle y_0, e^{-t_n\mathcal{A}^*}q_{t_n}\rangle_E + \int_0^{t_n} \langle v(s), \mathcal{B}^* e^{-(t_n-s)\mathcal{A}^*}q_{t_n}\rangle_U\, ds \\ &= \langle e^{-t_n\mathcal{A}}y_0, q_{t_n}\rangle_E + \int_0^{t_n} \langle v(t_n-s), \mathcal{B}^* e^{-s\mathcal{A}^*}q_{t_n}\rangle_U\, ds \\ &= \langle e^{-t_n\mathcal{A}}y_0, q_{t_n}\rangle_E + \int_0^{T} \langle \bar{v}(t_n-s), \mathcal{B}^* e^{-s\mathcal{A}^*}q_{t_n}\rangle_E\, ds.\end{aligned}$$

  The first term is treated by the weak-strong convergence property and using the strong continuity of the semi-group. The second term is treated in the same way by using:

  - The admissibility condition that leads to the weak convergence of $s \mapsto \mathcal{B}^* e^{-s\mathcal{A}^*}q_{t_n}$ in $L^2(0,T,U)$ and the strong convergence of the translations $s \mapsto \bar{v}(t_n - s)$ in $L^2(0,T,U)$.

■

Actually, we shall also encounter cases where the admissibility condition for $\mathcal{B}$ does not hold exactly as written above. More precisely, assume that there exists an Hilbert space F continuously and densely embedded in E and such that

$$\left(t \mapsto \mathcal{B}^* e^{-s\mathcal{A}^*} q_T\right) \in L^2(0,T;U), \ \ \forall q_T \in F,$$

and

$$\int_0^T \|\mathcal{B}^* e^{-(T-t)\mathcal{A}^*} q_T\|_U^2 \, dt \leq C^2 \|q_T\|_F^2, \ \ \forall q_T \in F.$$

In that case, we may consider the dual space F' (more precisely, its representation obtained by using E as a pivot space) and prove the following result

Theorem III.1.2 (Well-posedness in a dual sense - weaker form)

Under the assumptions above, for any $y_0 \in E$ and $v \in L^2(0,T;U)$, there exists a unique $y = y_{v,y_0} \in \mathcal{C}^0([0,T],F')$ such that

$$\langle y(t), q_t\rangle_{F',F} - \langle y_0, e^{-t\mathcal{A}^*} q_t\rangle_E = \int_0^t \langle v(s), \mathcal{B}^* e^{-(t-s)\mathcal{A}^*} q_t\rangle_U \, ds, \ \ \forall t \in [0,T], \forall q_t \in F.$$

Moreover, if F is stable by the semi-group generated by $\mathcal{A}^*$, the above definition can be extended to any initial data $y_0 \in F'$.

Here also we have seen the important role played by the adjoint problem (which is a backward in time parabolic problem)

$$-\partial_t q + \mathcal{A}^* q = 0, \tag{III.2}$$

III.2 Examples

Let Ω be a bounded smooth connected domain of $\mathbb{R}^d$. Let ω be a non empty open subset of Ω and Γ_0 a non empty open subset of $\partial\Omega$.

- **Distributed control for the heat equation.**

  We consider the problem

$$\begin{cases} \partial_t y - \Delta y = \mathbf{1}_\omega v, & \text{in } \Omega \\ y = 0, & \text{on } \partial\Omega. \end{cases}$$

  The natural state space is $E = L^2(\Omega)$, the control space is also $U = L^2(\Omega)$ (we could have defined $U = L^2(\omega)$ without any real difference), the domain of $\mathcal{A}$ is $D(\mathcal{A}) = H^2(\Omega) \cap H_0^1(\Omega)$, and the control operator is $\mathcal{B} = 1_\omega$, so that we get also $\mathcal{B}^* = 1_\omega$.

- **(Dirichlet) Boundary control for the heat equation.**

  Let us consider the problem

$$\begin{cases} \partial_t y - \Delta y = 0, & \text{in } \Omega \\ y = \mathbf{1}_{\Gamma_0} v, & \text{on } \partial\Omega. \end{cases}$$

  Here the control operator $\mathcal{B}$ is not so easy to define and it is in fact easier to define its adjoint $\mathcal{B}^*$ (through a formal integration by parts). More precisely, we set

$$\mathcal{B}^* \stackrel{\text{def}}{=} 1_{\Gamma_0} \partial_n.$$

In order for the admissibility condition for this operator to hold, we see that we have, for instance, to work in the space $F = H_0^1(\Omega)$. Indeed, in that case, one can show by standard arguments that

$$t \mapsto e^{-t\mathcal{A}^*} q_T \in L^2(0,T,H^2(\Omega)), \quad \forall q_T \in F,$$

and by trace theorems

$$t \mapsto \partial_n(e^{-t\mathcal{A}^*} q_T) \in L^2(0,T,H^{1/2}(\partial\Omega)) \subset L^2(0,T,L^2(\partial\Omega)).$$

Actually, one may use for any any of the spaces $F = D(\mathcal{A}^s)$ with $s > 1/2$.

- **Distributed control for parabolic systems.**

  In the last part of the course, we will be interested in coupled parabolic systems, as for instance the following problem

$$\begin{cases} \partial_t y - \Delta y + C(t,x)y = \mathbf{1}_\omega Bv, & \text{in } \Omega \\ y = 0, & \text{on } \partial\Omega, \end{cases} \tag{III.3}$$

  where y is now a n-component function. The state space is $E = (L^2(\Omega))^n$, the control space is $U = (L^2(\Omega))^m$, $B \in M_{n,m}(\mathbb{R})$ is the control matrix and $C(t,x) \in M_{n,n}(\mathbb{R})$ is the coupling matrix.

  In that case, the control operator is $\mathcal{B} = 1_\omega B$ and its adjoint is $\mathcal{B}^* = 1_\omega B^*$.

- **(Dirichlet) Boundary control for parabolic systems.**

  Similarly, we can consider the boundary control problem

$$\begin{cases} \partial_t y - \Delta y + C(t,x)y = 0, & \text{in } \Omega \\ y = \mathbf{1}_{\Gamma_0} Bv, & \text{on } \partial\Omega. \end{cases} \tag{III.4}$$

  The definition of the functional spaces and of the operator are clear.

- **More general examples:**

  Of course we may consider a large number of other examples such as : time- and or space-dependent diffusion coefficients, different diffusion operators for each component, first or second order coupling terms, non linear terms, etc ...

III.3 Controllability - Observability

The general definitions for approximate/exact/null- controllability questions are formally the same as before.

We have already seen in the first chapter that exact controllability for parabolic equations is certainly not a suitable notion. We may in fact prove that, in general, the set of reachable functions for the heat equation with a distributed control supported on a strict subset of Ω is a very small set. For instance, usual regularity properties for such PDEs show that any reachable target must be smooth (at least $\mathcal{C}^\infty$) in $\Omega \setminus \overline{\omega}$.

We will thus restrict our attention now on the approximate and null-controllability properties. By adapting the arguments given in the finite dimensional case, we can prove the following properties.

Theorem III.3.3 (Approximate controllability and Unique continuation)

Our system (III.1) is approximately controllable at time $T > 0$ if and only if the adjoint system (III.2) satisfies the unique continuation property with respect to the observation operator $\mathcal{B}^*$, namely : for any solution q of (III.2) with $q(T) \in F$, we have

$$\Big(\mathcal{B}^* q(t) = 0, \forall t \in (0,T)\Big) \Longrightarrow q \equiv 0.$$

With the semi-group notation, the Unique Continuation property writes

$$\left(\mathcal{B}^* e^{-(T-t)\mathcal{A}^*} q_T = 0, \forall t \in (0,T)\right) \Longrightarrow q_T = 0.$$

Notice that, if the semi-group generated by $-\mathcal{A}^*$ is analytic, then the unique continuation property does not depend on T, and thus so is the approximate controllability.

Proof :

- Assume that the Unique Continuation property does not hold. There exists $q_T \in F$, $q_T \neq 0$ such that $\mathcal{B}^* e^{-(T-t)\mathcal{A}^*} q_T = 0$. By definition, for any control v, we have

$$\langle y(T), q_T \rangle_{F',F} - \langle y_0, e^{-T\mathcal{A}^*} q_T \rangle_E = \int_0^T \langle v(s), \mathcal{B}^* e^{-(T-s)\mathcal{A}^*} q_T \rangle_U \, ds = 0, \tag{III.5}$$

  and if follows that

$$\langle y(T) - e^{-T\mathcal{A}} y_0, q_T \rangle_{F',F} = 0,$$

  which proves that the reachable space at time T cannot be dense in F'. Indeed, if $z \in F'$ is any element such that $\langle z, q_T \rangle_{F',F} \neq 0$, then $e^{-T\mathcal{A}} y_0 + \varepsilon z$ is not reachable for any $\varepsilon > 0$.

- Assume that the approximate controllability does not hold in F'. By the Hahn-Banach theorem, it means that there exists a $y_T \in F'$ and a $q_T \in F \setminus \{0\}$ such that

$$\langle y_T, q_T \rangle_{F',F} \geq \langle y(T), q_T \rangle_{F',F},$$

  for any control $v \in L^2(0,T,U)$.

  From (III.5) we deduce that, for any $v \in L^2(0,T,U)$

$$\int_0^T \langle v(s), \mathcal{B}^* e^{-(T-s)\mathcal{A}^*} q_T \rangle_U \, ds \leq \langle y_T - e^{-T\mathcal{A}} y_0, q_T \rangle_{F',F}.$$

  We apply this inequality to $v = \frac{1}{\delta} \mathcal{B}^* e^{-(T-s)\mathcal{A}^*} q_T$, with $\delta > 0$, which gives

$$\frac{1}{\delta} \int_0^T \| \mathcal{B}^* e^{-(T-s)\mathcal{A}^*} q_T \|_U^2 \, ds \leq \langle y_T - e^{-T\mathcal{A}} y_0, q_T \rangle_{F',F}.$$

  Letting δ going to 0 leads to

$$\int_0^T \| \mathcal{B}^* e^{-(T-s)\mathcal{A}^*} q_T \|_U^2 \, ds = 0$$

  and since $q_T \neq 0$, we obtained that the unique continuation property does not hold for the adjoint problem.

■

Theorem III.3.4 (Null controllability and Observability)

Our system (III.1) is null-controllable in E at time $T > 0$ if and only if the adjoint system (III.2) satisfies the following observability property with respect to the observation operator $\mathcal{B}^*$, namely :
There exists a $C > 0$ such that for any solution q of (III.2) with $q(T) \in F$, we have

$$\|q(0)\|_E^2 \leq C^2 \int_0^T \|\mathcal{B}^* q(t)\|_U^2 \, dt.$$

With the semi-group notation, the observability inequality writes

$$\|e^{-T\mathcal{A}^*}q_T\|_E^2 \leq C^2 \int_0^T \|\mathcal{B}^* e^{-(T-t)\mathcal{A}^*}q_T\|_U^2\,dt, \;\; \forall q_T \in F.$$

Remark III.3.5

If we are interested in the null-controllability with initial data in F', then the above inequalities should hold with $\|q(0)\|_F^2$ in the left-hand side.

Proof :

This result is a straightforward consequence of the following general result in functional analysis (which is itself a consequence of the closed graph theorem).

Lemma III.3.6 (see Proposition 12.1.2 in [TW09])

Let H_1, H_2, H_3 be three Hilbert spaces and $F : H_1 \to H_3$, $G : H_2 \to H_3$ be two bounded linear operators. Then the following properties are equivalent

1. *The range of F is included in the range of G.*
2. *There exists a $C > 0$ such that the following inequalities hold*

$$\|F^* x\|_{H_1} \leq C\|G^* x\|_{H_2}, \;\; \forall x \in H_3.$$

If those properties are true, there exists a bounded linear operator $L : H_1 \to H_2$ such that

$$F = G \circ L, \text{ and } \|L\|_{H_1 \to H_2} \leq C.$$

To prove the theorem, we apply the previous lemma with $H_2 = L^2(0,T;U)$, $H_1 = H_3 = E$, and

$$F : y_0 \in E \mapsto e^{-TA}y_0 \in E,$$

$$G : v \in L^2(0,T,U) \mapsto \int_0^T e^{-(T-s)\mathcal{A}}Bv(s)\,ds \in E,$$

(this integral being well-defined by duality as seen before).

■

There is no natural (and easy to manage) generalization of the Kalman rank criterion in the infinite dimension case. However, the Fattorini-Hautus test still holds under quite general assumptions but it will of course only gives an approximate controllability result .

Theorem III.3.7 (Fattorini-Hautus test)

Assume that:

- *$\mathcal{A}^*$ has a compact resolvant and a complete system of root vectors.*
- *$\mathcal{B}^*$ is a bounded operator from $D(\mathcal{A}^*)$ (with the graph norm) into U.*

We also assume that the semi-group generated by $-\mathcal{A}^*$ is analytic, even though the result can be adapted if it is not the case.
Then, our system (III.1) is approximately controllable at time $T > 0$ if and only if we have

$$(\mathit{Ker}\,\mathcal{B}^*) \cap \mathit{Ker}\,(\mathcal{A}^* - \lambda \mathrm{I}) = \{0\}, \;\; \forall \lambda \in \mathbb{C}.$$

In particular, the approximate controllability property does not depend on T.

For a proof of this result in the framework above which is more general than the original one by Fattorini, we refer to [Oli14].

Chapter IV

The heat equation

In this chapter we are interested in the controllability properties of a parabolic scalar equation of the heat type in a bounded domain. We will actually be a little bit more general by looking at the following equation.

Let Ω be a bounded connected smooth domain of $\mathbb{R}^d$. Let $\gamma \in \mathcal{C}^0(\overline{\Omega}, \mathbb{R})$ be a diffusion coefficient such that $\gamma_{\min} \stackrel{\text{def}}{=} \inf_\Omega \gamma > 0$ and $\alpha \in \mathcal{C}^0(\overline{\Omega}, \mathbb{R})$ a potential term. Let $\mathcal{A}$ be the differential operator defined by

$$(\mathcal{A}y)(x) = -\text{div}\,(\gamma(x)\nabla y) + \alpha(x)y. \tag{IV.1}$$

We shall consider the partial differential evolution equation given by

$$\partial_t y + \mathcal{A}y = 0, \text{ in } (0,T) \times \Omega. \tag{IV.2}$$

If we look at $\mathcal{A}$ as an unbounded operator in $L^2(\Omega)$ with domain $D(\mathcal{A}) = H^2(\Omega) \cap H^1_0(\Omega)$, we know that $\mathcal{A}$ is self-adjoint and with compact resolvent. As a consequence, we have a complete spectral theory for this operator:

- The spectrum Λ of $\mathcal{A} = \mathcal{A}^*$ is only made of positive eigenvalues, moreover Λ is locally finite, unbounded but satisfies the bound from below
$$\inf \Lambda > \inf_\Omega \alpha. \tag{IV.3}$$
- For each $\lambda \in \Lambda$, the eigenspace Ker $(\mathcal{A} - \lambda)$ is finite dimensional and we have the orthogonality property in $L^2(\Omega)$
$$\text{Ker}\,(A - \lambda) \perp \text{Ker}\,(A - \mu), \quad \forall \lambda \neq \mu \in \Lambda.$$
We denote by π_λ the orthogonal projection in $L^2(\Omega)$ onto the eigenspace Ker $(\mathcal{A} - \lambda)$.
- We have an orthogonal spectral decomposition of the space $L^2(\Omega)$. This means that for any $\psi \in L^2(\Omega)$ we have
$$\psi = \sum_{\lambda \in \Lambda} \pi_\lambda \psi, \tag{IV.4}$$
this family being summable in $L^2(\Omega)$, and we have the Bessel-Parseval equality
$$\|\psi\|^2_{L^2(\Omega)} = \sum_{\lambda \in \Lambda} \|\pi_\lambda \psi\|^2_{L^2(\Omega)}.$$
- For any $\psi \in H^1_0(\Omega)$, the sum (IV.4) is also converging in $H^1_0(\Omega)$ and there exists $C_1, C_2 > 0$, depending only on the coefficients γ and α, such that
$$C_1 \sum_{\lambda \in \Lambda} (1 + |\lambda|) \|\pi_\lambda \psi\|^2_{L^2} \leq \|\psi\|^2_{H^1} \leq C_2 \sum_{\lambda \in \Lambda} (1 + |\lambda|) \|\pi_\lambda \psi\|^2_{L^2}.$$

- $-\mathcal{A}$ generates a semi-group that can be explicitely computed as follows
$$e^{-t\mathcal{A}}\psi = \sum_{\lambda\in\Lambda} e^{-t\lambda}\pi_\lambda\psi, \ \ \forall \psi \in L^2(\Omega).$$
Notice in particular the following energy estimate
$$\|e^{-t\mathcal{A}}\psi\|_{L^2(\Omega)} \leq e^{-t\inf\Lambda}\|\psi\|_{L^2(\Omega)}, \ \ \forall \psi\in E, \forall t\geq 0. \tag{IV.5}$$
In the case where $\inf\Lambda > 0$, we see that the system is dissipative in $L^2(\Omega)$, see Remark IV.0.1.

- We shall need the following spaces
$$E_\mu \stackrel{\text{def}}{=} \bigoplus_{\substack{\lambda\in\Lambda\\ \lambda\leq\mu}} \text{Ker }(\mathcal{A}-\lambda). \tag{IV.6}$$
Let P_μ be the orthogonal projection in L^2 onto E_μ, which can be expressed as follows
$$P_\mu = \sum_{\substack{\lambda\in\Lambda\\ \lambda\leq\mu}} \pi_\lambda.$$
We can prove the following additional dissipation property
$$\|e^{-t\mathcal{A}}\psi\|_{L^2(\Omega)} \leq e^{-t\mu}\|\psi\|_{L^2(\Omega)}, \ \ \forall\psi\in E, \text{s.t. } P_\mu\psi = 0, \ \forall t\geq 0. \tag{IV.7}$$

We will see in the sequel that other qualitative properties for the spectrum of the operator will be needed to analyze the controllability of the system.

We will analyze two types of controls:

- The distributed control problem: Let ω be a non empty open subset of Ω. We look for a control $v\in L^2(]0,T[\times\omega) = L^2(0,T;U)$ with $U = L^2(\omega)$ such that the solution $y\in\mathcal{C}^0([0,T],E)$, with $E = L^2(\Omega)$, of the problem
$$\begin{cases} \partial_t y + \mathcal{A}y = \mathbf{1}_\omega v, \ \text{in } \Omega,\\ y = 0, \ \text{on } \partial\Omega,\\ y(0) = y_0 \end{cases} \tag{IV.8}$$
satisfies either $\|y(T) - y_T\|_E \leq \varepsilon$ (approximate controllability) or $y(T) = 0$ (null-controllability).

- The boundary control problem: Let Γ_0 be a non empty open subset of Γ. We look for a control $v\in L^2(]0,T[\times\Gamma_0) = L^2(0,T;U)$ with $U = L^2(\Gamma_0)$ such that the solution $y\in\mathcal{C}^0([0,T],E)$, with $E = H^{-1}(\Omega)$, of the problem
$$\begin{cases} \partial_t y + \mathcal{A}y = 0, \ \text{in } \Omega,\\ y = \mathbf{1}_{\Gamma_0} v, \ \text{on } \partial\Omega,\\ y(0) = y_0 \end{cases} \tag{IV.9}$$
satisfies either $\|y(T) - y_T\|_E \leq \varepsilon$ (approximate controllability) or $y(T) = 0$ (null-controllability).

Remark IV.0.1

From the point of view of controllability we can always assume, if necessary, that the potential α is non negative, which implies $\inf\Lambda > 0$ *(see (IV.3)), and thus all the eigenvalues are positive.*
Indeed, is one sets $\tilde{y} = e^{-at}y$ we see that $\tilde{y}$ solves the problem
$$\begin{cases} \partial_t\tilde{y} + (\mathcal{A}+a)\tilde{y} = \mathbf{1}_\omega e^{-at}v, \ \textit{in } \Omega,\\ y = 0, \ \textit{on } \partial\Omega,\\ y(0) = y_0, \end{cases}$$
which amounts at adding the constant a to α.

As a consequence of the previous remark, we will systematically assume in the sequel that $\alpha\geq 0$.

IV.1 Further spectral properties and applications

IV.1.1 The 1D case

We assume in this section that $\Omega = (0,1)$. From a spectral point of view this particularly implies that all the eigenvalues are simple, therefore we can choose one eigenfunction ϕ_λ in each eigenspace Ker $(\mathcal{A} - \lambda)$, that we shall take normalized in $L^2(\Omega)$. The projection operator π_λ is thus simply given for any $\lambda \in \Lambda$ by

$$\pi_\lambda \psi = \langle \psi, \phi_\lambda \rangle_{L^2} \phi_\lambda, \quad \forall \psi \in L^2(\Omega).$$

The second property which is specific to the 1D case[1] is the following asymptotic property, called Weyl's law

$$N(r) \underset{r\to\infty}{\sim} \bar{N}\sqrt{r},$$

for some constant $\bar{N} > 0$, where N is the counting function of the family Λ (see Section A.6). We will present a proof of a weaker (but sufficient) version of this result below.

IV.1.1.1 Spectral estimates

The properties stated in this section are very classical but we adopt here the formalism and proofs introduced in [ABM16] that have the advantage to being easy to extend to more general situations like the discrete setting for instance.

Proposition IV.1.2

Under the assumptions above, for both boundary and distributed control problems, we have

$$\mathcal{B}^* \phi_\lambda \neq 0, \ \forall \lambda \in \Lambda.$$

In particular, the heat equation is approximately controllable at any time $T > 0$ in both cases.

Proof :

In both cases, if we assume that $\mathcal{B}^*\phi_\lambda = 0$, it implies that there exists a point $a \in [0,1]$ such that $\phi_\lambda(a) = \phi'_\lambda(a) = 0$. Indeed, we either take a to be a boundary point of Ω, or a point inside the control domain ω.

Since ϕ_λ satisfies a second order linear homogeneous differential equation, this would imply $\phi_\lambda \equiv 0$ which is impossible.

The approximate controllability in both cases is now a consequence of the Fattorini-Hautus test (see Theorem III.3.7). ■

Let us introduce the notations

$$\partial_l \phi \stackrel{\text{def}}{=} -\phi'(0), \quad \text{and} \quad \partial_r \phi \stackrel{\text{def}}{=} \phi'(1),$$

for the left and right normal derivatives of a function $\phi : (0,1) \to \mathbb{R}$.

Theorem IV.1.3

Under the assumptions above, there exists $C_1(\alpha,\gamma,\omega) > 0$ and $C_2(\alpha,\gamma), C_3(\alpha,\gamma) > 0$ such that

$$\|\phi_\lambda\|^2_{L^2(\omega)} \geq C_1(\alpha,\gamma,\omega), \ \forall \lambda \in \Lambda,$$

$$|\partial_\bullet \phi_\lambda| \geq C_2(\alpha,\gamma)\sqrt{\lambda}, \ \forall \lambda \in \Lambda, \forall \bullet \in \{l,r\},$$
$$|\lambda - \mu| \geq C_2(\alpha,\gamma)\sqrt{\lambda}, \ \forall \lambda \neq \mu \in \Lambda,$$
$$N_\Lambda(r) \leq C_3(\alpha,\gamma)\sqrt{r}, \quad \forall r > 0.$$

[1] Weyl's law also holds in higher dimension but it becomes $N(r) \sim \bar{N} r^{\frac{2}{d}}$, where d is the space dimension

Remark IV.1.4 (Laplace operator)

For the standard Laplace operator $\gamma = 1$, $\alpha = 0$, the eigenfunctions and eigenvalues are explicitely given by

$$\Lambda = \{k^2\pi^2,\ k \in \mathbb{N}^*\},$$
$$\phi_\lambda(x) = \sqrt{2}\sin(\sqrt{\lambda}x),\quad \lambda \in \Lambda.$$

The properties proved in the above theorem are thus straightforward it this case. Moreover, there are clearly optimal.

We begin with the following lemma.

Lemma IV.1.5

Let ω be a non-empty open subset of Ω. There exists $C_1(\alpha,\gamma) > 0$ and $C_2(\alpha,\gamma,\omega) > 0$ such that we have, for any $\lambda \in \Lambda$,

$$\frac{1}{\lambda}|\partial_\bullet \phi_\lambda|^2 \geq C_1(\alpha,\gamma)\,\mathcal{R}_\lambda,\quad \forall \bullet \in \{l, r\},$$

and

$$\|\phi_\lambda\|^2_{L^2(\omega)} \geq C_2(\alpha,\gamma,\omega)\,\mathcal{R}_\lambda,$$

where we have defined

$$\mathcal{R}_\lambda \stackrel{\text{def}}{=} \inf_{x,y\in\Omega} \frac{|\phi_\lambda(x)|^2 + \frac{\gamma(x)}{\lambda}|\phi'_\lambda(x)|^2}{|\phi_\lambda(y)|^2 + \frac{\gamma(y)}{\lambda}|\phi'_\lambda(y)|^2}. \tag{IV.10}$$

Proof :

- By definition of $\mathcal{R}_\lambda$, and the fact that $\phi_\lambda(0) = 0$, we have

$$\frac{\gamma(0)}{\lambda}|\phi'_\lambda(0)|^2 \geq \mathcal{R}_\lambda\left(|\phi_\lambda(y)|^2 + \frac{\gamma(y)}{\lambda}|\phi'_\lambda(y)|^2\right) \geq \mathcal{R}_\lambda|\phi_\lambda(y)|^2,\quad \forall y \in \Omega.$$

  By integration over $y \in \Omega$, we can use the normalisation condition and the equation satisfied by ϕ_λ to find that

$$\frac{\gamma(0)}{\lambda}|\phi'_\lambda(0)|^2 \geq \mathcal{R}_\lambda.$$

  For λ large enough, we deduce that

$$\frac{\gamma(0)}{\lambda}|\phi'_\lambda(0)|^2 \geq \mathcal{R}_\lambda,$$

  which gives the claim for $\partial_l\phi_\lambda$. A similar proof holds for $\partial_r\phi_\lambda$.

- Let $(a, b) \subset \omega$ be a connected component of ω. The Sturm comparison theorem (see Theorem A.5.15 and Corollary A.5.16) implies that there is a $\lambda_0(\alpha,\gamma,\omega)$ such that for $\lambda \geq \lambda_0$, we can find two zeros $a_\lambda < b_\lambda$ of ϕ_λ such that $(a_\lambda, b_\lambda) \subset (a, b)$ and

$$b_\lambda - a_\lambda \geq (b - a)/2. \tag{IV.11}$$

  We multiply by ϕ_λ the equation satisfied by ϕ_λ on (a_λ, b_λ) and we integrate by parts, using that a_λ and b_λ are zeros of ϕ_λ. We obtain

$$\int_{a_\lambda}^{b_\lambda} \gamma|\phi'_\lambda|^2 + \alpha|\phi_\lambda|^2 = \lambda\int_{a_\lambda}^{b_\lambda}|\phi_\lambda|^2,$$

  and since we have assumed that $\alpha \geq 0$, we find that

$$\int_{a_\lambda}^{b_\lambda}|\phi_\lambda|^2 \geq \int_{a_\lambda}^{b_\lambda}\frac{\gamma}{\lambda}|\phi'_\lambda|^2. \tag{IV.12}$$

By definition of $\mathcal{R}_\lambda$ we have, for any $x, y \in \Omega$

$$|\phi_\lambda(x)|^2 + \frac{\gamma(x)}{\lambda}|\phi'_\lambda(x)|^2 \geq \mathcal{R}_\lambda \left(|\phi_\lambda(y)|^2 + \frac{\gamma(y)}{\lambda}|\phi'_\lambda(y)|^2 \right).$$

We can integrate this inequality with respect to $x \in (a_\lambda, b_\lambda)$ on the one hand and with respect to $y \in \Omega = (0,1)$ on the other hand to get

$$\int_{a_\lambda}^{b_\lambda} |\phi_\lambda|^2 + \int_{a_\lambda}^{b_\lambda} \frac{\gamma}{\lambda}|\phi'_\lambda|^2 \geq \mathcal{R}_\lambda(b_\lambda - a_\lambda) \int_0^1 \left(|\phi_\lambda|^2 + \frac{\gamma}{\lambda}|\phi'_\lambda|^2 \right) \geq \mathcal{R}_\lambda(b_\lambda - a_\lambda).$$

By (IV.12), the normalisation condition of ϕ_λ in $L^2(\Omega)$ and (IV.11), we arrive to

$$\int_{a_\lambda}^{b_\lambda} |\phi_\lambda|^2 \geq \mathcal{R}_\lambda \frac{b-a}{4},$$

so that, for $\lambda \geq \lambda_0$, we have

$$\int_\omega |\phi_\lambda|^2 \geq \int_{a_\lambda}^{b_\lambda} |\phi_\lambda|^2 \geq \mathcal{R}_\lambda \frac{b-a}{4}.$$

Since there is a finite number of eigenvalues that satisfy $\lambda < \lambda_0$, the claim is proved thanks to Proposition IV.1.2.

■

Now we propose a reformulation of the differential equation that will permit us to prove uniform lower bounds for the quantity $\mathcal{R}_\lambda$.

Lemma IV.1.6

Let $f : \Omega \to \mathbb{R}$ be a continuous function and $\lambda > 0$. Suppose that $u : \Omega \to \mathbb{R}$ satisfies the second-order differential equation (without any prescribed boundary conditions)

$$\mathcal{A}u(x) = \lambda u(x) + f(x), \quad \forall x \in \Omega, \tag{IV.13}$$

then the following equation holds

$$U'(x) = M(x)U(x) + Q(x)U(x) + F(x), \quad \forall x \in \Omega, \tag{IV.14}$$

where we have defined the vectors

$$U(x) \stackrel{\text{def}}{=} \begin{pmatrix} u(x) \\ \sqrt{\frac{\gamma(x)}{\lambda}} u'(x) \end{pmatrix} \text{ and } F(x) \stackrel{\text{def}}{=} \begin{pmatrix} 0 \\ -\frac{f(x)}{\sqrt{\gamma(x)\lambda}} \end{pmatrix}.$$

and the matrices

$$M(x) \stackrel{\text{def}}{=} \begin{pmatrix} 0 & \sqrt{\frac{\lambda}{\gamma(x)}} \\ -\sqrt{\frac{\lambda}{\gamma(x)}} & 0 \end{pmatrix} \text{ and } Q(x) \stackrel{\text{def}}{=} \begin{pmatrix} 0 & 0 \\ \frac{\alpha(x)}{\sqrt{\lambda\gamma(x)}} & \sqrt{\gamma(x)} \left(\frac{1}{\sqrt{\gamma}} \right)'(x) \end{pmatrix}.$$

The key-point of this formulation is that the large terms in $\sqrt{\lambda}$ only appear in the skew-symmetric matrix $M(x)$, while the matrix $Q(x)$ only contain bounded terms with respect to λ.

As a consequence of this particular structure, we can obtain the following estimates.

Lemma IV.1.7

With the same notations as in Lemma IV.1.6, and assuming that $\lambda \geq 1$, there exists $C \stackrel{\text{def}}{=} C(\alpha,\gamma)$, independent of λ, such that for any $x, y \in \Omega$, we have

$$\|U(y)\| \leq C(\alpha,\gamma)\left(\|U(x)\| + \left|\int_x^y \|F(s)\|\,ds\right|\right). \tag{IV.15}$$

Proof :

Let $x, y \in \Omega$. Without loss of generality we assume $y > x$. It is fundamental to notice that the matrices $(M(s))_s$ pairwise commute, so that the resolvant operator associated with $x \mapsto M(x)$ simply reads

$$S(y,x) \stackrel{\text{def}}{=} \exp\left(\int_x^y M(s)\,ds\right).$$

We can then use Duhamel's formula to deduce from the equation (IV.14) the following expression

$$U(y) = S(y,x)U(x) + \int_x^y S(y,s)\left(Q(s)U(s) + F(s)\right)\,ds. \tag{IV.16}$$

We use now the fact that the matrix $M(s)$ is skew symmetric for any s, and so is $\int_x^y M(s)\,ds$. It follows that the resolvant $S(y,s)$ is unitary $\|S(y,s)\| = 1$ for any y, s. We get

$$\|U(y)\| \leq \|U(x)\| + \left|\int_x^y \|F(s)\|\,ds\right| + \left|\int_x^y \|Q(s)\|\|U(s)\|\,ds\right|.$$

Gronwall's lemma finally yields

$$\|U(y)\| \leq \left(\|U(x)\| + \left|\int_x^y \|F(s)\|\,ds\right|\right)\exp\left(\left|\int_x^y \|Q(s)\|\,ds\right|\right),$$

which gives the result since $Q(s)$ is bounded uniformly in s and λ, by using the assumptions on the coefficient γ and α ∎

We can now prove the main Theorem of this section.

Proof (of Theorem IV.1.3):

A first remark is that it is enough to prove the claims for λ large enough and in particular we can assume without loss of generality that $\lambda \geq 1$.

- We begin with the proof of the first two points of the theorem. By definition, ϕ_λ is solution of the equation

$$\mathcal{A}\phi_\lambda = \lambda\phi_\lambda,$$

  which is exactly (IV.13) with $u = \phi_\lambda$ and $f = 0$. From Lemma IV.1.7 we deduce that there exists $C \stackrel{\text{def}}{=} C(\gamma,\alpha)$, independent of λ, such that for any $x, y \in \overline{\Omega}$,

$$|\phi_\lambda(y)|^2 + \frac{\gamma(y)}{\lambda}|\phi_\lambda'(y)|^2 \geq C\left(|\phi_\lambda(x)|^2 + \frac{\gamma(x)}{\lambda}|\phi_\lambda'(x)|^2\right), \tag{IV.17}$$

  which exactly proves that the quantity $\mathcal{R}_\lambda$ defined in (IV.10) is uniformly bounded from below. The claim thus immediately follows from Lemma IV.1.5.

- We shall now prove the third point in Theorem IV.1.3. For any two $\lambda > \mu$ in Λ with $\mu \geq 1$, we define

$$u(x) \stackrel{\text{def}}{=} \phi_\mu'(1)\phi_\lambda(x) - \phi_\lambda'(1)\phi_\mu(x),$$

in such a way that $u(1) = u'(1) = 0$ and

$$\mathcal{A}u = \lambda u + f,$$

with

$$f(x) \stackrel{\text{def}}{=} \phi'_\lambda(1)\,(\lambda-\mu)\,\phi_\mu(x), \quad \forall x \in \Omega.$$

Using the notations introduced in Lemma IV.1.6, we observe that by construction we have $U(1) = 0$ so that the estimate (IV.15) specialized in $x = 1$ leads to

$$\|U(y)\| \leq C \int_y^1 \|F(s)\|\,ds \leq C \int_0^1 \|F(s)\|\,ds, \;\; \forall y \in \Omega.$$

Using the expression for F and f, we find that

$$\|U(y)\| \leq \frac{C}{\sqrt{\gamma_{min}}} \left(\frac{\lambda-\mu}{\sqrt{\lambda}}|\phi'_\lambda(1)|\right) \int_0^1 |\phi_\mu(s)|\,ds, \; \forall y \in \Omega.$$

Thanks to the normalisation condition $\|\phi_\mu\|_{L^2(\Omega)} = 1$ and the expressions of U and u, we obtain for any $y \in \Omega$,

$$\left|\phi'_\mu(1)\phi_\lambda(y) - \phi'_\lambda(1)\phi_\mu(y)\right|^2 \leq \frac{C}{\gamma_{min}} \left(\frac{\lambda-\mu}{\sqrt{\lambda}}|\phi'_\lambda(1)|\right)^2.$$

We integrate this inequality with respect to $y \in (0,1)$ and we use the $L^2(\Omega)$ orthonormality of ϕ_λ and ϕ_μ to finally get

$$|\phi'_\lambda(1)|^2 \leq \left(\phi'_\lambda(1)\right)^2 + \left(\phi'_\mu(1)\right)^2 \leq \frac{C}{\gamma_{min}} \left(\frac{\lambda-\mu}{\sqrt{\lambda}}|\phi'_\lambda(1)|\right)^2,$$

and since $\phi'_\lambda(1) \neq 0$, we conclude that

$$\lambda - \mu \geq \bar{C}\sqrt{\lambda},$$

for some $\bar{C} > 0$ independent of λ and μ.

- Let us finally prove the estimate on the counting function N_Λ. We first observe that the estimate we proved above implies that

$$|\lambda - \mu| \geq \frac{C_1}{2}|\sqrt{\lambda} + \sqrt{\mu}|, \quad \forall \lambda \neq \mu \in \Lambda,$$

from which we deduce

$$|\sqrt{\lambda} - \sqrt{\mu}| \geq \frac{C_1}{2}, \;\; \forall \lambda \neq \mu \in \Lambda. \tag{IV.18}$$

Let us fix $r > 0$ and let $\lambda_1 < \cdots < \lambda_{N_\Lambda(r)}$ all the elements in $\Lambda \cap [0, r]$. We set $\lambda_0 = 0$.

We can write on the one hand

$$\sum_{k=1}^{N_\Lambda(r)} \left(\sqrt{\lambda_k} - \sqrt{\lambda_{k-1}}\right) = \sqrt{\lambda_{N_\Lambda(r)}} \leq \sqrt{r},$$

and on the other hand, by using (IV.18),

$$\sum_{k=1}^{N_\Lambda(r)} \left(\sqrt{\lambda_k} - \sqrt{\lambda_{k-1}}\right) \geq \frac{C_1}{2}(N_\Lambda(r) - 1) + \sqrt{\lambda_1} \geq C N_\Lambda(r),$$

with $C = \min(\sqrt{\lambda_1}, C_1/2)$. Combining the two inequalities above we obtain

$$N_\Lambda(r) \leq \frac{1}{C}\sqrt{r},$$

and the proof is complete.

■

IV.1.1.2 Approximate controllability

The results obtained in Theorem IV.1.3 and the Fattorini-Hautus test (Theorem III.3.7) immediately shows that both problems (IV.8) and (IV.9) are approximately controllable in 1D at any time $T > 0$.

IV.1.1.3 Null-controllability

We shall now prove the null-controllability of (IV.8) and (IV.9), still in 1D, by using the moments method. We already encountered this method in Section II.4 in order to deal with the controllability of finite dimensional linear differential systems.

The main difference here is that there is now a countable infinite number of frequencies in the system. In this framework, this strategy were for instance used in the seminal papers [FR71, FR75].

That is the reason why we will need to be able to prove the existence of a countable biorthogonal family functions to the set of all real exponential functions present in the definition of semigroup of the operator. Moreover, we shall need precise estimate on those families.

To begin with, let us introduce a few notations. For any $\lambda > 0$ we define $e[\lambda] \in L^2(0, +\infty)$ to be the exponential function

$$e[\lambda] := \left(t \longmapsto e^{-\lambda t}\right).$$

When evaluating this function at time t we shall write $e_t[\lambda] = e^{-\lambda t}$. This bracket notation is motivated by the fact that we shall need, later, to use the divided differences formalism recalled in Section A.3.

Let Λ be a family of real numbers in $(0, +\infty)$ such that

$$\sum_{\lambda \in \Lambda} \frac{1}{\lambda} < +\infty. \tag{IV.19}$$

We shall denote by R a remainder function associated to Λ (see Definition A.6.19). We will also assume the following gap condition

$$|\lambda - \mu| \geq \rho, \ \ \forall \lambda \neq \mu \in \Lambda. \tag{IV.20}$$

The theorems we need are the following (their proofs are postponed to Section IV.1.2).

Theorem IV.1.8 (Biorthogonal families of exponential functions in infinite horizon)

Let Λ satisfying (IV.19) and (IV.20).
There exists a family $(q_{\lambda,\infty})_{\lambda \in \Lambda}$ in $L^2(0, +\infty)$ satisfying

$$(q_{\lambda,\infty}, e[\mu])_{L^2(0,+\infty)} = \delta_{\lambda,\mu}, \quad \forall \lambda, \mu \in \Lambda,$$

as well as the estimate

$$\|q_{\lambda,\infty}\|_{L^2(0,+\infty)} \leq e^{\varepsilon(\lambda)\lambda}, \ \ \forall \lambda \in \Lambda,$$

where $\varepsilon : \mathbb{R} \to \mathbb{R}$ is a function such that $\lim_{s \to +\infty} \varepsilon(s) = 0$ that only depends on ρ and R.

Theorem IV.1.9 (Biorthogonal families of exponential functions in finite horizon)

Let Λ satisfying (IV.19) and (IV.20).
For any time $T > 0$, there exists a family $(q_{\lambda,T})_{\lambda \in \Lambda}$ in $L^2(0, T)$ satisfying

$$(q_{\lambda,T}, e[\mu])_{L^2(0,T)} = \delta_{\lambda,\mu}, \quad \forall \lambda, \mu \in \Lambda,$$

and the estimate

$$\|q_{\lambda,T}\|_{L^2(0,T)} \leq K e^{\varepsilon(\lambda)\lambda}, \ \ \forall \lambda \in \Lambda,$$

where $\varepsilon : \mathbb{R} \to \mathbb{R}$ is a function such that $\lim_{s \to +\infty} \varepsilon(s) = 0$ that only depends on ρ and R, and $K > 0$ is a constant depending only on T, ρ and R.

Those results can be made more precise if one has a sharper asymptotic behavior of the family Λ.

Theorem IV.1.10

Let $\bar{N} > 0$ and $\alpha \in (0,1)$.

- *If the counting function of Λ satisfies an a priori bound*

$$N(r) \leq \bar{N} r^{\alpha}, \quad \forall r > 0, \tag{IV.21}$$

then, in Theorems IV.1.8 and IV.1.9, we can take

$$\varepsilon(r) = \frac{C + \log r}{r^{1-\alpha}},$$

where C depends only on α and $\bar{N}$

- *If the counting function of Λ satisfies an even more precise estimate*

$$|N(r) - \bar{N} r^{\alpha}| \leq \tilde{N} r^{\alpha'}, \quad \forall r > 0, \tag{IV.22}$$

with $\tilde{N} > 0$ and $0 \leq \alpha' < \alpha$, then, in Theorems IV.1.8 and IV.1.9, we can take

$$\varepsilon(r) = \frac{C}{r^{1-\alpha}},$$

where C depends only on $\alpha, \alpha', \bar{N}, \tilde{N}$.

Using those results we can deduce the following two null-controllability results as follows.

Theorem IV.1.11 (Boundary null-controllability in 1D)

Assume that $d = 1$, $\Omega = (0,1)$. Let $\Gamma_0 = \{1\}$ for instance. For any $y_0 \in L^2(\Omega)$, and $T > 0$, there exists a control $v \in L^2(0,T)$ such that the solution of (IV.9) satisfies $y(T) = 0$.

Proof :

Let $T > 0$ be given. For any $v \in L^2(0,T)$, the solution y of (IV.9) satisfies

$$\langle y(T), \phi_\lambda \rangle_{H^{-1},H^1_0} - \langle y_0, e^{-\lambda T}\phi_\lambda \rangle_{H^{-1},H^1_0} = \int_0^T v(t) e^{-(T-t)\lambda} \partial_r \phi_\lambda \, dt, \;\; \forall \lambda \in \Lambda.$$

Hence, v is a null-control for our system if and only if the function $u(t) \stackrel{\text{def}}{=} v(T-t)$ satisfies

$$-\langle y_0, e^{-\lambda T}\phi_\lambda \rangle_{L^2} = \int_0^T u(t) e^{-\lambda t} \partial_r \phi_\lambda \, dt, \;\; \forall \lambda \in \Lambda,$$

where we used here that $y_0 \in L^2(\Omega)$. We are thus led to find a function $u \in L^2(0,T)$ that satisfies the following moment problem

$$\int_0^T u(t) e^{-\lambda t} \, dt = \frac{-\langle y_0, \phi_\lambda \rangle_{L^2} e^{-\lambda T}}{\partial_r \phi_\lambda}, \quad \forall \lambda \in \Lambda.$$

From the properties of the eigenvalues Λ given in Theorem IV.1.3 and Theorem IV.1.9, we know that there exists a biorthogonal family $(q_{\lambda,T})_{\lambda \in \Lambda}$ to the exponentials made upon the family Λ. It follows that, as we did in the finite dimensional setting, we may **formally** solve the moment problem above by defining

$$u(t) \stackrel{\text{def}}{=} \sum_{\mu \in \Lambda} u_\mu(t), \text{ with } u_\mu(t) \stackrel{\text{def}}{=} \frac{-\langle y_0, \phi_\mu \rangle_{L^2} e^{-\mu T}}{\partial_r \phi_\mu} q_{\mu,T}(t), \quad \forall \mu \in \Lambda.$$

Indeed, if this series makes sense (and if the following computation can be justified) we have for any $\lambda \in \Lambda$,

$$\int_0^T u(t)e^{-\lambda t}\,dt = \sum_{\mu\in\Lambda} \frac{-\langle y_0, \phi_\mu\rangle_{L^2} e^{-\mu T}}{\partial_r \phi_\mu} \underbrace{\int_0^T q_{\mu,T}(t)e^{-\lambda t}\,dt}_{=\delta_{\lambda,\mu}} = \frac{-\langle y_0, \phi_\lambda\rangle_{L^2} e^{-\lambda T}}{\partial_r \phi_\lambda},$$

and the claim will be proved. It remains to show the convergence of the series in $L^2(0,T)$. To this end, we will show that it is normally convergent. Indeed we have

$$\|u_\mu\|_{L^2(0,T)} \le \frac{\|y_0\|_{L^2} e^{-\mu T}}{|\partial_r \phi_\mu|} \|q_{\mu,T}\|_{L^2(0,T)}, \tag{IV.23}$$

and by the estimate given in Theorem IV.1.9, we deduce that

$$\begin{aligned}\|u_\mu\|_{L^2(0,T)} &\le K \frac{\|y_0\|_{L^2}}{|\partial_r \phi_\mu|} e^{-\mu T} e^{\varepsilon(\mu)\mu} \\ &\le C_{T,\Lambda,y_0} \frac{1}{|\partial_r \phi_\mu|} e^{-\mu(T-\varepsilon(\mu))}.\end{aligned}$$

Since $\lim_{+\infty} \varepsilon = 0$ we deduce that for, for some $\mu_0 > 0$, we have

$$\|u_\mu\|_{L^2(0,T)} \le C_{T,\Lambda,y_0} \frac{1}{|\partial_r \phi_\mu|} e^{-\mu T/2}, \quad \forall \mu \in \Lambda,\ \mu > \mu_0.$$

Finally, we use the bound from below for $|\partial_r \phi_\mu|$ given in Theorem IV.1.3, to deduce that

$$\|u_\mu\|_{L^2(0,T)} \le C \frac{e^{-\mu T/2}}{\sqrt{\mu}}, \quad \forall \mu \in \Lambda, \mu > \mu_0,$$

which proves, thanks to (A.15), that $\sum_{\mu\in\Lambda} \|u_\mu\|_{L^2(0,T)} < +\infty$ and concludes the proof. ■

We can use the same kind of proof in the case of the distributed control problem.

Theorem IV.1.12 (Distributed null-controllability in 1D)

Assume that $d = 1$, $\Omega = (0,1)$. Let ω be any non empty open subset of Ω. For any $y_0 \in L^2(\Omega)$, and $T > 0$, there exists a control $v \in L^2((0,T)\times\omega)$ such that the solution of (IV.8) satisfies $y(T) = 0$.

Proof :

We start with the same formulation as before, for any function $v \in L^2((0,T)\times\omega)$

$$\langle y(T), \phi_\lambda\rangle_{L^2} - \langle y_0, e^{-\lambda T}\phi_\lambda\rangle_{L^2} = \int_0^T \int_\omega v(t,x)e^{-(T-t)\lambda}\phi_\lambda(x)\,dx\,dt, \quad \forall \lambda \in \Lambda.$$

The solution vanishes at time T, if and only if the function $u(t,x) \stackrel{\text{def}}{=} v(T-t,x)$ satisfies the following space-time moment problem

$$\int_0^T \int_\omega u(t,x)e^{-\lambda t}\phi_\lambda(x)\,dx\,dt = -\langle y_0, \phi_\lambda\rangle_{L^2} e^{-\lambda T}, \ \forall \lambda \in \Lambda.$$

To solve this problem, we look for a biorthogonal family $(\tilde{q}_{\lambda,T})_{\lambda\in\Lambda}$ in $L^2((0,T)\times\omega)$ to the family of functions $\{(t,x) \in (0,T)\times\omega \mapsto \phi_\lambda(x)e^{-\lambda t}\}$. We propose the following family

$$\tilde{q}_{\lambda,T}(t,x) \stackrel{\text{def}}{=} \frac{\phi_\lambda(x)}{\|\phi_\lambda\|^2_{L^2(\omega)}} q_{\lambda,T}(t), \quad \forall (t,x) \in (0,T)\times\Omega, \quad \forall \lambda \in \Lambda,$$

and we indeed check, by the Fubini theorem, that for any $\lambda, \mu \in \Lambda$, we have

$$\int_0^T \int_\omega \tilde{q}_{\lambda,T}(t,x)\phi_\mu(x)e^{-\mu t}\,dt = \frac{1}{\|\phi_\lambda\|^2_{L^2(\omega)}}\left(\int_\omega \phi_\lambda\phi_\mu\,dx\right)\underbrace{\left(\int_0^T q_{\lambda,T}(t)e^{-\mu t}\,dt\right)}_{=\delta_{\lambda,\mu}} = \delta_{\lambda,\mu}.$$

Finally, we can define a formal null-control u by the series

$$u \stackrel{\text{def}}{=} \sum_{\mu\in\Lambda} u_\mu, \quad \text{with } u_\mu(t,x) \stackrel{\text{def}}{=} -\langle y_0, \phi_\mu\rangle_{L^2} e^{-\mu T}\tilde{q}_{\mu,T}(t,x).$$

It remains to check the convergence of this series by computing

$$\|u_\mu\|_{L^2((0,T)\times\omega)} \leq \|y_0\|_{L^2}e^{-\mu T}\|\tilde{q}_{\mu,T}\|_{L^2((0,T)\times\Omega)} \leq \|y_0\|_{L^2}e^{-\mu T}\frac{\|q_{\mu,T}\|_{L^2(0,T)}}{\|\phi_\mu\|_{L^2(\omega)}},$$

so that, for some $\mu_0 > 0$,

$$\|u_\mu\|_{L^2((0,T)\times\omega)} \leq K_{y_0,T,\Lambda}\frac{1}{\|\phi_\mu\|_{L^2(\omega)}}e^{-\mu T/2}, \quad \forall \mu \in \Lambda, \mu > \mu_0.$$

Using the bound from below for $\|\phi_\mu\|_{L^2(\omega)}$ in Theorem IV.1.3 and (A.15), we conclude the convergence in $L^2((0,T)\times\omega)$ of the series that defines u and the claim is proved. ■

IV.1.2 Biorthogonal family of exponentials

The main goal of this section is to prove Theorems IV.1.8 and IV.1.9. Let us consider a given family of positive numbers Λ that satisfies, for the moment, the summability condition (IV.19).

IV.1.2.1 Blaschke products

Let us define $\mathbb{C}^+$ to be the complex half-plane

$$\mathbb{C}^+ \stackrel{\text{def}}{=} \{z \in \mathbb{C}, \mathcal{R}e\, z > 0\}.$$

Proposition and Definition IV.1.13

Under assumption (IV.19), for any $L \subset \Lambda$, the following product

$$W_L(z) \stackrel{\text{def}}{=} \prod_{\mu\in L}\frac{\mu - z}{\mu + z}, \quad \forall z \in \mathbb{C}^+,$$

is well-defined and holomorphic on $\mathbb{C}^+$. Its zeros are exactly the points in L.
Moreover, we have the lower bound

$$|W_L(z)| \geq |W_L(\mathcal{R}e\, z)|, \quad \forall z \in \mathbb{C}^+. \tag{IV.24}$$

Proof :

In the case where L is finite, the existence and holomorphy are straightforward. Assume now that L is infinite and let us fix $M > 0$. We write

$$W_L(z) = W_L^-(z).W_L^+(z),$$

with

$$W_L^-(z) = \prod_{\substack{\mu\in L\\ \mu\leq 4M}}\frac{\mu - z}{\mu + z},$$

$$W_L^+(z) = \prod_{\substack{\mu\in L\\ \mu>4M}} \frac{\mu - z}{\mu+z}.$$

Since $L\cap[0,4M]$ is finite, the well-posedness and the properties of W_L^- are clear. Let us study the other factor W_L^+ on the open half-disk $D_M \stackrel{\text{def}}{=} \mathbb{C}^+\cap D(0,M)$.

For any $\mu > 4M$, and $z\in D_M$, we have

$$\left|1-\frac{\mu-z}{\mu+z}\right| = \left|\frac{2z}{\mu+z}\right| \leq \frac{2|z|}{\mu} < \frac{2M}{4M} = 1/2,$$

and thus, using that, for any $w\in\mathbb{C}$ such that $|w|<1/2$, we have

$$|\log(1+w)| \leq \frac{|w|}{1-|w|} \leq 2|w|,$$

we obtain that

$$\left|\log\left(\frac{\mu-z}{\mu+z}\right)\right| \leq \frac{4|z|}{\mu} \leq \frac{4M}{\mu}.$$

By using (IV.19) we get that the infinite product W_L^+ is uniformly convergent in D_M and has no zeros in D_M. The claim is proved.

Finally, to each $\mu\in L$ and $z\in\mathbb{C}^+$, since $\mu>0$ and $\mathcal{R}e\, z>0$, we have

$$\left|\frac{\mu-z}{\mu+z}\right|^2 = \frac{(\mu-\mathcal{R}e\, z)^2+(\mathcal{I}m\, z)^2}{(\mu+\mathcal{R}e\, z)^2+(\mathcal{I}m\, z)^2} \geq \frac{(\mu-\mathcal{R}e\, z)^2}{(\mu+\mathcal{R}e\, z)^2},$$

which implies the last claim. ■

In addition to the summability property (IV.19), we assume now that a weak gap condition is satisfied: there exists an integer $p\geq 1$ and a $\rho>0$ such that

$$\#\big(\Lambda\cap[\mu,\mu+\rho)\big) \leq p, \quad \forall \mu>0. \tag{IV.25}$$

For $p=1$, this is nothing but the usual gap condition (IV.20). The weaker assumption (IV.25) will be useful when we will tackle the boundary null-controllability issue for coupled parabolic systems in Section V.4.2.

With those assumptions, we can prove the following estimates on the Blaschke product W_L.

Proposition IV.1.14

Let $\rho>0$ and $p\geq 1$ an integer, we assume the weak gap condition (IV.25) and the summability condition (IV.19). Let R be a remainder function for Λ.

There exists a decreasing function $r\mapsto\varepsilon(r)$ such that $\lim_{r\to+\infty}\varepsilon(r)=0$, depending only on ρ, p and R such that, for any $L\subset\Lambda$, we have

$$\left|\frac{1}{W_L(x)}\right| \leq e^{\varepsilon(x)\,x}, \quad \forall x>0, \text{ such that } d(x,L)>\rho/2.$$

In particular, for any $\lambda\in\Lambda$, we have

$$\left|\frac{1}{W_{\Lambda\setminus\{\lambda\}}(\lambda)}\right| \leq e^{\varepsilon(\lambda)\,\lambda}.$$

Proof :

We fix $x>0$ such that $d(x,L)\geq\rho/2$. Our goal is to bound the quantity

$$\frac{1}{W_L(x)} = \prod_{\sigma\in L}\left|\frac{1+\frac{x}{\sigma}}{1-\frac{x}{\sigma}}\right|.$$

We introduce the following convergent products

$$Q(x) \stackrel{\text{def}}{=} \prod_{\sigma\in L}\left(1+\frac{x}{\sigma}\right),$$

$$D(x) \stackrel{\text{def}}{=} \prod_{\sigma\in L}\left|1-\frac{x}{\sigma}\right|,$$

and we study separately those two expressions.

1. **Bound from above for Q.**

   Let us fix some value of $0 < x_0 < x$ that will be determined later and we write Q as the product of Q_1 and Q_2 defined by

$$Q_1(x) \stackrel{\text{def}}{=} \prod_{\substack{\sigma\in L\\ \sigma<x_0}}\left(1+\frac{x}{\sigma}\right), \quad \text{and} \quad Q_2(x) \stackrel{\text{def}}{=} \prod_{\substack{\sigma\in L\\ \sigma\geq x_0}}\left(1+\frac{x}{\sigma}\right).$$

   - In Q_1 we have $\sigma < x_0 < x$ so that $1+\frac{x}{\sigma} \leq \frac{2x}{\sigma}$ and it follows from Proposition A.6.18 that

$$\begin{aligned}\log Q_1(x) &\leq \sum_{\substack{\sigma\in L\\ \sigma<x_0}}\log\left(\frac{2x}{\sigma}\right)\\ &\leq N(x_0)\log\left(\frac{2x}{\inf\Lambda}\right)\\ &\leq N(x_0)\log\left(2R(0)x\right)\\ &\leq R(0)x_0\log\left(2R(0)x\right)\end{aligned}$$

     where we have used (A.13) to get $N(x_0)/x_0 \leq R(0)$ and (A.14).
   - In Q_2 we use the bound $1+\frac{x}{\sigma} \leq e^{x/\sigma}$ to obtain

$$\log Q_2(x) \leq \sum_{\substack{\sigma\in L\\ \sigma\geq x_0}}\frac{x}{\sigma} \leq \sum_{\substack{\sigma\in\Lambda\\ \sigma>x_0/2}}\frac{x}{\sigma} \leq xR(x_0/2).$$

   Finally, we have proved that

$$\log Q(x) \leq R(0)x_0\log\left(2R(0)x\right) + xR(x_0/2).$$

   Choosing $x_0 = x/(\log x)^2$, we eventually get

$$\log Q(x) \leq x\left[R(0)\frac{\log(2R(0)x)}{(\log x)^2} + R\left(\frac{x}{2(\log x)^2}\right)\right]. \tag{IV.26}$$

2. **Bound from below for D.**

   We write D as a product of four terms

$$D_1(x) \stackrel{\text{def}}{=} \prod_{\substack{\sigma\in L\\ \sigma<x/2}}\left|1-\frac{x}{\sigma}\right|,$$

$$D_2(x) \stackrel{\text{def}}{=} \prod_{\substack{\sigma\in L\\ x/2\leq\sigma<x}}\left|1-\frac{x}{\sigma}\right|,$$

$$D_3(x) \stackrel{\text{def}}{=} \prod_{\substack{\sigma\in L\\ x\leq\sigma<2x}}\left|1-\frac{x}{\sigma}\right|,$$

$$D_4(x) \stackrel{\text{def}}{=} \prod_{\substack{\sigma\in L\\ 2x\leq\sigma}}\left|1-\frac{x}{\sigma}\right|.$$

- All the factors in D_1 are larger than 1 so that $D_1(x) \geq 1$.
- In the term D_2 we notice that we necessarily have $\sigma < x - \rho/2$ since $d(x, L) > \rho/2$. Therefore, we can write

$$\begin{aligned}\log D_2(x) &= \sum_{\substack{\sigma\in L\\ x/2\leq\sigma<x-\rho/2}} \log\left(\frac{x-\sigma}{\sigma}\right),\\ &\geq \sum_{\substack{\sigma\in L\\ x/2\leq\sigma<x-\rho/2}} \log\left(\frac{x-\sigma}{x}\right),\\ &=\log(\rho/2x)N_L(x-\rho/2) - \log(1/2)N_L(x/2) + \int_{x/2}^{x-\rho/2} \frac{1}{x-t} N_L(t)\,dt\\ &=\log(\rho/x)N_L(x-\rho) - \log(2)N_L(x-\rho/2) + \log(2)N(x/2) + \int_{\rho/2}^{x/2} \frac{1}{u} N_L(x-u)\,dt\\ &=(\log 2)\left[N_L(x/2) - N_L(x-\rho/2)\right] - \int_{\rho/2}^{x/2} \frac{N_L(x-\rho/2) - N_L(x-u)}{u}\,du.\end{aligned}$$

Let us remark that, since $L \subset \Lambda$, we have for any $s < r$

$$\begin{aligned}N_L(r) - N_L(s) &= \#\big(L \cap (s, r]\big)\\ &\leq \#\big(\Lambda \cap (s, r]\big)\\ &= N_\Lambda(r) - N_\Lambda(s).\end{aligned} \tag{IV.27}$$

We can then use the following two estimates on $N = N_\Lambda$ for $u \in [\rho/2, x/2]$.

- The first one comes from Proposition A.6.20

$$|N(x-u) - N(x-\rho/2)| \leq (x-\rho/2)R(x-u) \leq xR(x/2).$$

- The second one comes from the weak gap assumption (IV.25) that gives

$$\begin{aligned}|N(x-u) - N(x-\rho/2)| &\leq \#\big(\Lambda \cap [x-u, x-\rho/2]\big)\\ &\leq p\left(\frac{u-\rho/2}{\rho} + 1\right)\\ &\leq p\left(\frac{u}{\rho} + \frac{1}{2}\right)\\ &\leq \frac{2p}{\rho}u.\end{aligned}$$

  Indeed, $[x-u, x-\rho/2]$ can be split into $\lfloor (u-\rho/2)/\rho \rfloor + 1$ disjoint intervals of length less than ρ, each of them containing at most p elements of Λ.

Let us now combine the previous two estimates as follows

$$|N(x-u) - N(x-\rho/2)| \leq \sqrt{\frac{2p}{\rho}R(x/2)}\sqrt{u}\sqrt{x},$$

so that

$$\begin{aligned}\int_{\rho/2}^{x/2} \frac{N(x-\rho/2) - N(x-u)}{u}\,du &\leq \sqrt{\frac{2p}{\rho}R(x/2)}\sqrt{x}\int_{\rho/2}^{x/2} \frac{1}{\sqrt{u}}\,du\\ &\leq 2\sqrt{\frac{2p}{\rho}R(x/2)}\sqrt{x}\sqrt{x/2}\\ &\leq C_{p,\rho}\,x\sqrt{R(x/2)}.\end{aligned}$$

As a conclusion, we have proved that

$$\log D_2(x) \geq -(\log 2)xR(x/2) - C_{p,\rho}x\sqrt{R(x/2)}.$$

- The term D_3 is treated in a similar way as D_2:

$$\begin{aligned}
\log D_3(x) &= \sum_{\substack{\sigma\in L\\ x+\rho/2\leq\sigma<2x}} \log\left(\frac{\sigma-x}{\sigma}\right),\\
&\geq \sum_{\substack{\sigma\in L\\ x+\rho/2\leq\sigma<2x}} \log\left(\frac{\sigma-x}{2x}\right),\\
&= \log(1/2)N_L(2x) - \log(\rho/2x)N_L(x+\rho/2) - \int_{x+\rho/2}^{2x} \frac{1}{t-x}N_L(t)\,dt\\
&= -\log 2(N_L(2x) - N_L(x+\rho/2)) - \int_{\rho/2}^{x} \frac{N_L(x+u)-N_L(x+\rho/2)}{u}\,du,
\end{aligned}$$

we conclude by using (IV.27) and by combining the following two inequalities

$$|N(x+u) - N(x+\rho/2)| \leq 2xR(x+\rho/2) \leq 2xR(x),$$

$$|N(x+u) - N(x+\rho/2)| \leq \frac{2p}{\rho}u,$$

as we did for D_2.

- For the term D_4 we use that

$$1-u \geq e^{-2u}, \quad \forall u\in[0,1/2],$$

so that

$$\log D_4(x) \geq -2x\sum_{\substack{\sigma\in L\\ 2x\leq\sigma}} \frac{1}{\sigma} \geq -2x\sum_{\substack{\sigma\in\Lambda\\ 2x\leq\sigma}} \frac{1}{\sigma} \geq -2xR(2x).$$

All in all we have obtained

$$\begin{aligned}
\log\frac{1}{W(x)} &\leq \log Q(x) - \log D(x)\\
&\leq C_{p,\rho}\left[R(0)\frac{\log(2R(0)x)}{(\log x)^2} + R\left(\frac{x}{2(\log x)^2}\right) + R(x/2) + \sqrt{R(x/2)}\right]x,
\end{aligned}$$

which is the expected estimate if we define $\varepsilon(x)$ to be factor in front of x in the right-hand side. By the properties of R we have indeed that this ε is non increasing and that $\lim_{x\to\infty}\varepsilon(x)=0$. ■

Remark IV.1.15

Note that the estimates on the terms D_2 and D_3 really need the gap condition. For instance, the following basic estimate

$$\begin{aligned}
\log D_2(x) &\geq \sum_{\substack{\sigma\in L\\ x/2\leq\sigma<x-\rho/2}} \log\left(\frac{x-\sigma}{x}\right),\\
&\geq -(N(x)-N(x/2))\log(2x/\rho)\\
&\geq -xR(x/2)\log(2x/\rho),
\end{aligned}$$

is not sufficient since it may happen that $x\mapsto R(x/2)\log(x)$ does not tend to 0 at infinity.

Corollary IV.1.16

Let $\rho > 0$ and $p \geq 1$ an integer, we assume the weak gap condition (IV.25) and the summability condition (IV.19). Let R be a remainder function for Λ.
For any $k \geq 0$, there exists a decreasing function $r \mapsto \varepsilon(r)$ such that $\lim_{r\to+\infty} \varepsilon(r) = 0$, depending only on k, ρ, p and R such that, for any $L \subset \Lambda$, such that we have

$$\left| \left(\frac{1}{W_L}\right)^{(k)}(x) \right| \leq e^{\varepsilon(x)\, x}, \quad \forall x > 0, \textit{such that } \ d(x, L) > 3\rho/4.$$

Proof :

Let $x > 0$ such that $d(x, L) > 3\rho/4$. We also assume that $x > \rho/4$ (if not the result is straightforward). Let Γ_x be the circle in the complex plane centered at x and of radius $\rho/4$. By the Cauchy formula we have

$$\left(\frac{1}{W_L}\right)^{(k)}(x) = \frac{k!}{2i\pi} \int_{\Gamma_x} \frac{(1/W_L)(z)}{(z-x)^{k+1}}\, dz,$$

so that

$$\left| \left(\frac{1}{W_L}\right)^{(k)}(x) \right| \leq \frac{C_k}{\rho^{k+1}} \sup_{z\in\Gamma_x} \left| \frac{1}{W_L(z)} \right|.$$

We can use (IV.24) to deduce

$$\left| \left(\frac{1}{W_L}\right)^{(k)}(x) \right| \leq \frac{C_k}{\rho^{k+1}} \sup_{y\in[x-\rho/4, x+\rho/4]} \left| \frac{1}{W_L(y)} \right|.$$

Since any $y \in [x - \rho/4, x + \rho/4]$ satisfies $d(y, L) \geq \rho/2$, we can use Proposition IV.1.14 to obtain

$$\left| \left(\frac{1}{W_L}\right)^{(k)}(x) \right| \leq \frac{C_k}{\rho^{k+1}} e^{\varepsilon(y)y} \leq \frac{C_k}{\rho^{k+1}} e^{\varepsilon(x-\rho/4)(x+\rho/4)},$$

which concludes the proof of the claim, by changing the definition of the function ε. ∎

IV.1.2.2 Biorthogonal functions in $L^2(0, +\infty)$. Proof of Theorem IV.1.8

In this subsection we assume that Λ satisfies (IV.19) and (IV.20).

We start by observing that for any $\lambda, \mu \in \Lambda$ we have

$$(e[\lambda], e[\mu])_{L^2(0,+\infty)} = \frac{1}{\lambda + \mu}. \tag{IV.28}$$

For any subset L of Λ, we introduce the family $\mathcal{E}_L \stackrel{\text{def}}{=} \{e[\mu], \mu \in L\}$ in $L^2(0, +\infty)$. As defined in Section A.4, we introduce $\pi_{\mathcal{E}_L}$ the orthogonal projection in $L^2(0, +\infty)$ onto $\overline{\text{Span}\,\mathcal{E}_L}$.

For any **finite** subset L of Λ, we see by (IV.28) that the Gram matrix G_L of the family $\mathcal{E}_L$ in $L^2(0, +\infty)$ is just the Cauchy matrix

$$G_L \stackrel{\text{def}}{=} \left(\frac{1}{\lambda + \mu}\right)_{\substack{\lambda\in L \\ \mu\in L}},$$

whose determinant is explicitly computable (see Proposition A.4.14) as follows

$$\Delta_L = \left(\prod_{\lambda\in L} \frac{1}{2\lambda}\right) \prod_{\substack{\lambda,\mu\in L \\ \lambda<\mu}} \left(\frac{\lambda-\mu}{\lambda+\mu}\right)^2.$$

By usual results on Gram determinants (see Proposition A.4.9) we have that for any $\sigma \in \Lambda$ and any finite $L \subset \Lambda$ with $\sigma \notin L$,

$$\delta(e[\sigma], \mathcal{E}_L)^2 = \frac{\Delta_{\sigma \cup L}}{\Delta_L},$$

and we finally obtain the explicit formula

$$\delta(e[\sigma], \mathcal{E}_L)^2 = \frac{1}{2\sigma} \prod_{\mu \in L} \left(\frac{\sigma - \mu}{\sigma + \mu} \right)^2,$$

that can be written as follows using the Blaschke product W_L, introduced in the Definition and Proposition IV.1.13,

$$\delta(e[\sigma], \mathcal{E}_L) = \frac{1}{\sqrt{2\sigma}} |W_L(\sigma)| . \tag{IV.29}$$

A priori, this formula is only valid for a finite subset L of Λ. However, by Lemma A.4.7 we know that

$$\delta(\sigma, \mathcal{E}_{\Lambda \setminus \{\sigma\}}) = \lim_{n \to \infty} \delta(\sigma, \mathcal{E}_{\Lambda_n \setminus \{\sigma\}}),$$

where, for instance, we have chosen $\Lambda_n \stackrel{\text{def}}{=} \Lambda \cap [0, n]$. By (IV.29) and the uniform convergence property of the infinite product we get

$$\delta(\sigma, \mathcal{E}_{\Lambda \setminus \{\sigma\}}) = \frac{1}{\sqrt{2\sigma}} \lim_{n \to \infty} \left| W_{\Lambda_n \setminus \{\sigma\}}(\sigma) \right| = \frac{1}{2\sigma} \left| W_{\Lambda \setminus \{\sigma\}}(\sigma) \right| > 0.$$

Since this is true for any $\sigma \in \Lambda$, we deduce by Proposition A.4.12 that there exists a family $(q_{\sigma,\infty})_{\sigma \in \Lambda}$ in $L^2(0, +\infty)$ which is biorthogonal to $\mathcal{E}_\Lambda$, and it satisfies

$$\|q_{\sigma,\infty}\|_{L^2(0,+\infty)} = \frac{1}{\delta(\sigma, \mathcal{E}_{\Lambda \setminus \{\sigma\}})} = \sqrt{2\sigma} \left| \frac{1}{W_{\Lambda \setminus \{\sigma\}}(\sigma)} \right| .$$

The proof of Theorem IV.1.8 is thus complete thanks to the estimate given in Proposition IV.1.14, with $p = 1$ in the present case.

IV.1.2.3 Biorthogonal family on $(0, T)$. Proof of Theorem IV.1.9

Let us introduce the linear space spanned by all the exponential functions corresponding to Λ

$$E_\Lambda \stackrel{\text{def}}{=} \text{Span}(\mathcal{E}_\Lambda),$$

and the closures of this space in $L^2(0, \infty)$ and $L^2(0, T)$, $T > 0$, respectively

$$F_{\Lambda,\infty} \stackrel{\text{def}}{=} \overline{E_\Lambda}^{L^2(0,\infty)}, \quad F_{\Lambda,T} \stackrel{\text{def}}{=} \overline{E_\Lambda}^{L^2(0,T)}, \quad \forall T > 0.$$

Restriction operator. For a given $T > 0$ we define $\Gamma_{\Lambda,T}$ to be the restriction operator

$$\Gamma_{\Lambda,T} : f \in F_{\Lambda,\infty} \mapsto f_{|[0,T]} \in F_{\Lambda,T},$$

which is of course linear, continuous and onto. We shall now prove that this operator is invertible. More precisely we have:

Proposition IV.1.17

Assume that Λ satisfies (IV.19) and (IV.20) and let R be a remainder function for Λ. For any $T > 0$, there exists a $C_{R,\rho,T} > 0$, depending only on T, R and ρ, such that

$$\|f\|_{L^2(0,+\infty)} \leq C_{R,\rho,T} \|\Gamma_{\Lambda,T} f\|_{L^2(0,T)}, \quad \forall f \in F_{\Lambda,\infty}. \tag{IV.30}$$

Proof :

By density, is is enough to prove (IV.30) for $f \in E_\Lambda$. We will use a contradiction argument.

Let us fix a $T > 0$ and assume that this inequality is false: then there exists a sequence $(\Lambda^n)_n$ of subsets of $(0,+\infty)$ each of them satisfying the summability condition (IV.19) with remainder function R, the gap condition (IV.20), and a sequence of functions $f_n \in E_{\Lambda^n}$ such that

$$\|f_n\|_{L^2(0,+\infty)} = 1, \text{ and } \|\Gamma_{\Lambda^n,T} f_n\|_{L^2(0,T)} \leq 1/n.$$

Each f_n can be written

$$f_n(t) = \sum_{\lambda\in\Lambda^n} a_\lambda^n e_t[\lambda], \tag{IV.31}$$

where $a_\lambda^n \neq 0$ only for finitely many values of λ. From Theorem IV.1.8, we know that, for each n there exists a biorthogonal family $(q_{\lambda,\infty}^n)_{\lambda\in\Lambda^n}$ to E_{Λ^n} in $L^2(0,\infty)$ that satisfies

$$\|q_{\lambda,\infty}^n\|_{L^2(0,+\infty)} \leq e^{\varepsilon(\lambda)\lambda}, \quad \forall \lambda \in \Lambda^n,$$

where ε is a non increasing function tending to 0 at infinity which does depend on n since all the Λ^n share the same gap ρ and the same remainder function R.

Taking the inner product of (IV.31) by some $q_{\lambda,\infty}$ and using the biorthogonality property, we have, for any n and any $\lambda \in \Lambda^n$

$$a_\lambda^n = (f_n, q_{\lambda,\infty}^n)_{L^2(0,\infty)}.$$

From the Cauchy-Schwarz inequality and the bounds above, we deduce that

$$|a_\lambda^n| \leq e^{\varepsilon(\lambda)\lambda}, \quad \forall \lambda \in \Lambda^n, \forall n \geq 1.$$

Let us now introduce the holomorphic extension of f_n

$$f_n(z) \stackrel{\text{def}}{=} \sum_{\lambda\in\Lambda^n} a_\lambda^n e_z[\lambda], \quad \forall z \in \mathbb{C}.$$

By the estimate of a_λ^n above, we deduce

$$|f_n(z)| \leq \sum_{\lambda\in\Lambda^n} e^{\varepsilon(\lambda)\lambda} e^{-\lambda \mathcal{R}e\, z} = \sum_{\lambda\in\Lambda^n} e^{-\lambda(\mathcal{R}e\, z - \varepsilon(\lambda))}. \tag{IV.32}$$

Using (A.15), we get that the sequence $(f_n)_n$ is bounded on every half-plane $\mathbb{C}_\eta^+ = \{z \in \mathbb{C}, \mathcal{R}e\, z > \eta\}$, with $\eta > 0$. By Montel's theorem, we deduce that $(f_n)_n$ converges locally uniformly in $\mathbb{C}^+$ towards an holomorphic function f.

By construction of $(f_n)_n$ we also have $\|\Gamma_{\Lambda^n,T} f_n\|_{L^2(0,T)} \to 0$ when $n \to \infty$ which implies that $f = 0$ on $(0,T)$. Since f is holomophic in $\mathbb{C}^+$, we deduce by the isolated zeros principle that $f = 0$ everywhere in $\mathbb{C}^+$.

As a consequence, for any $S > T$, we have

$$\int_0^S |f_n(t)|^2\, dt \xrightarrow[n\to\infty]{} \int_0^S |f(t)|^2\, dt = 0.$$

We choose now

$$S \stackrel{\text{def}}{=} 2\varepsilon\left(\frac{1}{R(0)}\right).$$

By (A.14) and since ε is non increasing, we have

$$\varepsilon(\lambda) \leq \frac{S}{2}, \quad \forall \lambda \in \Lambda^n, \forall n \geq 1.$$

Therefore, with this choice of S, we deduce from (IV.32) that, for any $t > S$ and any $n \geq 1$,

$$|f_n(t)| \leq \sum_{\lambda\in\Lambda^n} e^{-\lambda t/2},$$

and thus, using (A.15), we get

$$\begin{aligned}\int_S^{+\infty} |f_n(t)|\,dt &\leq \sum_{\lambda\in\Lambda^n} \frac{2}{\lambda} e^{-\lambda S/2}\\ &\leq 2R(0) \sum_{\lambda\in\Lambda^n} e^{-\lambda S/2}\\ &\leq \frac{4(R(0))^2}{S}.\end{aligned}$$

It follows that

$$\int_S^{+\infty} |f_n(t)|\,dt \xrightarrow[S\to\infty]{} 0,$$

uniformly in n. Since $(f_n)_n$ is uniformly bounded on $[S,+\infty[$ this implies

$$\int_S^{+\infty} |f_n(t)|^2\,dt \leq C\int_S^{+\infty} |f_n(t)|\,dt \xrightarrow[S\to\infty]{} 0,$$

uniformly in n. All in all, we have finally proved that $\|f_n\|_{L^2(0,+\infty)} \to 0$ which is a contradiction with the initial assumption that $\|f_n\|_{L^2(0,+\infty)} = 1$. The claim is proved. ■

Conclusion. For any $\lambda\in\Lambda$, we set

$$q_{\lambda,T} \stackrel{\text{def}}{=} (\Gamma_{\Lambda,T}^{-1})^* q_{\lambda,\infty}, \tag{IV.33}$$

where $(q_{\lambda,\infty})_{\lambda\in\Lambda}$ is the biorthogonal family to $\mathcal{E}_\Lambda$ in $L^2(0,+\infty)$ given by Theorem IV.1.8. Notice that, by construction, we have $q_{\lambda,\infty}\in F_{\Lambda,\infty}$, so that formula (IV.33) makes sense.

We can now check that this family $(q_{\lambda,T})_{\lambda\in\Lambda}$ satisfies the required properties

- For any $\lambda,\mu\in\Lambda$, we have

$$(q_{\lambda,T}, e[\mu])_{L^2(0,T)} = ((\Gamma_{\Lambda,T}^{-1})^* q_{\lambda,\infty}, \Gamma_{\Lambda,T} e[\mu])_{L^2(0,T)} = (q_{\lambda,\infty}, (\Gamma_{\Lambda,T})^{-1}\Gamma_{\Lambda,T} e[\mu])_{L^2(0,+\infty)} = \delta_{\lambda,\mu}.$$

- For any $\lambda\in\Lambda$, we can use Proposition IV.1.18 to get

$$\|q_{\lambda,T}\|_{L^2(0,T)} \leq \|(\Gamma_{\Lambda,T}^{-1})^*\|\|q_{\lambda,\infty}\|_{L^2(0,+\infty)} = \|\Gamma_{\Lambda,T}^{-1}\|\|q_{\lambda,\infty}\|_{L^2(0,+\infty)} \leq C_{R,\rho,T}\|q_{\lambda,\infty}\|_{L^2(0,+\infty)},$$

  and thus, the bounds on $(q_{\lambda,\infty})_{\lambda\in\Lambda}$ are transferred to $(q_{\lambda,T})_{\lambda\in\Lambda}$ with the additional constant $C_{R,\rho,T}$ in front of the exponential. The claim is proved.

IV.1.2.4 Sharper estimates of the biorthogonal family on $(0,T)$.

In the case where we assume a suitable asymptotic behavior on the counting function of Λ, we can obtain an explicit estimate of the norm of the restriction operator $R_{T,\Lambda}$ as a function of T and then an explicit estimate of the norm of the biorthogonal family with respect to T and λ. In section IV.1.2.5 we will give, as an application, an estimate of the control cost for our parabolic PDE as a function of the control time.

More precisely, let us assume that, in addition to the gap condition (IV.20), the counting function N associated with Λ satisfies (IV.21) for some $\bar N$ and some $\alpha\in(0,1)$. This implies, of course the condition (IV.19). All the constants C_i in the statements and proofs of this section will only depend on the parameters $\bar N$ and α.

The remainder of this section will be devoted to the proof of the following result which is a refinement of Proposition IV.1.17.

Theorem IV.1.18

Under the above assumptions, there exists $C_1>0$ such that

$$\|f\|_{L^2(0,+\infty)} \leq C_1 e^{C_1 T^{-\frac{\alpha}{1-\alpha}}} \|\Gamma_{\Lambda,T} f\|_{L^2(0,T)}, \quad \forall f\in F_{\Lambda,\infty}.$$

The proof makes use of real and complex analysis tools. Our first goal will be to construct an entire function satisfying the following properties.

Proposition IV.1.19

For any $\tau > 0$, there exists an entire function $G_{\Lambda,\tau}$ satisfying:

1. *$G_{\Lambda,T}$ is of exponential type τ,*
2. *$G_{\Lambda,T}(0) = 1$,*
3. *$G_{\Lambda,T}(i\lambda) = 0$ for any $\lambda \in \Lambda$,*
4. *$G_{\Lambda,T}$ is square integrable on the real axis and satisfies*

$$\|G_{\Lambda,T}\|_{L^2(\mathbb{R})} \leq C_2 e^{C_2 \tau^{-\frac{\alpha}{1-\alpha}}}.$$

Proof :

This function will be found as the product of two functions:

- an entire function F_Λ, depending only on Λ, that cancels on $i\Lambda$ but which does not satisfy the expected estimates.
- a mutiplier function M_τ depending only on τ, α and $\bar{N}$ () that will let us produce a suitable bound on $G_{\Lambda,T}$.

• Step 1: An infinite product cancelling on $i\Lambda$.

Let us define

$$F_\Lambda(z) \stackrel{\text{def}}{=} \prod_{\lambda\in\Lambda} \left(1 + \frac{iz}{\lambda}\right), \quad z \in \mathbb{C},$$

which is well-defined and holomorphic on $\mathbb{C}$ since the series $\sum \frac{1}{\lambda}$ is convergent.

Lemma IV.1.20

The following estimate holds

$$|F_\Lambda(z)| \leq e^{C_3|z|^\alpha},$$

with

$$C_3 \stackrel{\text{def}}{=} \bar{N} \int_0^\infty \frac{dr}{(1+r)r^{1-\alpha}}.$$

Proof :

We will estimate the logarithm of $|F_\Lambda(z)|$ as follows

$$\log |F_\Lambda(z)| \leq \sum_{\lambda\in\Lambda} \log\left(1 + \frac{|z|}{\lambda}\right).$$

Using the summability formulas in Proposition A.6.18, we get

$$\begin{aligned}\log |F_\Lambda(z)| &\leq \int_0^{+\infty} \frac{|z|}{r(|z|+r)} N(r)\, dr \\ &= \int_0^{+\infty} \frac{1}{r(1+r)} N(r|z|)\, dr \\ &\leq \bar{N}|z|^\alpha \int_0^{+\infty} \frac{r^\alpha}{r(1+r)}\, dr\end{aligned}$$

■

- Step 2 : A suitable multiplier.

  The function F_Λ constructed above is clearly not square integrable on the real line. We will thus now construct a multiplier function M_τ, depending only on τ and of the parameters $\alpha, \bar{N}$ of the counting function N, whose role is to make the product $F_\Lambda M_\tau$ square integrable on the real line, still controlling its exponential type.

  To this end, we set
$$C_4 \stackrel{\text{def}}{=} 4\alpha C_3,$$
  and we consider a $\tau > 0$ such that
$$\tau < \frac{C_4^{\frac{1}{\alpha}}}{1-\alpha}. \tag{IV.34}$$
  Let $\mathsf{L} \subset (0, +\infty)$ be the following family
$$\mathsf{L} \stackrel{\text{def}}{=} \left\{ r_0 + \left(\frac{n}{C_4}\right)^{\frac{1}{\alpha}}, \; n \geq 1 \right\},$$
  with
$$r_0 \stackrel{\text{def}}{=} \left(\frac{(1-\alpha)\tau}{C_4}\right)^{-\frac{1}{1-\alpha}} - C_4^{-\frac{1}{\alpha}}.$$
  The assumption (IV.34) implies that $r_0 > 0$ and that
$$\inf \mathsf{L} = \left(\frac{(1-\alpha)\tau}{C_4}\right)^{-\frac{1}{1-\alpha}}. \tag{IV.35}$$
  It is very easy to prove that the counting function N_L associated with L satisfies
$$C_4(r-r_0)^\alpha - 1 \leq N_\mathsf{L}(r) \leq C_4 r^\alpha, \quad \forall r \geq 0,$$
  and of course
$$N_\mathsf{L}(r) = 0, \quad r \leq \inf \mathsf{L}.$$

Lemma IV.1.21

We have the property
$$\sum_{l \in \mathsf{L}} \frac{1}{l} \leq \tau.$$

Proof :

We apply the summation formulas of Proposition A.6.18
$$\begin{aligned}
\sum_{l \in \mathsf{L}} \frac{1}{l} &= \int_0^\infty \frac{1}{r^2} N_\mathsf{L}(r)\, dr \\
&= \int_{\inf \mathsf{L}}^\infty \frac{1}{r^2} N_\mathsf{L}(r)\, dr \\
&\leq C_4 \int_{\inf \mathsf{L}}^\infty \frac{1}{r^{2-\alpha}}\, dr \\
&= \frac{C_4}{1-\alpha} (\inf \mathsf{L})^{\alpha-1} \\
&= \tau.
\end{aligned}$$

■

We can now introduce the following multiplier

$$M_\tau(z) \stackrel{\text{def}}{=} \prod_{l\in\mathsf{L}} \frac{\sin(z/l)}{z/l}.$$

Lemma IV.1.22

- *There exists $C_5 > 0$ such that*

$$|M_\tau(z)| \leq C_5 e^{\tau|z|}, \quad \forall z \in \mathbb{C}.$$

- *There exists $C_6 > 0$ such that*

$$|M_\tau(x)| \leq e^{-\frac{C_4}{2\alpha}|x|^\alpha + C_6\tau^{-\frac{\alpha}{1-\alpha}}}, \quad \forall x \in \mathbb{R}.$$

Proof :

- We note that, for any complex number z, we have

$$\left|\frac{\sin z}{z}\right| = \left|\sum_{k\geq 0}(-1)^k \frac{z^{2k}}{(2k+1)!}\right| \leq \sum_{k\geq 0} \frac{|z|^{2k}}{(2k+1)!} \leq \sum_{k\geq 0} \frac{|z|^{2k}}{(2k)!} \leq e^{|z|}.$$

Thus, we clearly have

$$|M_\tau(z)| \leq e^{\left(\sum_{l\in\mathsf{L}}\frac{1}{l}\right)|z|} \leq e^{\tau|z|},$$

by Lemma IV.1.21.

- We simply write for any $x \geq 0$

$$|M_\tau(x)| \leq \prod_{l\in\mathsf{L}} \left|\frac{\sin(x/l)}{x/l}\right|,$$

and we use that the sinc function is less than 1 to obtain

$$|M_\tau(x)| \leq \prod_{\substack{l\in\mathsf{L}\\ l\geq x}} \frac{l}{x}.$$

Taking the logarithm, it follows

$$\begin{aligned}
\log|M_\tau(x)| &\leq \sum_{\substack{l\in\mathsf{L}\\ l\geq x}} \log\left(\frac{l}{x}\right)\\
&= -\int_0^x \frac{N_{\mathsf{L}}(r)}{r}\,dr\\
&= -\int_{\inf\mathsf{L}}^x \frac{N_{\mathsf{L}}(r)}{r}\,dr\\
&\leq \int_{\inf\mathsf{L}}^x \frac{1-C_4(r-r_0)^\alpha}{r}\,dr\\
&= \log(x/\inf\mathsf{L}) - C_4\int_{\inf\mathsf{L}}^x \left(\frac{1}{(r-r_0)^{1-\alpha}} - \frac{r_0}{r(r-r_0)^{1-\alpha}}\right)dr\\
&\leq \log(x/\inf\mathsf{L}) - \frac{C_4}{\alpha}\Big((x-r_0)^\alpha - (\inf\mathsf{L}-r_0)^\alpha\Big) + C_4\int_{r_0}^\infty \frac{r_0}{r(r-r_0)^{1-\alpha}}\,dr\\
&\leq \log(x/\inf\mathsf{L}) - \frac{C_4}{\alpha}\Big((x-r_0)^\alpha - (\inf\mathsf{L}-r_0)^\alpha\Big) + C_4 r_0^\alpha\int_1^\infty \frac{1}{r(r-1)^{1-\alpha}}\,dr.
\end{aligned}$$

Using that $(\inf \mathsf{L} - r_0)^\alpha = \frac{1}{C_4}$, the sublinearity of the function $r \mapsto r^\alpha$ and (IV.35), we deduce that

$$\log |M_\tau(x)| \leq \log x - \frac{C_4}{\alpha} x^\alpha + \frac{1}{1-\alpha}\left(\log \frac{1-\alpha}{C_4} + \log \tau\right) + \frac{1}{\alpha} + \frac{C_4}{\alpha} r_0^\alpha + C_4 r_0^\alpha \int_1^\infty \frac{1}{r(r-1)^{1-\alpha}}\, dr.$$

Since $r_0 \leq \inf \mathsf{L}$ and using (IV.35), we obtain that for some $C_7 > 0$ (depending only on α, and $\bar{N}$), we have

$$\log |M_\tau(x)| \leq \log x - \frac{C_4}{\alpha} x^\alpha + \frac{\log \tau}{1-\alpha} + C_7\left(1 + \tau^{-\frac{\alpha}{1-\alpha}}\right).$$

The claim comes the comparison between the logarithm and the power functions $x \mapsto x^\alpha$ and $\tau \mapsto \tau^{\frac{\alpha}{1-\alpha}}$.

■

- Step 3 : The final construction.

  We can now consider the product function

  $$G_{\Lambda,\tau}(z) \stackrel{\text{def}}{=} F_\Lambda(z) M_\tau(z), \quad \forall z \in \mathbb{C}.$$

  Using the above estimates on F_Λ and M_τ, we know that $G_{\Lambda,\tau}$ is entire, of exponential type τ, and that it cancels at each $z = i\lambda$, $\lambda \in \Lambda$ and finally satisfies $G_{\Lambda,\tau}(0) = 1$.

  Moreover, for any $x \in \mathbb{R}$, we have

  $$|G_{\Lambda,\tau}(x)| \leq e^{-C_3|x|^\alpha + C_6 \tau^{-\frac{\alpha}{1-\alpha}}},$$

  so that, for some $C_2 > 0$ we have and so $\|G_{\Lambda,\tau}\|_{L^2} \leq C_2 e^{C_2 \tau^{-\frac{\alpha}{1-\alpha}}}$.

■

Estimates on sums of real exponentials and on generalized Müntz polynomials.

Proposition IV.1.23

There exists $C_8 > 0$, such that for any $\tau > 0$ and any function f in $\mathcal{E}_{\Lambda\cup\{0\}} = \mathrm{Span}(e[0], e[\lambda], \lambda \in \Lambda)$ that we write

$$f = a_0 + \sum_{\lambda \in \Lambda} a_\lambda e[\lambda],$$

we have the estimate

$$\lim_{t \to +\infty} |f(t)| = |a_0| \leq C_8 e^{C_8 \tau^{-\frac{\alpha}{1-\alpha}}} \|a\|_{L^2(0,2\tau)}.$$

Proof :

Applying the Paley-Wiener theorem to the function $G_{\Lambda,T}$ built in Proposition IV.1.19, we get the existence of a function $g_{\Lambda,\tau} \in L^2(\mathbb{R})$ such that

$$G_{\Lambda,\tau}(z) = \int_{-\tau}^{\tau} g_{\Lambda,\tau}(t) e^{itz}\, dt,$$

and

$$\|g_{\Lambda,\tau}\|_{L^2(\mathbb{R})} = \|G_{\Lambda,\tau}\|_{L^2(\mathbb{R})} \leq C_2 e^{C_2 \tau^{-\frac{\alpha}{1-\alpha}}}.$$

We compute the following integral

$$\begin{aligned}\int_{-\tau}^{\tau} f(t+\tau)g_{\Lambda,\tau}(t)\,dt &= f_0\int_{-\tau}^{\tau} g_{\Lambda,\tau}(t)\,dt + \sum_{\lambda\in\Lambda} f_\lambda e^{-\lambda\tau}\int_{-\tau}^{\tau} e^{-\lambda t}g_{\Lambda,\tau}(t)\,dt\\ &= f_0 G_{\Lambda,\tau}(0) + \sum_{\lambda\in\Lambda} f_\lambda e^{-\lambda\tau} G_{\Lambda,\tau}(i\lambda)\\ &= f_0,\end{aligned}$$

by using the properties of $G_{\Lambda,\tau}$. The conclusion follows from the Cauchy-Schwarz inequality and the estimate of $\|g_{\Lambda,\tau}\|_{L^2(\mathbb{R})}$. ■

We recall from Appendix A.7 that the set of Müntz polynomial functions $M(\Lambda\cup\{0\})$ is the set of functions defined as

$$p(x) = p_0 + \sum_{\lambda\in\Lambda} p_\lambda x^\lambda, \quad x\in[0,+\infty),$$

where only a finite number of coefficients p_λ are non zero.

Proposition IV.1.24

There exists $C_9 > 0$ such that for any τ satisfying (IV.34) and $\tau < 1$ we have

$$|p(0)| \leq C_9 e^{C_9\tau^{-\frac{\alpha}{1-\alpha}}}\|p\|_{L^\infty(1-\tau,1)}, \quad \forall p\in M(\Lambda\cup\{0\}).$$

Proof :

We set

$$f(t) \overset{\text{def}}{=} p(e^{-t}), \quad \forall t>0.$$

By construction, we have $f\in\mathcal{E}_{\Lambda\cup\{0\}}$ so that we can apply Proposition IV.1.23. Since $p(0)=p_0$ we get

$$|p(0)| \leq C_8 e^{C_8\tau^{-\frac{\alpha}{1-\alpha}}}\|f\|_{L^2(0,2\tau)}.$$

Since $\tau<1$, we can bound the L^2 norm by the L^∞ norm

$$\begin{aligned}|p(0)| &\leq C_8 e^{C_8\tau^{-\frac{\alpha}{1-\alpha}}}\|f\|_{L^\infty(0,2\tau)}\\ &\leq C_8 e^{C_8\tau^{-\frac{\alpha}{1-\alpha}}}\|p\|_{L^\infty(e^{-2\tau},1)}.\end{aligned}$$

Since $e^{-2\tau}\geq 1-2\tau$, we finally get

$$|p(0)| \leq C_8 e^{C_8\tau^{-\frac{\alpha}{1-\alpha}}}\|p\|_{L^\infty(1-2\tau,1)},$$

and the claim is proved by changing τ in $\tau/2$ and adapting the constant accordingly. ■

Theorem IV.1.25

Let $s>0$ and A be a closed subset of $[0,1]$ whose Lebesgue measure is at least s. Under the same assumptions as above, we have

$$\|p\|_{L^\infty(0,\inf A)} \leq C_9 e^{C_9 s^{-\frac{\alpha}{1-\alpha}}}\|p\|_{L^\infty(A)}, \quad \forall p\in M(\Lambda\cup\{0\}).$$

Proof :

Let $L_0\subset\Lambda\cup\{0\}$ be the finite subset corresponding to the non zero coefficients of p in the basis of $M(\Lambda)$. We define the interval $I_s=[1-s,1]$.

Let T_{L_0,I_s} be the generalized Tchebychev polynomial corresponding to L_0 and to the set I_s as defined in Appendix A.7.

We use Theorem A.7.31 with $I = I_s$ (since $|A| \geq s = |I_s|$ and $\sup A \leq 1 = \sup I_s$) and we deduce that

$$\|p\|_{L^\infty(0,\inf A)} \leq |T_{L_0,I_s}(0)| \, \|p\|_{L^\infty(A)}.$$

Applying Proposition IV.1.24 to T_{L_0,I_s} (and τ replaced by s) we get

$$|T_{L_0,I_s}(0)| \leq C_9 e^{C_9 s^{-\frac{\alpha}{1-\alpha}}},$$

and the claim is proved. ∎

We can now move to a similar L^2 estimate.

Theorem IV.1.26

There exists $C_{10} > 0$ such that for any $0 < s < 1$, we have

$$\|p\|_{L^2(0,1)} \leq C_{10} e^{C_{10} s^{-\frac{\alpha}{1-\alpha}}} \|p\|_{L^2(1-s,1)}, \quad \forall p \in M(\Lambda \cup \{0\}).$$

Proof :

For any $s > 0$ and $p \in M(\Lambda \cup \{0\})$, we introduce the compact set

$$A_s = \left\{ x \in [1-s,1], \;\; |p(x)| \leq \sqrt{\frac{2}{s}} \|p\|_{L^2(1-s,1)} \right\},$$

and

$$B_s = [1-s,1] \setminus A_s.$$

Integrating $|p|^2$ on B_s we get

$$\|p\|^2_{L^2(1-s,1)} \geq \int_{B_s} |p|^2 \geq \frac{2}{s} \|p\|^2_{L^2(1-s,1)} |B_s|,$$

from which we deduce that

$$|B_s| \leq s/2,$$

and consequently

$$|A_s| \geq s/2.$$

We apply Theorem IV.1.25 to this set A_s to get

$$\|p\|_{L^\infty(0,1-s)} \leq C_9 e^{C_9 s^{-\frac{\alpha}{1-\alpha}}} \|p\|_{L^\infty(A_s)} \leq C_9 e^{C C_9 s^{-\frac{\alpha}{1-\alpha}}} \sqrt{\frac{2}{s}} \|p\|_{L^2(1-s,1)},$$

and consequently

$$\|p\|_{L^2(0,1-s)} \leq C_9 e^{C_9 s^{-\frac{\alpha}{1-\alpha}}} \sqrt{\frac{2}{s}} \|p\|_{L^2(1-s,1)},$$

and finally

$$\|p\|^2_{L^2(0,1)} \leq C_9 \left(1 + e^{2C_9 s^{-\frac{\alpha}{1-\alpha}}} \frac{2}{s} \right) \|p\|^2_{L^2(1-s,1)}.$$

The claim is proved. ∎

We can now come back to our original problem and prove the expected result.

Proof (of Theorem IV.1.18):

We set $\lambda_0 \stackrel{\text{def}}{=} \min\left(\frac{1}{R(0)}, 1\right)$. Let $f = \sum_{\lambda\in\Lambda} a_\lambda e[\lambda] \in \mathcal{E}_\Lambda$ and let $0 \leq \tilde{t} \leq +\infty$. By using straightforward changes of variable we get

$$\begin{aligned}
\int_0^{\tilde{t}} |f(t)|^2\,dt &= \int_0^{\tilde{t}} \left|\sum_{\lambda\in\Lambda} a_\lambda e^{-\lambda t}\right|^2 dt \\
&= \int_0^{\tilde{t}} \left|\sum_{\lambda\in\Lambda} a_\lambda e^{-(\lambda-\lambda_0/2)t}\right|^2 e^{-\lambda_0 t}\,dt \\
&= \frac{1}{\lambda_0}\int_0^{\lambda_0\tilde{t}} \left|\sum_{\lambda\in\Lambda} a_\lambda e^{-\frac{\lambda-\lambda_0/2}{\lambda_0}t}\right|^2 e^{-t}\,dt \\
&= \int_{e^{-\lambda_0\tilde{t}}}^1 \left|\sum_{\lambda\in\Lambda} a_\lambda x^{\frac{\lambda-\lambda_0/2}{\lambda_0}}\right|^2 dx.
\end{aligned} \tag{IV.36}$$

Let us define a new family $\tilde{\Lambda}$ as follows

$$\tilde{\Lambda} \stackrel{\text{def}}{=} \left\{\frac{\lambda-\lambda_0/2}{\lambda_0},\ \ \lambda\in\Lambda\right\} \subset (0,+\infty).$$

By (A.14), we see that $\inf\Lambda \geq \lambda_0$, and thus we have

$$\inf\tilde{\Lambda} = \frac{(\inf\Lambda)-\lambda_0/2}{\lambda_0} \geq \frac{1}{2}.$$

In particular, the counting function $\tilde{N}$ of this new family satisfies

$$\tilde{N}(r) = 0,\ \ \forall r < \frac{1}{2},$$

and, moreover

$$\tilde{N}(r) = N(\lambda_0/2+\lambda_0 r) \leq N(2\lambda_0 r) \leq N(2r) \leq \bar{N}2^\alpha r^\alpha,\ \ \forall r \geq \frac{1}{2},$$

since $\lambda_0 \leq 1$. Therefore, $\tilde{N}$ satisfies the same assumption as (IV.21) with $\bar{N}$ changed into $\bar{N}2^\alpha$.

We then apply Theorem IV.1.26 to $q(x) \stackrel{\text{def}}{=} \sum_{\lambda\in\Lambda} p_\lambda x^{\left(\frac{\lambda}{\lambda_0}-\frac{1}{2}\right)} \in M(\tilde{\Lambda})$, that we reformulate by using formula (IV.36) with $\tilde{t} = +\infty$ and $\tilde{t} = -\log(1-s)/\lambda_0$. It follows

$$\int_0^{+\infty} |f(t)|^2\,dt \leq C_{10}e^{C_{10}s^{-\frac{\alpha}{1-\alpha}}} \int_0^{-\frac{\log(1-s)}{\lambda_0}} |f(t)|^2\,dt.$$

Since $-\log(1-s) \leq 2s$ for any $s \in (0,1/2)$, we deduce that

$$\int_0^{+\infty} |f(t)|^2\,dt \leq C_{10}e^{C_{10}s^{-\frac{\alpha}{1-\alpha}}} \int_0^{\frac{2s}{\lambda_0}} |f(t)|^2\,dt,$$

from which, for any $T < \frac{1}{\lambda_0}$, we can set $s = T\lambda_0/2$ and obtain

$$\int_0^{+\infty} |f(t)|^2\,dt \leq C_{10}e^{C_{10}\left(\frac{T\lambda_0}{2}\right)^{-\frac{\alpha}{1-\alpha}}} \int_0^T |f(t)|^2\,dt,$$

and the proof is complete for $T \leq \frac{1}{\lambda_0}$. For $T > \frac{1}{\lambda_0}$, the result is a straightforward consequence of the previous case. ■

IV.1.2.5 Control cost estimate.

By using the material of the previous section, we can now obtain an estimate of the control cost for the heat equation (both in the distributed and boundary control cases).

Theorem IV.1.27

The null-control v obtained in Theorem IV.1.11 (resp. in IV.1.12) satisfies the estimate

$$\|v\|_{L^2(0,T)} \leq Ce^{\frac{C}{T}}\|y_0\|_{L^2(\Omega)},$$

$$\left(\text{resp. } \|v\|_{L^2((0,T)\times\omega)} \leq Ce^{\frac{C}{T}}\|y_0\|_{L^2(\Omega)}\right),$$

where C does not depend on T and y_0.

Proof :

We use here the precised Weyl's law for the elliptic operator $\mathcal{A}$ defined in (IV.1) that says that the counting function associated with $\Lambda = \mathrm{Sp}(\mathcal{A})$ satisfies

$$|N(r) - \bar{N}\sqrt{r}| \leq \tilde{N},$$

where $\bar{N}$ depends only on the coefficients appearing in the operator $\mathcal{A}$.

With this information at hand and the sharp estimates given by Theorems IV.1.10 and by Theorem IV.1.18, we can conclude that the biorthogonal family $(q_{\lambda,T})_{\lambda\in\Lambda}$ associated to the spectrum of $\mathcal{A}$ satisfy the estimate

$$\|q_{\lambda,T}\|_{L^2(0,T)} \leq Ce^{\frac{C}{T}}e^{C\sqrt{\lambda}}, \quad \forall \lambda \in \Lambda,$$

where C does not depend on T.

Coming back, for instance, to the proof of Theorem IV.1.11, we find that for any $\mu \in \Lambda$, the estimate (IV.23) now leads to

$$\|u_\mu\|_{L^2(0,T)} \leq C\frac{\|y_0\|}{\sqrt{\mu}}e^{\frac{C}{T}}e^{-\mu T}e^{C\sqrt{\mu}}.$$

Young's inequality gives

$$C\sqrt{\mu} \leq \frac{\mu T}{2} + \frac{C^2}{2T},$$

so that we finally obtain

$$\|u_\mu\|_{L^2(0,T)} \leq C'\frac{\|y_0\|}{\sqrt{\mu}}e^{\frac{C'}{T}}e^{-\mu T/2},$$

for a constant C' that does not depend on T. The end of the proof is now the same and the cost of the control is obtained by summing all those estimates that makes appear the expected exponential term depending on T.

The argument is exactly the same in the case of the distributed control. ■

IV.1.3 The multi-D case

This will be the opportunity to encounter our first *Carleman estimate*. Those are weighted a priori estimate on solutions of PDEs that imply many important qualitative properties for those PDEs such as unique continuation, spectral estimates, and so on. We refer for instance to the references [LRL11] and [Cor07].

We first state the following two estimates without proof. We shall actually give the proof of a slightly more general estimate in Section IV.3.

Theorem IV.1.28 (Boundary Carleman estimate)

Let Γ be a non empty open subset of $\partial\Omega$. There exists a function $\varphi \in \mathcal{C}^2(\overline{\Omega})$, a $C > 0$ and $s_0 > 0$ such that, for any $u \in H^2(\Omega) \cap H^1_0(\Omega)$ and any $s \geq s_0$, we have

$$s^3 \|e^{s\varphi} u\|^2_{L^2(\Omega)} + s\|e^{s\varphi}\nabla u\|^2_{L^2(\Omega)} \leq C\left(\|e^{s\varphi}\Delta u\|^2_{L^2(\Omega)} + s\|e^{s\varphi}\partial_n u\|^2_{L^2(\Gamma)}\right). \tag{IV.37}$$

Theorem IV.1.29 (Interior Carleman estimate)

Let ω be a non empty open subset of Ω. There exists a function $\varphi \in \mathcal{C}^2(\overline{\Omega})$, a $C > 0$ and a $s_0 > 0$ such that, for any $u \in H^2(\Omega) \cap H^1_0(\Omega)$ and any $s \geq s_0$, we have

$$s^3 \|e^{s\varphi} u\|^2_{L^2(\Omega)} + s\|e^{s\varphi}\nabla u\|^2_{L^2(\Omega)} \leq C\left(\|e^{s\varphi}\Delta u\|^2_{L^2(\Omega)} + s^3\|e^{s\varphi} u\|^2_{L^2(\omega)}\right). \tag{IV.38}$$

Proposition IV.1.30

Let $\omega \subset \Omega$ and $\Gamma \subset \partial\Omega$ as before, then the eigenfunctions of $\mathcal{A}$ satisfy

$$\|\phi\|_{L^2(\omega)} \neq 0, \text{ and } \|\partial_n \phi\|_{L^2(\Gamma)} \neq 0, \quad \forall \phi \in \mathit{Ker}\,(\mathcal{A} - \lambda) \setminus \{0\}, \forall \lambda \in \Lambda.$$

Proof :

We start from the equation satisfied by ϕ under the following form

$$-\gamma(\Delta\phi) - 2\nabla\phi\cdot\nabla\gamma - (\Delta\gamma)\phi + \alpha\phi = \lambda\phi,$$

which gives

$$\Delta\phi = \frac{\alpha-\lambda}{\gamma}\phi - 2\frac{\nabla\phi\cdot\nabla\gamma}{\gamma} - \frac{\Delta\gamma}{\gamma}\phi.$$

We deduce the pointwise inequality

$$|\Delta\phi| \leq C_{\alpha,\gamma}(1+|\lambda|)|\phi| + C_\gamma|\nabla\phi|.$$

- Assume first that $\phi = 0$ on ω. We can apply (IV.38) in which the observation term cancels and we get

  $$s^3\|e^{s\varphi}\phi\|^2_{L^2(\Omega)} + s\|e^{s\varphi}\nabla\phi\|^2_{L^2(\Omega)} \leq C(1+\lambda^2)\|e^{s\varphi}\phi\|^2_{L^2(\Omega)} + C\|e^{s\varphi}\nabla\phi\|^2_{L^2(\Omega)}.$$

  Taking s large enough (depending on k) we can conclude that

  $$s^3\|e^{s\varphi}\phi\|^2_{L^2(\Omega)} + s\|e^{s\varphi}\nabla\phi\|^2_{L^2(\Omega)} \leq 0,$$

  which implies $\phi = 0$ and thus a contradiction.

- If we assume that $\partial_n\phi = 0$ on Γ, we apply the same reasoning with the other Carleman estimate.

■

Remark IV.1.31

The reasoning above shows that for $s = C_1\lambda^{2/3}$ we have

$$s^3\|e^{s\varphi}\phi\|^2_{L^2(\Omega)} + s\|e^{s\varphi}\nabla\phi\|^2_{L^2(\Omega)} \leq Cs^3\|e^{s\varphi}\phi\|^2_{L^2(\omega)},$$

and thus

$$C_1^3 s^3 e^{2s\inf\varphi}\|\phi\|^2_{L^2(\Omega)} \leq Cs^3 e^{2s\sup\varphi}\|\phi\|^2_{L^2(\omega)}.$$

Since $\|\phi\|_{L^2(\Omega)} = 1$, we deduce

$$\|\phi\|^2_{L^2(\omega)} \geq Ce^{-C_3 s} = Ce^{-C_4\lambda^{2/3}}.$$

Similarly, we can show

$$\|\partial_n\phi\|^2_{L^2(\Gamma)} \geq Ce^{-C\lambda^{2/3}}.$$

However, with the above elements, we have proved the approximate controllability properties for the heat equation. Indeed, using the Fattorini-Hautus theorem (Theorem III.3.7), we see that the claim of Proposition IV.1.30 exactly gives the following result.

Theorem IV.1.32

Under the above assumptions, both problems (IV.8) and (IV.9) are approximately controllable from any initial data $y_0 \in L^2(\Omega)$ and at any time $T > 0$.

IV.2 The method of Lebeau and Robbiano

In order to deal with the null-controllability problem in dimension greater than 1, we will need a much stronger spectral property for the eigenfunctions of $\mathcal{A}$.

More precisely, we will prove the following spectral inequality (taken from [LRL11], see also [LR95]) that will be crucial in our analysis.

Theorem IV.2.33 (Lebeau-Robbiano spectral inequality)

Let Ω as before and ω a non empty open subset of Ω. There exists a $C > 0$ depending only on α, γ, ω such that: for any $\mu > 0$ we have

$$\|\phi\|_{L^2(\Omega)} \leq Ce^{C\sqrt{\mu}}\|\phi\|_{L^2(\omega)}, \quad \forall \phi \in E_\mu,$$

where E_μ is defined in (IV.6).

Remark IV.2.34

The above spectral inequality (as well as the proof below of the controllability result) does not hold for the boundary control problem. This is very easy to see, even in 1D for instance, that for any two eigenvalues $\lambda \neq \mu$, we can find a non trivial linear combination $\phi = a_\lambda\phi_\lambda + a_\mu\phi_\mu$ such that $\partial_x\phi_{|x=0} = 0$.

The above spectral inequality can be proved by means of another kind of global elliptic Carleman estimate that will be proved in Section IV.3. We only give here the simplified version of this Carleman estimate that we need at that point and proceed to the proof of the spectral inequality.

Proposition IV.2.35

Let Ω and ω as before. Let $T^* > 0$ be given and we set $Q = (0, T^*) \times \Omega$. There exists a positive function $\varphi \in \mathcal{C}^2(\overline{Q})$ such that $\nabla_x \varphi(T^*, .) = 0$ and $C, s_0 > 0$ such that:
For any $s \geq s_0$, and any function $u \in \mathcal{C}^2(\overline{Q})$ satisfying $u(0, .) = 0$ and $u = 0$ on $[0, T] \times \partial\Omega$, we have the estimate

$$s^3 e^{2s\varphi(T^*)} \int_\Omega |u(T^*, .)|^2 \leq C s e^{2s\varphi(T^*)} \int_\Omega |\nabla_x u(T^*, .)|^2 + Cs \int_\omega |e^{s\varphi(0,.)} \partial_\tau u(0, .)|^2 + 2\|e^{s\varphi}(\partial_\tau^2 u - \mathcal{A}u)\|^2_{L^2(Q)}.$$

Proof (of Theorem IV.2.33):

Let us consider any element $v \in E_\mu$, that we write

$$v = \sum_{\substack{\lambda \in \Lambda \\ \lambda \leq \mu}} v_\lambda \in E_\mu,$$

with $v_\lambda \in \mathrm{Ker}\,(\mathcal{A} - \lambda)$ for each λ. We define the function $u : Q \to \mathbb{R}$ as follows

$$u(\tau, x) = \sum_{\substack{\lambda \in \Lambda \\ \lambda \leq \mu}} \frac{\sinh(\sqrt{\lambda}\tau)}{\sqrt{\lambda}} v_\lambda(x).$$

This function is the unique solution of the following Cauchy problem for the elliptic augmented operator $\partial_\tau^2 - \mathcal{A}$, indeed we have

$$u(0, .) = 0, \quad \partial_\tau u(0, .) = v, \quad (\partial_\tau^2 - \mathcal{A})(u) = 0.$$

We can apply the above Carleman estimate to this particular function u and find

$$s^3 e^{2s\varphi(T)} \int_\Omega |u(T^*, .)|^2 \leq Cs \int_\omega |e^{s\varphi(0,.)} v|^2 + C s e^{2s\varphi(T)} \int_\Omega |\nabla_x u(T^*, .)|^2. \tag{IV.39}$$

Let us compute the norms at time T^*:

- Since the v_λ are pairwise orthogonal in $L^2(\Omega)$, we simply have

$$\int_\Omega |u(T^*, .)|^2 = \sum_{\substack{\lambda \in \Lambda \\ \lambda \leq \mu}} \frac{\|v_\lambda\|^2_{L^2}}{\lambda} |\sinh(\sqrt{\lambda}T^*)|^2 \geq \frac{1}{\mu} \sum_{\substack{\lambda \in \Lambda \\ \lambda \leq \mu}} \|v_\lambda\|^2_{L^2} |\sinh(\sqrt{\lambda}T^*)|^2. \tag{IV.40}$$

- For the gradient term, we first observe that

$$\begin{aligned}\int_\Omega |\nabla_x u(T^*, .)|^2 \leq C \int_\Omega \gamma |\nabla_x u(T^*, .)|^2 &= C\langle \mathcal{A}u(T, *), u(T^*, .)\rangle_{L^2(\Omega)} - C \int_\Omega \alpha |u(T^*, .)|^2 \\ &\leq C\langle \mathcal{A}u(T, *), u(T^*, .)\rangle_{L^2(\Omega)} + C \int_\Omega |u(T^*, .)|^2.\end{aligned}$$

  Then we use that, for any λ, λ', we have

$$\langle \mathcal{A}v_\lambda, v_{\lambda'}\rangle_{L^2} = \lambda \|v_\lambda\|^2_{L^2} \delta_{\lambda,\lambda'},$$

  to write

$$\langle \mathcal{A}u(T^*, .), u(T^*, .)\rangle = \sum_{\substack{\lambda \in \Lambda \\ \lambda \leq \mu}} \|v_\lambda\|^2_{L^2} |\sinh(\sqrt{\lambda}T^*)|^2.$$

  Using (IV.40), we have finally proved that

$$\int_\Omega |\nabla_x u(T^*, .)|^2 \leq C(1 + \mu) \int_\Omega |u(T^*, .)|^2. \tag{IV.41}$$

Using (IV.41) in (IV.39), we have finally obtained

$$s^3 e^{2s\varphi(T)} \int_\Omega |u(T^*,.)|^2 \leq Cs \int_\omega |e^{s\varphi(0,.)}v|^2 + Cse^{2s\varphi(T)}(1+\mu)\int_\Omega |u(T^*,.)|^2.$$

Since this inequality holds for any value of s, large enough, we see that we can choose $s = \tilde{C}\sqrt{\mu}$ for some $\tilde{C}$ in order to absorb the last term by the left-hand side term of the inequality. It remains, for this particular value of s

$$\mu^{3/2} e^{C\sqrt{\mu}\varphi(T)} \int_\Omega |u(T^*,.)|^2 \leq C\sqrt{\mu}\int_\omega |e^{C\sqrt{\mu}\varphi(0,.)}v|^2,$$

and then, changing the values of the constants if necessary, we get

$$\int_\Omega |u(T^*,.)|^2 \leq \frac{C}{\mu} e^{C\sqrt{\mu}} \|v\|^2_{L^2(\omega)}.$$

To conclude, we use the inequality $|\sinh(t)/t| \geq 1$ for any $t\in\mathbb{R}$, to write

$$\int_\Omega |u(T^*,.)|^2 = \sum_{\substack{\lambda\in\Lambda\\ \lambda\leq\mu}} \|v_\lambda\|^2_{L^2} \left|\frac{\sinh(\sqrt{\lambda}T^*)}{\sqrt{\lambda}}\right|^2 \geq C_{T^*} \sum_{\substack{\lambda\in\Lambda\\ \lambda\leq\mu}} \|v_\lambda\|^2_{L^2} = C_{T^*}\|v\|^2_{L^2}.$$

■

With this inequality at hand we can prove a partial observability inequality and a related partial distributed controllability result. We recall that we assume that all the eigenvalues of $\mathcal{A}$ are positive.

Proposition IV.2.36

There exists a $C>0$ such that for any time $\tau>0$ and any $\mu>0$, we have the following inequality

$$\|e^{-\tau\mathcal{A}}q_T\|^2_E \leq \frac{Ce^{C\sqrt{\mu}}}{\tau}\int_0^\tau \|e^{-(\tau-s)\mathcal{A}}q_T\|^2_{L^2(\omega)}\,ds, \quad \forall q_T\in E_\mu.$$

Note that the operator $\mathcal{A}$ is self-adjoint and thus the adjoint operator that we should have put in this inequality is nothing but $\mathcal{A}^* = \mathcal{A}$. Moreover, we also have $\mathcal{B} = \mathcal{B}^* = \mathbf{1}_\omega$ which explains the form of the right hand side.

Proof :

Since the space E_μ is stable by the operator $\mathcal{A}$ (it is built upon its eigenfunctions), we know that $e^{-(\tau-s)\mathcal{A}}q_T$ belongs to E_μ as soon as $q_T\in E_\mu$. Therefore, we can apply the Lebeau-Robbiano spectral inequality to this particular element of E_μ

$$\|e^{-(\tau-s)\mathcal{A}}q_T\|^2_{L^2(\Omega)} \leq Ce^{C\sqrt{\mu}}\|e^{-(\tau-s)\mathcal{A}}q_T\|^2_{L^2(\omega)}.$$

By the dissipation estimate (IV.5), we find that

$$\|e^{-\tau\mathcal{A}}q_T\|^2_{L^2(\Omega)} \leq Ce^{C\sqrt{\mu}}\|e^{-(\tau-s)\mathcal{A}}q_T\|^2_{L^2(\omega)},$$

(with λ_1 possibly negative). We can now integrate this inequality with respect to s on $(0,\tau)$ to find

$$\tau\|e^{-\tau\mathcal{A}}q_T\|^2_{L^2(\Omega)} \leq Ce^{C\sqrt{\mu}}\int_0^\tau \|e^{-(\tau-s)\mathcal{A}}q_T\|^2_{L^2(\omega)},$$

which gives the result. ■

For any $\mu>0$, and $\tau>0$, we consider the following finite dimensional control problem

$$\begin{cases} \partial_t y + \mathcal{A}y = P_\mu(1_\omega v(t,x)) \\ y(0) = y_{0,\mu}\in E_\mu, \end{cases} \tag{IV.42}$$

with $v \in L^2(0,\tau;E_\mu)$. Since E_μ is stable by $\mathcal{A}$, this problem can be recast in the ODE form

$$y'(t) + A_\mu y = B_\mu v,$$

by setting $A_\mu = \mathcal{A}_{|E_\mu}$ and $B_\mu = P_\mu(1_\omega.)$. The state space is $E = E_\mu$ and the control space is also $U = E_\mu$ with their natural inner product.

We observe that

$$A_\mu^* = A_\mu, \text{ and } B_\mu^* = B_\mu.$$

Corollary IV.2.37

For any $\mu > 0$, $\tau > 0$ and $y_{0,\mu} \in E_\mu$, the partial control System (IV.42) is null-controllable at time τ and more precisely, there exists control $v_\mu \in L^2(0,\tau,E_\mu)$ such that the solution satisfies $y(\tau) = 0$ and such that

$$\|v_\mu\|_{L^2(0,\tau;E_\mu)} \leq C\frac{e^{C\sqrt{\mu}}}{\sqrt{\tau}}\|y_{0,\mu}\|_{E_\mu}.$$

Proof :

We simply use the results we proved in the finite dimensional framework and in particular the second point of Theorem II.7.25. ■

Proposition IV.2.38

For any $\mu > 0$, $\tau > 0$ and $y_0 \in E$, there exists a control $v_\mu \in L^2(0,\tau,L^2(\Omega))$ for our original system (IV.8) such that

$$P_\mu y(\tau) = 0,$$

and

$$\|v_\mu\|_{L^2(0,\tau;E)} \leq C\frac{e^{C\sqrt{\mu}}}{\sqrt{\tau}}\|y_0\|_E,$$
$$\|y(\tau)\|_E \leq C_2 e^{C_2\sqrt{\mu}}\|y_0\|_E.$$

Proof :

We take v_μ to be the control for the partial control system obtained in Corollary IV.2.37 with the initial data $y_{0,\mu} = P_\mu y_0$. Let y be the solution of the full system associated with this control

$$\partial_t y + \mathcal{A}y = 1_\omega v_\mu, \;\; y(0) = y_0.$$

We apply the projector P_μ (which commutes with $\mathcal{A}$) to get

$$\partial_t(P_\mu y) + \mathcal{A}(P_\mu y) = P_\mu(1_\omega v_\mu), \;\; (P_\mu y)(0) = P_\mu y_0.$$

This proves that $P_\mu y$ is the (unique) solution of (IV.42), and by construction we have $P_\mu y(\tau) = 0$. Moreover, since P_μ is an orthogonal projection in E, we have

$$\|v_\mu\|_{L^2(0,\tau;E)} \leq Ce^{C\sqrt{\mu}}\|P_\mu y_0\|_E \leq Ce^{C\sqrt{\mu}}\|y_0\|_E.$$

Finally, we write the Duhamel formula

$$y(\tau) = y_0 + \int_0^\tau e^{-(\tau-s)\mathcal{A}}\mathcal{B}v_\mu(s)\,ds,$$

and take the norm in E

$$\|y(\tau)\|_E \leq \|y_0\|_E + \int_0^\tau \|e^{-(\tau-s)\mathcal{A}}\mathcal{B}v_\mu(s)\|_E\,ds.$$

We use now the dissipation estimate for $\mathcal{A}$ (IV.5) (with $\lambda_1 > 0$ here) and the fact that $\mathcal{B} = 1_\omega$ is bounded with norm 1. It follows

$$\|y(\tau)\|_E \leq \|y_0\|_E + C\int_0^\tau \|v_\mu(s)\|_E\,ds \leq \|y_0\|_E + C\sqrt{\tau}\|v_\mu\|_{L^2(0,\tau;E)},$$

and the conclusion follows by the estimate we got on the norm of v_μ. ■

Corollary IV.2.39

For any $\mu > 0$, $0 < \tau < T$ and $y_0 \in E$, there exists a control $v_\mu \in L^2(0,\tau,L^2(\Omega))$ such that

$$\|v_\mu\|_{L^2(0,\tau;E)} \leq C\frac{e^{C\sqrt{\mu}}}{\sqrt{\tau}}\|y_0\|_E,$$
$$\|y(\tau)\|_E \leq C_2 e^{C_2\sqrt{\mu}-\frac{\tau\mu}{2}}\|y_0\|_E.$$

Proof :

The idea is to use the previous proposition on the time interval $(0,\tau/2)$. This gives us a control $w_\mu \in L^2(0,\tau/2;E)$ such that $P_\mu y(\tau/2) = 0$ and

$$\|w_\mu\|_{L^2(0,\tau/2;E)} \leq C\frac{e^{C\sqrt{\mu}}}{\sqrt{\tau}}\|y_0\|_E,$$
$$\|y(\tau/2)\|_E \leq C_2 e^{C_2\sqrt{\mu}}\|y_0\|_E.$$

Now, on the second half of the time interval we *do nothing* in order to take advantage of the natural dissipation of the system and to the fact that all frequencies less than μ have been killed at time $\tau/2$. It means that the control we finally consider is

$$v_\mu(t) = \begin{cases} w_\mu(t), & \text{for } t \in (0,\tau/2), \\ 0, & \text{for } t \in (\tau/2,\tau). \end{cases}$$

It is clear that v_μ and w_μ have the same L^2-norm. Moreover, since $v_\mu = 0$ on $(\tau/2,\tau)$, we have

$$y(\tau) = e^{-\frac{\tau}{2}\mathcal{A}}y(\tau/2),$$

and thus, since $P_\mu y(\tau/2) = 0$, it follows by (IV.7)

$$\|y(\tau)\|_E \leq e^{-\frac{\tau}{2}\mu}\|y(\tau/2)\|_E \leq C_2 e^{C_2\sqrt{\mu}-\frac{\tau\mu}{2}}\|y_0\|_E.$$

■

Theorem IV.2.40 (Lebeau-Robbiano null-controllability theorem [LR95])

For any $T > 0$, the heat-like equation (IV.2), is null-controllable at time T.

Proof :

The idea is to split the time interval $(0,T)$ into small subintervals of size τ_j, $j \geq 1$ with

$$\sum_{j\geq 1}\tau_j = T,$$

and to apply successively a partial control as in the previous corollary with a cut frequency μ_j that tends to infinity when $j \to \infty$.

More precisely, we set

$$\tau_j = \frac{T}{2^j}, \text{ and } \mu_j = \beta(2^j)^2,$$

with $\beta > 0$ to be determined later.

Let $T_j = \sum_{k=1}^j \tau_k$, for $j \geq 1$.

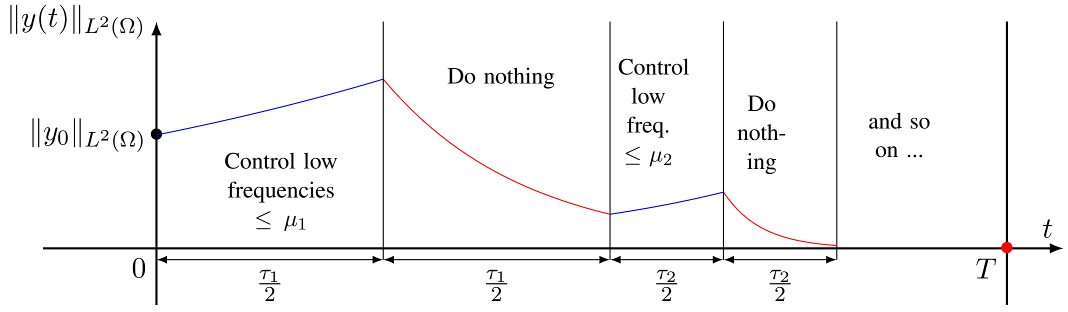

Figure IV.1: The Lebeau-Robbiano method

- During the time interval $(0, \tau_1) = (0, T_1)$, we apply a control v_1 as given by Corollary IV.2.39 with $\mu = \mu_1$, in such a way that

$$\|v_1\|_{L^2(0,T_1,E)} \le C\frac{e^{C\sqrt{\mu_1}}}{\sqrt{\tau_1}}\|y_0\|_E,$$

$$\|y(T_1)\|_E \le C_2 e^{C_2\sqrt{\mu_1}-\frac{\tau_1\mu_1}{2}}\|y_0\|_E.$$

- During the time interval $(\tau_1, \tau_1 + \tau_2)$ we apply a control v_2 as given by Corollary IV.2.39 with $\mu = \mu_2$, in such a way that

$$\|v_2\|_{L^2(T_1,T_2;E)} \le C\frac{e^{C\sqrt{\mu_2}}}{\sqrt{\tau_2}}\|y(T_1)\|_E,$$

$$\|y(T_2)\|_E \le C_2^2 e^{C_2(\sqrt{\mu_1}+\sqrt{\mu_2})-\frac{\tau_1\mu_1}{2}-\frac{\tau_2\mu_2}{2}}\|y_0\|_E.$$

- And so on, by induction we build a control v_j on the time interval (T_{j-1}, T_J) such that

$$\|v_j\|_{L^2(T_{j-1},T_j;E)} \le C\frac{e^{C\sqrt{\mu_j}}}{\sqrt{\tau_j}}\|y(T_{j-1})\|_E,$$

$$\|y(T_j)\|_E \le C_2^j e^{C_2\sum_{k=1}^j\sqrt{\mu_k}-\frac{1}{2}\sum_{k=1}^j\tau_k\mu_k}\|y_0\|_E.$$

- By construction, we have

$$\begin{aligned} C_2\sum_{k=1}^j\sqrt{\mu_k} - \frac{1}{2}\sum_{k=1}^j\tau_k\mu_k &= C_2\sqrt{\beta}\sum_{k=1}^j 2^k - \frac{\beta}{2}T\sum_{k=1}^j 2^k \\ &= (C_2\sqrt{\beta} - \tfrac{\beta}{2}T)(2^{j+1}-1) \end{aligned}$$

We are thus led to choose β large enough so that

$$\tilde{\beta} \stackrel{\text{def}}{=} \frac{\beta}{2}T - C_2\sqrt{\beta} > 0,$$

and we have obtained that for any j,

$$\|y(T_j)\|_E \le C_3 C_2^j e^{-\tilde{\beta}2^{j+1}}\|y_0\|_E.$$

- Going back to the estimate of the norm of v_j, we have

$$\begin{aligned} \|v_j\|_{L^2(T_{j-1},T_j;E)} &\le C\frac{e^{C\sqrt{\mu_j}}}{\sqrt{\tau_j}}\|y(T_{j-1})\|_E \\ &\le \frac{CC_3}{\sqrt{T}}2^{j/2}C_2^{j-1}e^{C\sqrt{\beta}2^j-\tilde{\beta}2^j}\|y_0\|_E. \end{aligned}$$

Wee that we can choose β even larger to ensure that

$$\bar{\beta} \stackrel{\text{def}}{=} \tilde{\beta} - C\sqrt{\beta} > 0.$$

We finally got the estimate

$$\|v_j\|_{L^2(T_{j-1},T_j;E)} \leq \frac{CC_3}{\sqrt{T}} 2^{j/2} C_2^{j-1} e^{-\bar{\beta}2^j} \|y_0\|_E.$$

- All the previous estimates show that

$$\sum_{j\geq 1} \|v_j\|^2_{L^2(T_{j-1},T_j;E)} < +\infty,$$

and in particular the function v that is obtained by gluing all together the $(v_j)_j$ is an element of $L^2(0,T;E)$. The associated solution y of the PDE is continuous in time on $[0,T]$ with values in E and satisfies

$$\|y(T_j)\| \leq C_3 C_2^j e^{-\tilde{\beta}2^{j+1}} \|y_0\|_E \xrightarrow[j\to\infty]{} 0.$$

This implies $y(T) = 0$, since $T_j \to T$ as $j \to \infty$.

The claim is proved.

■

Remark IV.2.41

A careful inspection of the proof shows that one can take β of the form

$$\beta = \frac{\alpha}{T^2},$$

with $\alpha > 0$ large enough independent of T. It follows that $\tilde{\beta}$ and $\bar{\beta}$ will be proportional to $1/T$ and therefore we can obtain the following estimate on the control cost

$$\|v\|_{L^2(0,T;E)} \leq Ce^{\frac{C}{T}} \|y_0\|_E.$$

This exponential behavior of the cost in the limit $T \to 0$ is actually optimal.

IV.3 Global elliptic Carleman estimates and applications

As we have seen below, the Carleman inequalities aim at giving **global** weighted estimates of a solution of a PDE (here we shall specifically consider elliptic PDEs) as a function of source terms and of some **partial information** on the solution itself either on a part of the boundary, or on a part of the domain. For a more complete discussion about those kind of estimates (including some insights on the profound reasons why they are true) we refer for instance to [LRL11, Erv17].

IV.3.1 The basic computation

Let Ω be a Lipschitz domain of $\mathbb{R}^d$ and $\varphi \in \mathcal{C}^2(\overline{\Omega}, \mathbb{R})$ be a smooth function to be determined later.

Proposition IV.3.42

For any $u \in \mathcal{C}^2(\overline{\Omega}, \mathbb{R})$, and any $s \geq 0$, we set $v = e^{s\varphi}u$. The following inequality holds

$$\begin{aligned} s^3 \int_\Omega \big(2(D^2\varphi)(\nabla\varphi, \nabla\varphi) - \Delta\varphi|\nabla\varphi|^2\big)|v|^2 + s\int_\Omega \big[2(D^2\varphi)(\nabla v, \nabla v) + \Delta\varphi|\nabla v|^2\big] \\ - s^3\int_{\partial\Omega} |\nabla\varphi|^2\partial_n\varphi|v|^2 - s\int_{\partial\Omega} \partial_n\varphi|\partial_n v|^2 \\ \leq - 2s\int_\Omega v\nabla v\cdot\nabla\Delta\varphi + s^2\int_\Omega |\Delta\varphi|^2|v|^2 \\ - s\int_{\partial\Omega}\partial_n\varphi|\nabla_\| v|^2 - 2s\int_{\partial\Omega}\partial_n v(\nabla_\| v\cdot\nabla_\|\varphi) + 2s\int_{\partial\Omega}\Delta\varphi v\partial_n v \\ + \|e^{s\varphi}(\Delta u)\|^2_{L^2(\Omega)}. \end{aligned}$$

Proof :

We first write the following derivation formulas

$$\nabla e^{s\varphi} = (s\nabla\varphi)e^{s\varphi},$$

$$\Delta e^{s\varphi} = s^2|\nabla\varphi|^2 e^{s\varphi} + s(\Delta\varphi)e^{s\varphi}.$$

Then we set $f = \Delta u$ and we compute

$$\nabla v = e^{s\varphi}(\nabla u) + (\nabla e^{s\varphi})u = e^{s\varphi}(\nabla u) + s\nabla\varphi(e^{s\varphi}u) = e^{s\varphi}(\nabla u) + s(\nabla\varphi)v,$$

$$\Delta v = \Delta(e^{s\varphi}u) = (\Delta e^{s\varphi})u + 2(\nabla e^{s\varphi})\cdot(\nabla u) + e^{s\varphi}(\Delta u),$$

which gives

$$\Delta v = s^2|\nabla\varphi|^2 v + s(\Delta\varphi)v + 2s(\nabla\varphi)\cdot(\nabla v - sv(\nabla\varphi)) + e^{s\varphi}f,$$

and finally

$$\Delta v = -s^2|\nabla\varphi|^2 v + s(\Delta\varphi)v + 2s\nabla\varphi\cdot\nabla v + e^{s\varphi}f. \tag{IV.43}$$

We write this formula in the following form

$$\underbrace{\Big(\Delta v + s^2|\nabla\varphi|^2 v\Big)}_{=M_1 v} + \underbrace{\Big(-2s\nabla\varphi\cdot\nabla v - 2s\Delta\varphi v\Big)}_{=M_2 v} = e^{s\varphi}f - s(\Delta\varphi)v.$$

We write

$$\begin{aligned} 2(M_1v, M_2v)_{L^2} &\leq \|M_1 v\|^2_{L^2} + 2(M_1v, M_2v)_{L^2} + \|M_2v\|^2_{L^2} = \|M_1v + M_2v\|^2_{L^2(\Omega)} \\ &= \|e^{s\varphi}f - s(\Delta\varphi)v\|^2_{L^2} \leq 2\|e^{s\varphi}f\|^2_{L^2} + 2s^2\|(\Delta\varphi)v\|^2_{L^2}. \end{aligned}$$

The two right-hand side terms are the ones we expect in the inequality. Let us now compute the inner product $(M_1v, M_2v)_{L^2}$. We denote by I_{ij} the inner product of the term number i of M_1v with the term number j of M_2v.

- Term I_{11} : We perform two integration by parts

$$\begin{aligned}
I_{11} &= -2s\int_\Omega (\nabla\varphi\cdot\nabla v)\Delta v = -2s\sum_i\int_\Omega \partial_i\varphi\partial_i v\Delta v\\
&= 2s\sum_i\int_\Omega \partial_i\nabla\varphi\cdot\nabla v\partial_i v + 2s\sum_i\int_\Omega \partial_i\varphi\nabla\partial_i v\cdot\nabla v - 2s\int_{\partial\Omega}(\nabla\varphi\cdot\nabla v)\partial_n v\\
&= 2s\int_\Omega D^2\varphi(\nabla v,\nabla v) + s\sum_i\int_\Omega \partial_i\varphi\partial_i\big(|\nabla v|^2\big) - 2s\int_{\partial\Omega}(\nabla\varphi\cdot\nabla v)\partial_n v\\
&= 2s\int_\Omega D^2\varphi(\nabla v,\nabla v) - s\int_\Omega \Delta\varphi|\nabla v|^2 + s\int_{\partial\Omega}\partial_n\varphi|\nabla v|^2 - 2s\int_{\partial\Omega}(\nabla\varphi\cdot\nabla v)\partial_n v.
\end{aligned}$$

- Term I_{12}: We perform one integration by parts

$$\begin{aligned}
I_{12} &= -2s\int_\Omega \Delta\varphi\Delta v v\\
&= 2s\int_\Omega(\Delta\varphi)|\nabla v|^2 + 2s\int_\Omega(\nabla\Delta\varphi\cdot\nabla v)v - 2s\int_{\partial\Omega}\Delta\varphi v\partial_n v.
\end{aligned}$$

- Term I_{21} : We perform one integration by parts

$$\begin{aligned}
I_{21} &= -2s^3\int_\Omega |\nabla\varphi|^2(\nabla\varphi\cdot\nabla v)v\\
&= -s^3\int_\Omega |\nabla\varphi|^2(\nabla\varphi\cdot\nabla)|v|^2\\
&= -s^3\int_\Omega |\nabla\varphi|^2\big(\mathrm{div}(|v|^2\nabla\varphi) - \Delta\varphi|v|^2\big)\\
&= s^3\int_\Omega \nabla\big(|\nabla\varphi|^2\big)\cdot\nabla\varphi|v|^2 - s^3\int_{\partial\Omega}\partial_n\varphi|\nabla\varphi|^2|v|^2 + s^3\int_\Omega(\Delta\varphi)|\nabla\varphi|^2|v|^2\\
&= s^3\int_\Omega \big(2D^2\varphi.(\nabla\varphi,\nabla\varphi) + \Delta\varphi|\nabla\varphi|^2\big)|v|^2 - s^3\int_{\partial\Omega}\partial_n\varphi|\nabla\varphi|^2|v|^2
\end{aligned}$$

- The term I_{22} is left unchanged

$$I_{22} = -2s^3\int_\Omega(\Delta\varphi)|\nabla\varphi|^2|v|^2.$$

Adding all the above terms and gathering all of them lead to the expected inequality. For the boundary terms, we make use of the following formulas

$$|\nabla f|^2 = |\partial_n f|^2 + |\nabla_\| f|^2,$$
$$(\nabla f\cdot\nabla g) = \partial_n f\partial_n g + \nabla_\| f\cdot\nabla_\| g.$$

■

If one wants to get some interesting information from the above huge inequality, we see that first two (volumic) terms in the left-hand side needs to have the good sign, at least on some large enough part of the domain and/or the boundary. More precisely, we would like that, for some $\beta>0$ and some subsets $K\subset\Omega$ and $\Sigma\subset\partial\Omega$, we have

$$2D^2\varphi + \Delta\varphi \text{ is uniformly } \beta\text{-coercive on } K, \tag{IV.44}$$

$$2D^2\varphi(\nabla\varphi,\nabla\varphi) - \Delta\varphi|\nabla\varphi|^2 \geq \beta|\nabla\varphi|^2, \text{ on } K, \tag{IV.45}$$

$$|\nabla\varphi| \geq \beta, \text{ on } K, \tag{IV.46}$$

$$\partial_n\varphi \leq -\beta, \text{ on } \Sigma. \tag{IV.47}$$

Let us point out that we cannot expect those assumptions to be valid all together with $K=\Omega$ and $\Sigma=\partial\Omega$:

- Imagine that assumption (IV.46) holds with $K = \Omega$, then we know that φ has to achieve its maximum on the boundary $\partial\Omega$ which proves that (IV.47) cannot hold for $\Sigma = \partial\Omega$.
- Imagine that (IV.44) holds for $K = \Omega$, then by taking the trace we deduce that

$$(d+2)\Delta\varphi \geq d\beta, \text{ in } \Omega,$$

and thus, by the Stokes formula,

$$\int_{\partial\Omega} \partial_n \varphi = \int_\Omega \Delta\varphi \geq \frac{d}{d+2}\beta|\Omega| > 0,$$

which prevents (IV.47) to be true with $\Sigma = \partial\Omega$.

Therefore, we will need to relax our requirements on K and Σ and that will lead to the observation terms in the final Carleman estimate.

More precisely, it is possible to build suitable weight functions as stated in the following result whose proof is postponed to Section IV.3.4.

Lemma IV.3.43

1. **Boundary observation :** *Let $\Gamma \subset \partial\Omega$. There exists a $\beta > 0$ and a function φ satisfying (IV.44), (IV.45) and (IV.46) with $K = \Omega$ and (IV.47) with $\Sigma = \partial\Omega \setminus \Gamma$.*

   Moreover, we can choose φ that satisfies

$$\nabla_{\parallel}\varphi = 0, \quad \textit{on } \partial\Omega.$$

2. **Interior observation :** *Let $\omega \subset \Omega$ a non empty open subset of Ω. There exists a $\beta > 0$ and a function φ satisfying (IV.44), (IV.45) and (IV.46) with $K = \Omega \setminus \omega$, and (IV.47) with $\Sigma = \partial\Omega$.*

   Moreover, we can choose φ that satisfies

$$\nabla_{\parallel}\varphi = 0, \quad \textit{on } \partial\Omega.$$

IV.3.2 Proof of the boundary Carleman estimate

We may now prove Theorem IV.1.28. For the moment we shall not use the fact that v satisfies any boundary condition in order to identify the precise point where this property will be used.

We take a function φ associated with Γ, as in the first point of Lemma IV.3.43.

We apply the inequality of Proposition IV.3.42 with this particular function φ using its properties to get

$$\begin{aligned} s^3\beta^3 \int_\Omega |v|^2 + s\beta \int_\Omega |\nabla v|^2 + s^3\beta^3 \int_{\partial\Omega\setminus\Gamma} |v|^2 + s\beta \int_{\partial\Omega\setminus\Gamma} |\partial_n v|^2 \\ \leq \|\nabla\varphi\|_\infty^3 s^3 \int_\Gamma |v|^2 + s\|\nabla\varphi\|_\infty \int_\Gamma |\partial_n v|^2 + s\|\nabla\varphi\|_\infty \int_{\partial\Omega} |\nabla_\parallel v|^2 + 2s\|\Delta\varphi\|_{L^\infty} \int_{\partial\Omega} |v||\partial_n v| \\ + 2\|e^{s\varphi}(\Delta u)\|_{L^2(\Omega)}^2 - 2s\int_\Omega v\nabla v\cdot\nabla\Delta\varphi + 2s^2\int_\Omega |\Delta\varphi|^2|v|^2. \end{aligned}$$

Adding the terms $s^3\beta^3 \int_\Gamma |v|^2$ and $s\beta \int_\Gamma |\partial_n v|^2$ on both sides of the inequality gives

$$\begin{aligned} s^3\beta^3 \int_\Omega |v|^2 + s\beta \int_\Omega |\nabla v|^2 + s^3\beta^3 \int_{\partial\Omega} |v|^2 + s\beta \int_{\partial\Omega} |\partial_n v|^2 \\ \leq 2\|\nabla\varphi\|_\infty^3 s^3 \int_\Gamma |v|^2 + 2s\|\nabla\varphi\|_\infty \int_\Gamma |\partial_n v|^2 + s\|\nabla\varphi\|_\infty \int_{\partial\Omega} |\nabla_\parallel v|^2 + 2s\|\Delta\varphi\|_{L^\infty} \int_{\partial\Omega} |v||\partial_n v| \\ + 2\|e^{s\varphi}(\Delta u)\|_{L^2(\Omega)}^2 - 2s\int_\Omega v\nabla v\cdot\nabla\Delta\varphi + 2s^2\int_\Omega |\Delta\varphi|^2|v|^2. \end{aligned}$$

We see that the left-hand side terms give global information on v and ∇v in Ω and on v and $\partial_n v$ on $\partial\Omega$.

The last two terms can be bounded as follows

$$-2s\int_\Omega v\nabla v\cdot\nabla\Delta\varphi + 2s^2\int_\Omega |\Delta\varphi|^2|v|^2 \leq C_\varphi s\|v\|_{L^2}\|\nabla v\|_{L^2} + C_\varphi s^2\|v\|_{L^2}$$
$$\leq C_\varphi s^2\|v\|_{L^2}^2 + C_\varphi\|\nabla v\|_{L^2}^2.$$

We observe that the powers of s in those terms are less than the powers of s on similar terms in the left-hand side of the inequality. Therefore, there exists a $s_0 > 0$ depending only on φ, such that those terms can be absorbed in the inequality. We get

$$s^3\beta^3\int_\Omega |v|^2 + s\beta\int_\Omega|\nabla v|^2 + s^3\beta^3\int_{\partial\Omega}|v|^2 + s\beta\int_{\partial\Omega}|\partial_n v|^2$$
$$\leq C_\varphi s^3\int_\Gamma |v|^2 + C_\varphi s\int_\Gamma|\partial_n v|^2 + C_\varphi s\int_{\partial\Omega}|\nabla_\parallel v|^2 + C_\varphi s\int_{\partial\Omega}|v||\partial_n v| + 2\|e^{s\varphi}(\Delta u)\|^2_{L^2(\Omega)}.$$

The fourth term in the right-hand side can be estimated by using the Cauchy-Schwarz and Young inequalities as follows

$$C_\varphi s\int_{\partial\Omega}|v||\partial_n v| \leq \tilde{C}_\varphi s^2\int_{\partial\Omega}|v|^2 + \tilde{C}_\varphi\int_{\partial\Omega}|\partial_n v|^2.$$

It follows (thanks to the low powers in s of those terms) that, for s large enough, we can absorb those contributions by the left-hand side terms in our inequality.

It remains the following inequality

$$s^3\beta^3\int_\Omega |v|^2 + s\beta\int_\Omega|\nabla v|^2 + s^3\beta^3\int_{\partial\Omega}|v|^2 + s\beta\int_{\partial\Omega}|\partial_n v|^2$$
$$\leq C_\varphi s^3\int_\Gamma |v|^2 + C_\varphi s\int_\Gamma|\partial_n v|^2 + C_\varphi s\int_{\partial\Omega}|\nabla_\parallel v|^2 + 2\|e^{s\varphi}(\Delta u)\|^2_{L^2(\Omega)},$$

which is valid for any function u without any assumption on the boundary conditions.

The only term which is not an observation term is the third one in the right-hand side. At that point, we need to consider the boundary condition for u. Indeed, if we assume that $u = 0$ (or equivalently $v = 0$) on $\partial\Omega\setminus\overline{\Gamma}$, we deduce that $\nabla_\parallel v = 0$ on $\partial\Omega\setminus\overline{\Gamma}$ and thus we have

$$s^3\beta^3\int_\Omega |v|^2 + s\beta\int_\Omega|\nabla v|^2 + s\beta\int_{\partial\Omega}|\partial_n v|^2$$
$$\leq C_\varphi s^3\int_\Gamma |v|^2 + C_\varphi s\int_\Gamma|\partial_n v|^2 + C_\varphi s\int_\Gamma|\nabla_\parallel v|^2 + 2\|e^{s\varphi}(\Delta u)\|^2_{L^2(\Omega)},$$

which is a first suitable Carleman estimate with observation on Γ.

The announced estimate is a particular case of the above inequality in the case where $v = 0$ on the whole boundary $\partial\Omega$ (and thus $\nabla_\parallel v = 0$)

$$s^3\beta^3\int_\Omega |v|^2 + s\beta\int_\Omega|\nabla v|^2 + s\beta\int_{\partial\Omega}|\partial_n v|^2 \leq C_\varphi s\int_\Gamma|\partial_n v|^2 + 2\|e^{s\varphi}(\Delta u)\|^2_{L^2(\Omega)}.$$

We just finally need to go back to the function u. We first note that

$$|v| = e^{s\varphi}|u|,$$

and

$$\nabla v = e^{s\varphi}(\nabla u) + (\nabla e^{s\varphi})u = e^{s\varphi}(\nabla u) + s(\nabla\varphi)\underbrace{e^{s\varphi}u}_{=v},$$

so that we have

$$s|e^{s\varphi}\nabla u|^2 \leq s|\nabla v|^2 + s^3|\nabla\varphi|^2|v|^2.$$

Moreover,

$$\partial_n v = e^{s\varphi}(\partial_n u) + u(\partial_n e^{s\varphi}) = e^{s\varphi}(\partial_n u),$$

since $u = 0$ on the boundary. The claim is proved.

IV.3.3 Proof of the distributed Carleman estimate

We may now prove Theorem IV.1.29. We take a function φ associated with ω, as in the second point of Lemma IV.3.43.

We apply the inequality of Proposition IV.3.42 with this particular function φ using its properties to get, for any function v that vanishes on the boundary

$$\beta^3 s^3 \int_{\Omega\setminus\omega} |v|^2 + s\beta \int_{\Omega\setminus\omega} |\nabla v|^2 + s\beta \int_{\partial\Omega} |\partial_n v|^2 \leq C_\varphi s^3 \int_\omega |v|^2 + C_\varphi s \int_\omega |\nabla v|^2 + 2\|e^{s\varphi}(\Delta u)\|^2_{L^2(\Omega)} + 2s^2 \int_\Omega |\Delta\varphi|^2 |v|^2 - 2s \int_\Omega v\nabla v \cdot \nabla\Delta\varphi$$

Adding the terms $s^3\beta^3 \int_\omega |v|^2$ and $s\beta \int_\omega |\nabla v|^2$ on both sides of the inequality gives (with another value of the constant C_φ)

$$\beta^3 s^3 \int_{\Omega} |v|^2 + s\beta \int_{\Omega} |\nabla v|^2 + s\beta \int_{\partial\Omega} |\partial_n v|^2 \leq C_\varphi s^3 \int_\omega |v|^2 + C_\varphi s \int_\omega |\nabla v|^2 + 2\|e^{s\varphi}(\Delta u)\|^2_{L^2(\Omega)} + 2s^2 \int_\Omega |\Delta\varphi|^2 |v|^2 - 2s \int_\Omega v\nabla v \cdot \nabla\Delta\varphi,$$

and we can now absorb the last two terms as we did previously, by assuming that $s \geq s_0$ for some s_0 depending only on the weight function φ. We finally get

$$\beta^3 s^3 \int_{\Omega} |v|^2 + s\beta \int_{\Omega} |\nabla v|^2 + s\beta \int_{\partial\Omega} |\partial_n v|^2 \leq C_\varphi s^3 \int_\omega |v|^2 + C_\varphi s \int_\omega |\nabla v|^2 + 2\|e^{s\varphi}(\Delta u)\|^2_{L^2(\Omega)}$$

This is actually a Carleman estimate with observation terms in ω but we would like a little bit more, namely to obtain a similar estimate without observation terms containing derivatives of v. Let us show how to obtain such an estimate.

To begin with we consider a small non-empty observation domain ω_0 such that $\overline{\omega_0} \subset \omega$ and we apply the above Carleman estimate to this new observation domain (this imply to use a weight function φ adapted to this new observation domain). It follows that

$$\beta^3 s^3 \int_{\Omega} |v|^2 + s\beta \int_{\Omega} |\nabla v|^2 + s\beta \int_{\partial\Omega} |\partial_n v|^2 \leq C s^3 \int_{\omega_0} |v|^2 + C s \int_{\omega_0} |\nabla v|^2 + 2\|e^{s\varphi}(\Delta u)\|^2_{L^2(\Omega)},$$

and we will now show how to get rid of the term $\int_{\omega_0} |\nabla v|^2$. Let η be a non-negative smooth function compactly supported in ω and such that $\eta = 1$ in ω_0. We write by an integration by parts

$$s \int_{\omega_0} |\nabla v|^2 \leq s \int_\omega \eta |\nabla v|^2 = -s \int_\omega v\nabla v \cdot \nabla\eta - s \int_\omega \eta v(\Delta v).$$

Then we use the equation satisfied by v (see (IV.43)) that we recall here

$$\Delta v = e^{s\varphi}(\Delta u) + s(\Delta\varphi)v - s^2|\nabla\varphi|^2 v + 2s\nabla\varphi \cdot \nabla v,$$

to obtain

$$s \int_{\omega_0} |\nabla v|^2 \leq C_\varphi \left(s \int_\omega |v||\nabla v| + s \int_\omega |v| e^{s\varphi} |\Delta u| + s^2 \int_\omega |v|^2 + s^3 \int_\omega |v|^2 + s^2 \int_\omega |v||\nabla v| \right).$$

Since $s \geq s_0$, we deduce

$$s \int_{\omega_0} |\nabla v|^2 \leq C_\varphi \left(s^2 \int_\omega |v||\nabla v| + s \int_\omega |v| e^{s\varphi} |\Delta u| + s^3 \int_\omega |v|^2 \right).$$

The last term is the observation term we would like to keep at the end. The second term can be bounded by the Cauchy-Schwarz and Young inequalities

$$s\int_\omega |v|e^{s\varphi}|\Delta u| \leq 2s^2\int_\omega |v|^2 + 2\int_\omega |e^{s\varphi}(\Delta u)|^2 \leq 2s^2\int_\omega |v|^2 + 2\|e^{s\varphi}(\Delta u)\|^2_{L^2(\Omega)}.$$

Finally, we also use the Cauchy-Schwarz inequality and the refined Young inequality to bound the first term as follows

$$s^2\int_\omega |v||\nabla v| = \int_\omega s^{3/2}|v|s^{1/2}|\nabla v| \leq \frac{\varepsilon}{2}s\int_\omega |\nabla v|^2 + \frac{1}{2\varepsilon}s^3\int_\omega |v|^2 \leq \frac{\varepsilon}{2}s\int_\Omega |\nabla v|^2 + \frac{1}{2\varepsilon}s^3\int_\omega |v|^2,$$

so that we can take ε small enough (depending only on φ) such that the term in ∇v is absorbed by the corresponding term in the left-hand side of the inequality. The proof is complete.

IV.3.4 Construction of the weight functions

Our goal is to prove Lemma IV.3.43. We begin by constructing a first function with particular properties.

Lemma IV.3.44

Let U be a bounded domain of $\mathbb{R}^d$ of class $\mathcal{C}^2$ and $V \subset U$ a non empty open subset of U. There exists a function $\psi \in \mathcal{C}^2(\overline{U})$ such that:

- *$\psi = d(., \partial U)$ in a neighborhood of ∂U. In particular $\psi = 0$ and $\partial_n \psi = -1$ on ∂U.*
- *$\psi > 0$ in U.*
- *$\nabla\psi \neq 0$ in the compact $K \stackrel{\text{def}}{=} \overline{U}\setminus V$. In particular, there exists $\alpha > 0$ such that*

$$|\nabla\psi| \geq \alpha, \text{ in } K.$$

Proof :

Using the Morse lemma, we can find a function $\tilde{\psi}$ that satisfies the first two properties and which has a finite number of critical points in U, let say $x_1, ..., x_n$, see for instance [TW09]. Then we choose n distinct points $y_1, ..., y_n$ in V. There exists a diffeomorphism G from U into itself such that $G(y_i) = x_i$ and such that $G(y) = y$ in a neighborhood of ∂U. This can be done by considering the flow of a suitable compactly supported vector field. We easily check that $\psi = \tilde{\psi}\circ G$ satisfies all the required properties. ■

We may now prove the second point of Lemma IV.3.43. We apply the previous lemma with $U = \Omega$ and $V = \omega$. We set $\varphi = e^{\lambda\psi}$ for $\lambda \geq 0$. and perform the following computations

$$\nabla\varphi = \lambda(\nabla\psi)\varphi,$$

$$D^2\varphi = \lambda(D^2\psi)\varphi + \lambda^2(\nabla\psi)\otimes(\nabla\psi)\varphi,$$

$$\Delta\varphi = \lambda(\Delta\psi)\varphi + \lambda^2|\nabla\psi|^2\varphi.$$

- We first compute

$$2D^2\varphi + \Delta\varphi = \lambda\big(2(D^2\psi) + (\Delta\psi)\big)\varphi + \lambda^2\big(2(\nabla\psi)\otimes(\nabla\psi) + |\nabla\psi|^2\big)\varphi,$$

and we see that for any $\xi \in \mathbb{R}^d$

$$\begin{aligned}\frac{1}{\varphi}(2D^2\varphi + \Delta\varphi).(\xi,\xi) &\geq \lambda^2(2|\nabla\psi\cdot\xi|^2 + |\nabla\psi|^2|\xi|^2) - \lambda C_\psi|\xi|^2\\ &\geq (\lambda^2|\nabla\psi|^2 - \lambda C_\psi)|\xi|^2.\end{aligned}$$

Therefore, since $\nabla\psi$ does not vanish in K, we can choose λ large enough so that

$$\frac{1}{\varphi}(2D^2\varphi + \Delta\varphi).(\xi,\xi) \geq C\lambda^2|\nabla\psi|^2|\xi|^2, \text{ in } K,$$

and since $\varphi \geq 1$, we get

$$2D^2\varphi + \Delta\varphi \geq C\lambda^2|\nabla\psi|^2, \text{ in } K.$$

- We compute now

$$\begin{aligned} 2D^2\varphi.(\nabla\varphi,\nabla\varphi) - \Delta\varphi|\nabla\varphi|^2 &= \lambda^2\varphi^2\big(2D^2\varphi.(\nabla\psi,\nabla\psi) - \Delta\varphi|\nabla\psi|^2\big) \\ &= \lambda^2\varphi^2\big(\lambda^2|\nabla\psi|^4\varphi + 2\lambda D^2\psi.(\nabla\psi,\nabla\psi)\varphi - \lambda(\Delta\psi)|\nabla\psi|^2\varphi\big) \\ &\geq \phi^3(\lambda^4\alpha^4 - C_\psi\lambda^3), \text{ in } K. \end{aligned}$$

Here also, for λ large enough we deduce that

$$2D^2\varphi.(\nabla\varphi,\nabla\varphi) - \Delta\varphi|\nabla\varphi|^2 \geq \lambda^4\alpha^4, \text{ in } K.$$

Let us now prove the first point of Lemma IV.3.43. To this end, we consider a bounded open set U that contains Ω and such that $\partial\Omega \cap U \subset \Gamma$. Then we choose some non empty open subset V such that $\overline{V} \cap \overline{\Omega} = \emptyset$.

We build a function φ related with this choice of U and V, and we easily see that its restriction to Ω satisfies all the required properties since

$$\partial\Omega \setminus \Gamma \subset \partial U.$$

IV.3.5 A Carleman estimate for augmented elliptic operators with special boundary conditions

For $T^* > 0$, we set $Q = (0,T^*) \times \Omega$ be a *time-space* domain (even though the time variable here has nothing to do with the physical time of the initial problem). We consider the augmented elliptic operator

$$\Delta_{\tau,x} \stackrel{\text{def}}{=} \partial_\tau^2 + \Delta,$$

where the operator Δ (as well as ∇) only concerns the space variables. The complete gradient operator will be denoted by

$$\nabla_{\tau,x} \stackrel{\text{def}}{=} (\partial_\tau, \nabla).$$

Note that all the analysis below still apply with Δ replaced by the general elliptic operator $-\mathcal{A}$, with suitable regularity assumptions on γ.

Lemma IV.3.45

Let $\omega \subset \Omega$ be a non-empty open subset of Ω. There exists a weight function $\varphi \in \mathcal{C}^2(\overline{Q})$ that satisfies the assumptions (IV.44), (IV.45) and (IV.46) on the time-space domain Q and moreover

$$\begin{aligned} \partial_n\varphi < 0, &\quad \textit{on } (0,T^*) \times \partial\Omega, \\ (-\partial_\tau\varphi) \leq -\beta, &\quad \textit{on } \{0\} \times (\Omega \setminus \omega), \\ \partial_\tau\varphi \leq -\beta, &\quad \textit{on } \{T^*\} \times \Omega, \\ \nabla_x\varphi(T^*,.) = 0, &\quad \textit{in } \Omega. \end{aligned}$$

We use this function φ in Proposition IV.3.42 on the domain Q for any function u that satisfies

$$\begin{cases} u(0,.) = 0, & \text{in } \Omega, \\ u(\tau,.) = 0, & \text{on } \partial\Omega \text{ for any } \tau \in (0,T^*). \end{cases}$$

Observe that u does not vanish for $\tau = T^*$ so that u does not satisfy homogeneous boundary condition on ∂Q. This is why the Carleman estimate we will prove is different from the one developed above.

We obtain

$$s^3\beta^3 \int_Q |v|^2 + s\beta \int_Q |\nabla_{\tau,x} v|^2 + s^3\beta^3 \int_\Omega |v(T^*,.)|^2 + \beta s \int_\Omega |\partial_\tau v(T^*,.)|^2 + \beta s \int_{\Omega\setminus\omega} |\partial_\tau v(0,.)|^2$$
$$\leq -s \int_\Omega \partial_\tau\varphi(T^*,.)|\nabla_x v(T^*,.)|^2 + 2\|e^{s\varphi}(\Delta_{\tau,x} u)\|^2_{L^2(Q)}$$
$$-2s \int_Q v\nabla_{\tau,x} v \cdot \nabla_{\tau,x}\Delta_{\tau,x}\varphi + 2s^2 \int_Q |\Delta_{\tau,x}\varphi|^2|v|^2.$$

The last two terms can be asborbed for $s \geq s_0$ as before, and we can add the observation term at time $\tau = 0$ on ω on both sides of the inequality to obtain

$$s^3\beta^3 \int_Q |v|^2 + s\beta \int_Q |\nabla_{\tau,x} v|^2 + s^3\beta^3 \int_\Omega |v(T^*,.)|^2 + \beta s \int_\Omega |\partial_\tau v(T^*,.)|^2 + \beta s \int_\Omega |\partial_\tau v(0,.)|^2$$
$$\leq Cs \int_\omega |\partial_\tau v(0,.)|^2 + Cs \int_\Omega |\nabla_x v(T^*,.)|^2 + C\|e^{s\varphi}(\Delta_{\tau,x} u)\|^2_{L^2(Q)}.$$

Coming back to the function u, and using that φ does not depend on x at $\tau = T^*$, we have finally obtained the following Carleman estimate.

Proposition IV.3.46

For any $s \geq s_1$, any $u \in \mathcal{C}^2(\overline{Q})$ such that $u(0,.) = 0$ and $u(t,.) = 0$ on $\partial\Omega$ for any $t \in (0,T^*)$, we have

$$s^3 \int_Q |e^{s\varphi}u|^2 + s \int_Q |e^{s\varphi}\nabla_{\tau,x} u|^2 + s \int_\Omega |e^{s\varphi(0,.)}\partial_\tau u(0,.)|^2$$
$$+ s^3 e^{2s\varphi(T^*)} \int_\Omega |u(T^*,.)|^2 + s e^{2s\varphi(T^*)} \int_\Omega |\partial_\tau u(T^*,.)|^2$$
$$\leq Cs \int_\omega |e^{s\varphi(0,.)}\partial_\tau u(0,.)|^2 + Cs e^{2s\varphi(T^*)} \int_\Omega |\nabla_x u(T^*,.)|^2 + C\|e^{s\varphi}(\Delta_{\tau,x} u)\|^2_{L^2(Q)}.$$

Remark IV.3.47

All the above elliptic Carleman estimates can be adapted to more general differential operators, like $-$div $(\gamma\nabla\cdot)$ for a smooth enough diffusion coefficient γ (and even for in some non-smooth cases).

IV.4 The Fursikov-Imanuvilov approach

Contrary to the Lebeau-Robbiano strategy that amounts to build, step by step, a null-control for our problem, the method proposed by Fursikov and Imanuvilov in [FI96] consists in directly proving the observability inequality on the adjoint problem.

IV.4.1 Global parabolic Carleman estimates

We shall derive and use now a new kind of Carleman estimates. Those inequalities will directly concern the solutions of the parabolic operator under study.

The control time $T > 0$ is fixed and we set $\theta(t) = \frac{1}{t(T-t)}$. We give the following result without proof (see [FI96], [Cor07] or [TW09]) since it follows very similar lines as the ones of the proof of the elliptic Carleman estimate (but with more technicalities).

Theorem IV.4.48

Let ω be a non empty open subset of Ω. There exists a function $\varphi \in \mathcal{C}^2(\overline{\Omega})$ such that

$$\sup_{\Omega} \varphi < 0, \text{ and } \inf_{\Omega\setminus\omega} |\nabla \varphi| > 0,$$

and for which we have the following property: for any $d \in \mathbb{R}$, there exists $s_0 > 0$ and $C > 0$ such that the following estimate holds for any $s \geq s_0$ and any $u \in \mathcal{C}^2([0,T] \times \overline{\Omega})$ such that $u = 0$ on $(0,T) \times \partial\Omega$

$$\int_0^T \int_\Omega (s\theta)^d \left|e^{s\theta\varphi} u\right|^2 + \int_0^T \int_\Omega (s\theta)^{d-2} \left|e^{s\theta\varphi} \nabla u\right|^2 \\ \leq C \left(\int_0^T \int_\omega (s\theta)^d \left|e^{s\theta\varphi} u\right|^2 + \int_0^T \int_\Omega (s\theta)^{d-3} \left|e^{s\theta\varphi} (\partial_t u \pm \Delta u)\right|^2 \right).$$

The sign $\pm$ in the parabolic operator just means that the estimate holds true for both operators $\partial_t - \Delta$ and $\partial_t + \Delta$.

As usual we can extend, by density, this estimate to less regular functions u as soon as all the terms in the inequality make sense.

Remark IV.4.49

A careful inspection of the proof shows that the same estimate holds with the following additional terms in the left-hand side

$$\int_0^T \int_\Omega (s\theta)^{d-4} \left|e^{s\theta\varphi} \partial_t u\right|^2 + \int_0^T \int_\Omega (s\theta)^{d-4} \left|e^{s\theta\varphi} \Delta u\right|^2.$$

Notice that, since φ is negative and $\theta(t) \to \infty$ when $t \to 0$ or $t \to T$, all the weights in this estimate are exponentially small near $t = 0$ and $t = T$. This explains why the estimate holds without any assumption on the values of u at time $t = 0$ or $t = T$.

IV.4.2 Another proof of the null-controllability of the heat equation

With the above estimate at hand, we can directly prove the observability inequality we need.

Theorem IV.4.50

With the same assumption as before, there exists $C > 0$ such that, for any solution q of the adjoint problem

$$-\partial_t q - \Delta q = 0,$$

with $q(T) \in L^2(\Omega)$, then we have

$$\|q(0)\|^2_{L^2(\Omega)} \leq C^2 \int_0^T \int_\omega |q(t,x)|^2 \, dt \, dx.$$

As a consequence, we have proved the null-controllability of the heat equation for any time $T > 0$.

Proof :

We choose $d = 0$ and take some $s \geq s_0$; then we apply the Carleman estimate above to the function q. Only keeping the first term in the left-hand side, we get

$$\int_0^T \int_\Omega \left| e^{s\theta\varphi} q \right|^2 \leq C \int_0^T \int_\omega \left| e^{s\theta\varphi} q \right|^2.$$

Since $\varphi < 0$ and $\theta > 0$, we easily see that $e^{s\theta\varphi} \leq 1$. Moreover, we restrict the left-hand side integral to the time interval $(T/4, 3T/4)$ to get

$$\int_{\frac{T}{4}}^{\frac{3T}{4}} \int_\Omega \left| e^{s\theta\varphi} q \right|^2 \leq C \int_0^T \int_\omega |q|^2.$$

On the interval $(T/4, 3T/4)$ we have $\theta(t) \leq 16/3T^2$. We deduce that

$$e^{2s\varphi} \geq e^{32/3T^2 \inf \varphi}, \text{ on } (T/4, 3T/4) \times \Omega.$$

We have thus obtained for another value of C

$$\int_{\frac{T}{4}}^{\frac{3T}{4}} \int_\Omega |q|^2 \leq C \int_0^T \int_\omega |q|^2.$$

We use now the dissipation property of the (backward) heat equation which gives

$$\|q(0)\|_{L^2}^2 \leq \|q(s)\|_{L^2(\Omega)}^2, \ \forall s \in (0, T).$$

By integration on $(T/4, 3T/4)$ we get

$$\|q(0)\|_{L^2}^2 \leq \frac{2}{T} \int_{\frac{T}{4}}^{\frac{3T}{4}} \|q(s)\|_{L^2(\Omega)}^2,$$

and the claim is proved by combining the last two inequalities. ■

Chapter V

Coupled parabolic equations

In this chapter, we would like to investigate controllability properties for coupled systems like (III.3) and (III.4). A particular attention will be paid to the case where $\text{rank}B < n$, that is when there are less controls than components in the system. We refer to the survey paper [AKBGBT11] even though many results were published on this topic after this survey.

V.1 Systems with as many controls as components

Let us first discuss the case where $\text{rank}B = n$ (which implies that $m \geq n$). We can remove some (useless) columns to B and assume that $m = n$ and that B is invertible.

Theorem V.1.1

Let ω be a non empty open subset of Ω and $T > 0$ and assume that B is a square invertible $n \times n$ matrix. Then, System (III.3) is null-controllable at time T.

Notice that we do not make any structure assumption on the coupling matrix $C(t, x)$, we only assume that $C \in L^\infty((0, T) \times \Omega)$.

Proof :

We propose a proof based on the global parabolic Carleman estimate. The adjoint system associated with (III.3) reads

$$-\partial_t q - \Delta q + C^*(t, x)q = 0,$$

which can be also written, component-by-component for any $i \in \{1, ..., n\}$, as follows

$$-\partial_t q_i - \Delta q_i = -\sum_j c_{ji}(t, x)q_j.$$

We apply to each q_i the Carleman estimate given in Theorem IV.4.48, with $d = 0$, the same value of $s \geq s_0$ and, of course, the same weight function φ. It follows that

$$\int_0^T \int_\Omega \left|e^{s\theta\varphi} q_i\right|^2 \leq C \int_0^T \int_\omega \left|e^{s\theta\varphi} q_i\right|^2 + C \sum_j \int_0^T \int_\Omega (s\theta)^{-3} |e^{s\theta\varphi} q_j|^2.$$

We sum over i all those inequalities and we observe that on $(0, T)$, the function θ^{-3} is bounded to deduce that, for all $s \geq s_0$

$$\sum_i \int_0^T \int_\Omega \left|e^{s\theta\varphi} q_i\right|^2 \leq C \sum_i \int_0^T \int_\omega \left|e^{s\theta\varphi} q_i\right|^2 + \frac{C}{s^3} \sum_j \int_0^T \int_\Omega |e^{s\theta\varphi} q_j|^2.$$

We see that, for s large enough (depending only on the data !), the last term is absorbed by the left-hand side term. We deduce that

$$\sum_i \int_0^T \int_\Omega \left|e^{s\theta\varphi} q_i\right|^2 \leq C \sum_i \int_0^T \int_\omega \left|e^{s\theta\varphi} q_i\right|^2.$$

Using the same arguments as in Theorem IV.4.50, we arrive at

$$\sum_i \int_{\frac{T}{4}}^{\frac{3T}{4}} \int_\Omega |q_i|^2 \leq C \sum_i \int_0^T \int_\omega |q_i|^2.$$

Still denoting by $|.|$ the Euclidean norm in $\mathbb{R}^n$, this reads

$$\int_{\frac{T}{4}}^{\frac{3T}{4}} \int_\Omega |q|^2 \leq C \int_0^T \int_\omega |q|^2.$$

We use now the fact that B is an invertible matrix to deduce that for some other constant C, we have

$$\int_{\frac{T}{4}}^{\frac{3T}{4}} \int_\Omega |q|^2 \leq C \int_0^T \int_\omega |B^* q|^2. \tag{V.1}$$

We would like now to use the dissipation argument. Because of the coupling terms we cannot simply use the estimate (IV.5) for the heat equation. Instead we will prove an energy estimate for the backward equation which implies that $\|q(0)\|_{L^2(\Omega)}$ can be bounded, up to a multiplicative constant, by $\|q(s)\|_{L^2(\Omega)}$ for any $s \geq 0$.

To this end we multiply the adjoint equation (in the sense of the Euclidean inner product of $\mathbb{R}^n$) by $q(t,x)$ and we integrate over Ω. It follows that

$$-\int_\Omega (\partial_t q)\cdot q\,dx - \int_\Omega \Delta q \cdot q\,dx = -\int_\Omega (C^* q)\cdot q\,dx.$$

Integrating by parts the second term it follows that

$$-\frac{1}{2}\frac{d}{dt}\int_\Omega |q|^2\,dx + \int_\Omega |\nabla q|^2\,dx = -\int_\Omega (C^* q)\cdot q\,dx \leq \|C\|_{L^\infty} \int_\Omega |q|^2\,dx,$$

in particular we have

$$-\frac{d}{dt}\|q(t)\|^2_{L^2(\Omega)} \leq 2\|C\|_{L^\infty} \|q(t)\|^2_{L^2(\Omega)}.$$

Using the Gronwall inequality we deduce that

$$\|q(t)\|_{L^2(\Omega)} \leq e^{(s-t)\|C\|_{L^\infty}} \|q(s)\|_{L^2(\Omega)}, \quad \forall 0 \leq t < s \leq T,$$

and in particular

$$\|q(0)\|_{L^2(\Omega)} \leq e^{T\|C\|_{L^\infty}} \|q(s)\|_{L^2(\Omega)}, \quad \forall 0 \leq s \leq T.$$

Combining this inequality with (V.1) we obtain

$$\|q(0)\|^2_{L^2(\Omega)} \leq C \int_0^T \int_\omega |B^* q|^2,$$

and the observability inequality is proved as well as the null-controllability by duality. ∎

V.2 Boundary versus distributed controllability

We first notice that, for the scalar problems we have studied before, the boundary and distributed controllability problems are in fact equivalent in some sense.

- Distributed controllability $\Rightarrow$ Boundary controllability:

  Imagine that you are able to prove the null-controllability for our system for any choice of Ω and ω, then we can prove the boundary controllability by considering an extended domain $\widetilde{\Omega}$ that contains Ω and which is built in such a way that $\overline{\Omega} \cap \widetilde{\Omega} \subset \Gamma_0$ (see Figure V.1). Then we choose a region $\omega \subset \widetilde{\Omega} \setminus \overline{\Omega}$.

  We then extend our initial data y_0 to the whole domain $\widetilde{\Omega}$ and apply the controllability result with control supported in ω on the new extended problem, let $\tilde{y} \in \mathcal{C}^0([0,T], L^2(\widetilde{\Omega}))$ be the corresponding controlled solution. Since $\omega \cap \Omega = \emptyset$, we see that the restriction of$\tilde{y}$ on Ω, $y = \tilde{y}_{|\Omega}$ satisfies the heat equation (without source term) in Ω. Moreover, since $\tilde{y}$ vanishes on $\partial\widetilde{\Omega}$ we see in particular that y vanishes on $\partial\Omega \setminus \Gamma_0$ by construction of the extended domain $\widetilde{\Omega}$.

  It remains to set $v = \tilde{y}_{|\Gamma_0}$ in the trace sense, which is an element of $L^2(0,T;H^{\frac{1}{2}}(\Gamma_0))$ which is an admissible boundary control for the original problem.

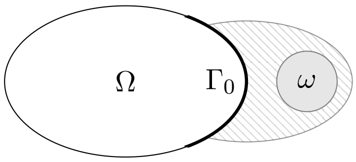

Figure V.1: Distributed controllability implies boundary controllability

- Boundary controllability $\Rightarrow$ Distributed controllability:

  A similar reasoning shows that the converse implication is true, see Figure V.2.

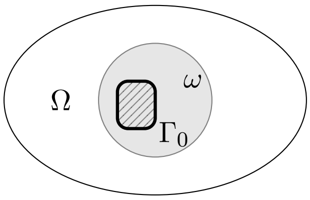

Figure V.2: Boundary controllability implies distributed controllability

The same arguments show that boundary and distributed controllability are equivalent problems in the case where $m = \mathrm{rankB} = n$.

However, in the sequel of this chapter we shall consider coupled parabolic systems with less controls than components in the system $m < n$. One can easily see that, in this case, the above reasoning does not hold anymore and in fact we will see that the boundary and distributed controllability systems may really present different behaviors.

V.3 Distributed control problems

V.3.1 Constant coefficient systems with less controls than equations

In this section we assume that $C(t,x)$ is a constant matrix C, that $m = \mathrm{rank}B < n$.

Proposition V.3.2

A necessary condition for the null- or approximate- controllability for (III.3) is that the pair (C,B) is controllable.

Proof :

Let y be any solution of (III.3) and ϕ_λ an eigenfunction of the Laplace operator associated with the eigenvalue λ. We deduce that the quantity

$$z(t) \stackrel{\text{def}}{=} \langle y(t), \phi_\lambda \rangle_{L^2} \in \mathbb{R}^n,$$

solves the following equation

$$\frac{d}{dt}z + \lambda z + Cz = Bv_\lambda(t), \tag{V.2}$$

where $v_\lambda(t) = \langle v(t,.), 1_\omega \phi_\lambda \rangle_{L^2} \in \mathbb{R}^m$. Then, the controllability of (III.3) implies the one of (V.2), which itself implies that the pair $(C + \lambda \text{Id}, B)$ is controllable and so is the pair (C, B). ■

Theorem V.3.3

Under the above assumptions and if we assume that the pair (C, B) is controllable, then the system (III.3) is approximately controllable for any time $T > 0$.

Proof :

The adjoint system reads

$$-\partial_t q - \Delta q + C^* q = 0.$$

Each eigenvalue of $-\Delta + C^*$ is of the form $\lambda = \sigma + \mu$ where $\sigma \in \text{Sp}(-\Delta)$ and $\mu \in \text{Sp}(C^*)$ and any element in Ker $((-\Delta + C^*) - \lambda)$ can be written

$$\Phi_\lambda = \sum_{\substack{\sigma \in \text{Sp}(-\Delta) \\ \mu \in \text{Sp}(C^*) \\ \lambda = \sigma + \mu}} \sum_{i=1}^{n_\sigma} v_{\sigma,i}(x) \Phi_{\mu,i},$$

where $(v_{\sigma,i})_{1 \leq i \leq n_\sigma}$ is an orthonormal family of Ker $(-\Delta - \sigma)$ and $(\Phi_{\mu,i})_{1 \leq i \leq n_\sigma} \subset$ Ker $(C^* - \mu)$.

When we apply the observation operator $\mathcal{B}^* = 1_\omega B^*$, we obtain

$$\mathcal{B}^* \Phi_\lambda = \sum_{\substack{\sigma \in \text{Sp}(-\Delta) \\ \mu \in \text{Sp}(C^*) \\ \lambda = \sigma + \mu}} \sum_{i=1}^{n_\sigma} (1_\omega v_{\sigma,i})(x) B^* \Phi_{\mu,i}.$$

Assume now that $\mathcal{B}^* \Phi_\lambda = 0$. This implies, by the Lebeau-Robbiano spectral inequality (Theorem IV.2.33), that we actually have

$$0 = \sum_{\substack{\sigma \in \text{Sp}(-\Delta) \\ \mu \in \text{Sp}(C^*) \\ \lambda = \sigma + \mu}} \sum_{i=1}^{n_\sigma} v_{\sigma,i}(x) B^* \Phi_{\mu,i}, \quad \forall x \in \Omega.$$

Since all the functions $(v_{\sigma,i})_{\sigma,i}$ are orthonormal, we can take the $L^2(\Omega)$ norm and obtain

$$0 = \sum_{\substack{\sigma \in \text{Sp}(-\Delta) \\ \mu \in \text{Sp}(C^*) \\ \lambda = \sigma + \mu}} \sum_{i=1}^{n_\sigma} \|B^* \Phi_{\mu,i}\|^2 .$$

This implies that $B^* \Phi_{\mu,i} = 0$ for any μ and any i. Since the pair (B, C) is controllable and $\Phi_{\mu,i} \in$ Ker $(C^* - \mu)$, the finite-dimensional Fattorini-Hautus test leads to $\Phi_{\mu,i} = 0$ for any μ and any i and finally, we find that $\Phi_\lambda = 0$.

It follows that our adjoint system satisfies the (infinite dimensional) Fattorini-Hautus test from which we deduce the approximate controllability of the system. ■

Actually, a stronger result can be obtained by using Carleman estimates.

Theorem V.3.4

Under the above assumptions the system (III.3) is null-controllable for any time $T > 0$.

Proof :

To simplify a little bit the proof we assume that $n = 2$ and $m = 1$; however the same proof easily extends to the general case.

Let us introduce the Kalman matrix $K = (B, CB)$ and we perform the change of variable $y = Kz$ to obtain

$$K\partial_t z - K\Delta z + CKz = 1_\omega Bv,$$

Since K is invertible and $KC = \tilde{C}Z$ and $B = K\tilde{B}$, with

$$\tilde{C} = \begin{pmatrix} 0 & c_{12} \\ 1 & c_{22} \end{pmatrix}, \quad \tilde{B} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

the system is transformed into a *cascade system*

$$\partial_t z - \Delta z + \tilde{C}z = 1_\omega \tilde{B}v,$$

that we write

$$\begin{cases} \partial_t z_1 - \Delta z_1 + \quad c_{12}z_2 & = 1_\omega v, \\ \partial_t z_2 - \Delta z_2 + z_1 + c_{22}z_2 & = 0. \end{cases}$$

The corresponding adjoint system is

$$\begin{cases} -\partial_t q_1 - \Delta q_1 + \quad q_2 & = 0, \\ -\partial_t q_2 - \Delta q_2 + c_{12}q_1 + c_{22}q_2 & = 0, \end{cases}$$

and the observation operator if $\mathcal{B}^* = 1_\omega B^* = 1_\omega \begin{pmatrix} 1 & 0 \end{pmatrix}$, which is nothing but the operator that takes the restriction on ω to the **first** component of the adjoint state.

We notice that the approximate observability is clear from the elliptic Carleman estimate.

In other words, the observability inequality we need to prove for this adjoint system is

$$\|q_1(0)\|^2_{L^2(\Omega)} + \|q_2(0)\|^2_{L^2(\Omega)} = \|q(0)\|^2_{L^2} \leq C \int_0^T \int_\omega |q_1|^2.$$

As we have seen before, we already know how to prove the same inequality but with an other observation term on ω involving the term q_2 but here we do not want this term in the inequality. The only way to get rid of this term is to express q_2 as a function of q_1 by using the first equation $q_2 = \partial_t q_1 + \Delta q_1$. However, this will make appear high derivatives of q_1 that are not allowed.

We thus need to come back at the Carleman estimate level. To simplify the computations, we define the quantities

$$J(d, f, U) \stackrel{\text{def}}{=} \int_0^T \int_U (s\theta)^d \left| e^{s\theta\varphi} f \right|^2.$$

With those notation, we write the parabolic Carleman estimate for q_1 with $d = d_1$ and for q_2 with another value $d = d_2$. Moreover, we will take into account some of the terms allowed by Remark IV.4.49. For q_1 we get

$$J(d_1, q_1, \Omega) + J(d_1 - 2, \nabla q_1, \Omega) \leq CJ(d_1, q_1, \omega) + CJ(d_1 - 3, \partial_t q_1 + \Delta q_1, \Omega),$$

and for q_2

$$J(d_2, q_2, \Omega) + J(d_2 - 2, \nabla q_2, \Omega) + J(d_2 - 4, \partial_t q_2, \Omega) + J(d_2 - 4, \Delta q_2, \Omega) \\ \leq CJ(d_2, q_2, \omega) + CJ(d_2 - 3, \partial_t q_2 + \Delta q_2, \Omega),$$

We use now the equations satisfied by q_1 and q_2, to get

$$J(d_1,q_1,\Omega)+J(d_1-2,\nabla q_1,\Omega)\leq CJ(d_1,q_1,\omega)+CJ(d_1-3,q_2,\Omega), \tag{V.3}$$

$$J(d_2,q_2,\Omega)+J(d_2-2,\nabla q_2,\Omega)+J(d_2-4,\partial_t q_2,\Omega)+J(d_2-4,\Delta q_2,\Omega)\\ \leq CJ(d_2,q_2,\omega)+CJ(d_2-3,q_1,\Omega)+CJ(d_2-3,q_2,\Omega), \tag{V.4}$$

In order to perform the following computations we choose now $d_1=7$ and $d_2=4$ and we add (V.3) that we multiply by some $\varepsilon>0$ and (V.4). We obtain

$$\varepsilon J(7,q_1,\Omega)+\varepsilon J(5,\nabla q_1,\Omega)+J(4,q_2,\Omega)+J(2,\nabla q_2,\Omega)+J(0,\partial_t q_2,\Omega)+J(0,\Delta q_2,\Omega)\\ \leq C\varepsilon J(7,q_1,\omega)+C\varepsilon J(4,q_2,\Omega)+CJ(4,q_2,\omega)+CJ(1,q_1,\Omega)+CJ(1,q_2,\Omega).$$

By chosing $\varepsilon>0$ small enough (depending only on the data) we can absorb the second term in the right-hand side by the third one of the left-hand side. This value of ε being now fixed, we will not make it appear in the sequel. Moreover, we use that

$$(s\theta)^1=(s\theta)^4(s\theta)^{-3}\leq\frac{C}{s^3}(s\theta)^4,$$

$$(s\theta)^1=(s\theta)^7(s\theta)^{-6}\leq\frac{C}{s^6}(s\theta)^7,$$

to say that, for a well chosen s_1 (depending only on the data), and any $s\geq s_1$, we can absorb the last two terms in the right-hand side by the first and third of the left-hand side.

To sum up, we have now the following estimate

$$J(7,q_1,\Omega)+J(5,\nabla q_1,\Omega)+J(4,q_2,\Omega)+J(2,\nabla q_2,\Omega)+J(0,\partial_t q_2,\Omega)+J(0,\Delta q_2,\Omega)\\ \leq CJ(7,q_1,\omega)+CJ(4,q_2,\omega).$$

We still have two observation terms and we would like to get rid of the one in q_2. It seems that we do not have make great progresses compared to the estimate obtained in Section V.1. However, the additional term in the left-hand side, as well as the different powers of $(s\theta)$ in both terms is a real progress.

First of all we replace the observation set ω in the above estimate by a smaller one ω_0 (such that $\overline{\omega_0}\subset\omega$). This requires of course to consider a slightly different weight function but we do not change the notation. We consider now a function η compactly supported in ω and such that $0\leq\eta\leq 1$ and $\eta=1$ in ω_0. It follows, by using the first equation of the system that

$$\begin{aligned}J(4,q_2,\omega_0)&=\int_0^T\int_{\omega_0}(s\theta)^4\left|e^{s\theta\varphi}q_2\right|^2\\ &\leq\int_0^T\int_{\omega}\eta(s\theta)^4\left|e^{s\theta\varphi}q_2\right|^2\\ &=\int_0^T\int_{\omega}\eta(s\theta)^4e^{2s\theta\varphi}q_2(\partial_t q_1+\Delta q_1).\end{aligned}$$

We evaluate now the term (referred to as I_1) in $\partial_t q_1$ and the one (referred to as I_2) in Δq_1 independently.

- In the term I_1, we perform an integration by parts in time (observing that there is no boundary term since the weight $e^{2s\theta\varphi}$ is exponentially flat in 0 and T.

$$I_1=-\int_0^T\int_{\omega}\eta(s\theta)^4e^{2s\theta\varphi}(\partial_t q_2)q_1-\int_0^T\int_{\omega}\eta s^4\theta^3(4\theta'+2s\theta\theta'\varphi)e^{2s\theta\varphi}q_2q_1.$$

Using that $\theta' \leq C\theta^2$, and the Cauchy-Schwarz inequality (with a suitable repartition of the weights $(s\theta)^\bullet$ in both terms), we get (for $s \geq 1$)

$$\begin{aligned} I_1 \leq & \int_0^T \int_\omega \eta(s\theta)^4 e^{2s\theta\varphi} |q_1 \partial_t q_2| + C \int_0^T \int_\omega \eta(s\theta)^6 e^{2s\theta\varphi} |q_2 q_1| \\ \leq & CJ(0, \partial_t q_2, \Omega)^{\frac{1}{2}} J(8, q_1, \omega)^{\frac{1}{2}} + CJ(4, q_2, \Omega)^{\frac{1}{2}} J(8, q_1, \omega)^{\frac{1}{2}}. \end{aligned}$$

Observe that we have mentioned Ω instead of ω in the terms concerning q_2 since we actually don't care that there are supported in ω (we will absorb them by left-hand side terms of the estimate). However, it is crucial that the terms in q_1 are localised in ω; those will contribute to the observation term at the end.

- In the term I_2 we perform three successive integrations by parts in space (without boundary terms since η is compactly supported), in order to make all the derivatives apply on q_2 instead of q_1. It follows

$$\begin{aligned} I_2 = & - \int_0^T \int_\omega \eta(s\theta)^4 e^{2s\theta\varphi} \nabla q_2 \cdot \nabla q_1 - \int_0^T \int_\omega (s\theta)^4 e^{2s\theta\varphi} q_2 (\nabla \eta + 2s\theta \nabla \varphi) \cdot \nabla q_1 \\ = & \int_0^T \int_\omega \eta(s\theta)^4 e^{2s\theta\varphi} (\Delta q_2) q_1 + \int_0^T \int_\omega (s\theta)^4 e^{2s\theta\varphi} q_1 (\nabla \eta + 2s\theta \nabla \varphi) \cdot \nabla q_2 \\ & + \int_0^T \int_\omega (s\theta)^4 e^{2s\theta\varphi} \nabla q_2 \cdot (\nabla \eta + 2s\theta \nabla \varphi) q_1 \\ & + \int_0^T \int_\omega (s\theta)^4 e^{2s\theta\varphi} (\Delta \eta + 2s\theta \Delta \varphi + 2s\theta \nabla \varphi \cdot \nabla \eta + 4s^2\theta^2 |\nabla \varphi|^2) q_2 q_1 \\ \leq & C \int_0^T \int_\omega (s\theta)^4 e^{2s\theta\varphi} |\Delta q_2| |q_1| + C \int_0^T \int_\omega (s\theta)^5 e^{2s\theta\varphi} |q_1| |\nabla q_2| + C \int_0^T \int_\omega (s\theta)^6 e^{2s\theta\varphi} |q_1| |q_2| \\ \leq & CJ(0, \Delta q_2, \Omega)^{\frac{1}{2}} J(8, q_1, \omega)^{\frac{1}{2}} + CJ(2, \nabla q_2, \Omega)^{\frac{1}{2}} J(8, q_1, \omega)^{\frac{1}{2}} + CJ(4, q_2, \Omega)^{\frac{1}{2}} J(8, q_1, \omega)^{\frac{1}{2}}. \end{aligned}$$

We gather the bound on I_1 and the one on I_2 and we use Young's inequality to obtain

$$\begin{aligned} J(7, q_1, \Omega) + J(5, \nabla q_1, \Omega) + J(4, q_2, \Omega) + J(2, \nabla q_2, \Omega) + J(0, \partial_t q_2, \Omega) + J(0, \Delta q_2, \Omega) \\ \leq CJ(7, q_1, \omega) + CJ(8, q_1, \omega). \end{aligned}$$

We finally obtained an estimate with a unique local observation term in q_1

$$J(7, q_1, \Omega) + J(5, \nabla q_1, \Omega) + J(4, q_2, \Omega) + J(2, \nabla q_2, \Omega) + J(0, \partial_t q_2, \Omega) + J(0, \Delta q_2, \Omega) \leq CJ(8, q_1, \omega).$$

We retain from this inequality only the terms in q_1 and q_2

$$J(7, q_1, \Omega) + J(4, q_2, \Omega) \leq CJ(8, q_1, \omega),$$

from which the observability inequality can proved the same way as before, by using dissipation estimates on q. ■

V.3.2 Variable coefficient cascade systems - The good case

In the case where the coupling coefficients in the system depend on x, we will see that the controllability properties of the system may be quite different.

If we assume that the *significant* coupling coefficients (i.e. the ones that are responsible for the indirect action of one controlled component of the system on the non-controlled components) do not identically vanish inside the control domain ω, the analysis is simpler. More precisely, as an example, we consider the following 2×2 system

$$\begin{cases} \partial_t z_1 - \Delta z_1 + c_{11}(x) z_1 + c_{12}(x) z_2 & = 1_\omega v, \\ \partial_t z_2 - \Delta z_2 + c_{21}(x) z_1 + c_{22}(x) z_2 & = 0, \end{cases} \tag{V.5}$$

and we assume that c_{21} does not identically vanish in ω, and more precisely : there exists a non-empty $\omega_0 \subset \omega$ such that

$$\exists \omega_0 \subset \omega, \quad \text{s.t.} \quad \inf_{\omega_0} |c_{21}| > 0. \tag{V.6}$$

Using similar techniques as in the scalar case, based on elliptic Carleman estimates, we can prove the following result.

Proposition V.3.5

Under the assumption (V.6), the system (V.5) is approximately controllable for any time $T > 0$.

Proof :

We will use the Fattorini-Hautus criterion. Let q be a (complex) eigenfunction of the adjoint elliptic operator associated with the (complex) eigenvalue λ. We assume that $\mathcal{B}^* q = 1_\omega q_1 = 0$ and we would like to prove that $q = 0$. The equation satisfied by q are

$$\begin{cases} -\Delta q_1 + c_{11}(x) q_1 + c_{21}(x) q_2 &= \lambda q_1, \\ -\Delta q_2 + c_{12}(x) q_1 + c_{22}(x) q_2 &= \lambda q_2. \end{cases}$$

By assumption, we have $q_1 = 0$ in ω_0 and $\inf_{\omega_0} |c_{21}| > 0$ so that the first equation leads to $q_2 = 0$ in ω_0. We apply now the global elliptic Carleman estimate given in Theorem IV.1.29 (for the observation domain ω_0) to q_1 and q_2 and we sum the two inequalities to obtain for any $s \geq s_0$,

$$s^3 \|e^{s\varphi} q_1\|^2_{L^2(\Omega)} + s^3 \|e^{s\varphi} q_2\|^2_{L^2(\Omega)} \leq C \Big(\|e^{s\varphi} \Delta q_1\|^2_{L^2(\Omega)} + \|e^{s\varphi} \Delta q_2\|^2_{L^2(\Omega)} + s^3 \|e^{s\varphi} q_1\|^2_{L^2(\omega_0)} + s^3 \|e^{s\varphi} q_2\|^2_{L^2(\omega_0)} \Big).$$

Since $q_1 = q_2 = 0$ in ω_0 and using the equations to express Δq_1 and Δq_2, we get

$$s^3 \|e^{s\varphi} q_1\|^2_{L^2(\Omega)} + s^3 \|e^{s\varphi} q_2\|^2_{L^2(\Omega)} \leq C \big(\max_{i,j} \|c_{ij}\|^2_{L^\infty} + |\lambda|^2 \big) \big(\|e^{s\varphi} q_1\|^2_{L^2(\Omega)} + \|e^{s\varphi} q_2\|^2_{L^2(\Omega)} \big).$$

Taking s large enough gives

$$s^3 \|e^{s\varphi} q_1\|^2_{L^2(\Omega)} + s^3 \|e^{s\varphi} q_2\|^2_{L^2(\Omega)} \leq 0,$$

and the claim is proved. ■

In fact the following, much stronger, result holds.

Proposition V.3.6

Under the same assumption (V.6), the system (V.5) is null-controllable at any time $T > 0$ (even if we allow the coefficients c_{ij} to depend on time).

Proof :

The strategy we used in Section V.1 can be applied exactly in the same way for such variable coefficients cascade systems. The only point is to be able to express q_2 as a function of q_1 in ω_0 by writing

$$q_2 = \frac{1}{c_{21}} \Big(\partial_t q_1 + \Delta q_1 - c_{11} q_1 \Big).$$

Details are left to the reader. ■

V.3.3 Variable coefficient cascade systems - The not so good case

In this section we will consider particular cascade systems in which the support of the coupling terms do not intersect the control region.

$$\begin{cases} \partial_t y + \mathcal{A} y + C(x) y = \mathbf{1}_\omega B v, & \text{in } \Omega \\ y = 0, & \text{on } \partial\Omega, \end{cases} \tag{V.7}$$

with

$$B = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \text{ and } C(x) = 0, \text{ in } \omega.$$

It is clear that the strategies relying on Carleman estimates are not usable in such a case since we will not be able to remove the unwanted observation term at the end as we did in Section V.1.

The general analysis of such systems (in particular in higher dimensions) remains an open problem at that time. We will concentrate here on the case of the 2×2 systems in the cascade form, that is we assume that the coupling matrix reads

$$C(x) = \begin{pmatrix} 0 & 0 \\ c_{21}(x) & 0 \end{pmatrix}. \tag{V.8}$$

Most of the analysis will rely on a precise knowledge of the eigenelements of the operator

$$\mathcal{L}^* = \mathcal{A} + C(x)^*.$$

V.3.3.1 Description of the spectrum of $\mathcal{L}^*$

A very simple analysis, using the Fredholm alternative, gives us the structure of the spectrum of $\mathcal{L}^*$.

Proposition V.3.7 (Spectrum of $\mathcal{L}^*$)

We have $\mathrm{Sp}(\mathcal{L}^*) = \mathrm{Sp}(\mathcal{A})$. For any $\lambda \in \mathrm{Sp}(\mathcal{A})$, let $n_\lambda = \dim \operatorname{Ker} (\mathcal{A} - \lambda)$ and $(\phi_{\lambda,i})_{i \in [\![1,n_\lambda]\!]}$ be an orthonormal family of eigenfunctions of $\mathcal{A}$ associated with λ. For each $i \in [\![1, n_\lambda]\!]$ we define

$$I_{\lambda,i}(c_{21}) \stackrel{\text{def}}{=} \int_\Omega c_{21} |\phi_{\lambda,i}|^2 \, dx.$$

1. *For each $i \in [\![1, n_\lambda]\!]$, the vector-valued function*

$$\Phi_{\lambda,i} = \begin{pmatrix} \phi_{\lambda,i} \\ 0 \end{pmatrix},$$

is an eigenfunction of $\mathcal{L}^*$.

2. *For each $i \in [\![1, n_\lambda]\!]$ such that $I_{\lambda,i}(c_{21}) = 0$, there exists an eigenfunction of $\mathcal{L}^*$ of the form*

$$\tilde{\Phi}_{\lambda,i} = \begin{pmatrix} \tilde{\phi}_{\lambda,i} \\ \phi_{\lambda,i} \end{pmatrix},$$

where $\tilde{\phi}_{\lambda,i}$ is a solution of $(\mathcal{A} - \lambda)\tilde{\phi}_{\lambda,i} = -c_{21}\phi_{\lambda,i}$.

3. *For each $i \in [\![1, n_\lambda]\!]$ such that $I_{\lambda,i}(c_{21}) \neq 0$, there exists a generalised eigenfunction of $\mathcal{L}^*$ satisfying $(\mathcal{L}^* - \lambda)(\Psi_{\lambda,i}) = \Phi_{\lambda,i}$ of the form*

$$\tilde{\Phi}_{\lambda,i} = \frac{1}{I_{\lambda,i}(c_{21})} \begin{pmatrix} \tilde{\phi}_{\lambda,i} \\ \phi_{\lambda,i} \end{pmatrix},$$

where $\tilde{\phi}_{\lambda,i}$ is any solution of $(\mathcal{A} - \lambda)\tilde{\phi}_{\lambda,i} = -\Big(c_{21} - I_{\lambda,i}(c_{21})\Big)\phi_{\lambda,i}$.

Finally, the family $\{\Phi_{\lambda,i}, \tilde{\Phi}_{\lambda,i},\ \lambda \in \Lambda, i \in [\![1, n_\lambda]\!]\}$ is linearly independent and complete in $(L^2(\Omega))^2$.

V.3.3.2 Approximate controllability in any dimension

By using the Fattorini-Hautus test, we known that the study of the approximate controllability of our system amounts at determining whether or not the eigenfunctions of $\mathcal{L}^*$ belong to the kernel of $\mathcal{B}^* = 1_\omega B^*$.

In any dimension, we have a sufficient approximate controllability condition which is the following.

Theorem V.3.8

Assume that c_{21} is continuous not identically zero and that $c_{21} \geq 0$, then the 2×2 system (V.7) with C given by (V.8) is approximately controllable at any time $T > 0$.

Proof :

By assumption on c_{21}, we know that for any $\lambda \in \Lambda$ and any $i \in [\![1, n_\lambda]\!]$ the number $I_{\lambda,i}(c_{21})$ cannot vanish since it is the integral of a non-negative function which is not identically zero. Indeed, by Proposition IV.1.30, we know that any eigenfunction of $\mathcal{A}$ cannot identically vanish on the non-empty open subset $\{x \in \Omega, c_{21}(x) > 0\}$.

Therefore, we know from Proposition V.3.7 that every eigenfunction $\Phi \in \mathrm{Ker}\,(\mathcal{L}^* - \lambda)$ can be written

$$\Phi = \sum_{i=1}^{n_\lambda} a_{\lambda,i}\Phi_{\lambda,i}.$$

By definition of the observation operator $\mathcal{B}^*$ we thus have

$$\mathcal{B}^*\Phi = 1_\omega \left(\sum_{i=1}^{n_\lambda} a_{\lambda,i}\phi_{\lambda,i}\right).$$

Therefore, if $\mathcal{B}^*\Phi = 0$ we deduce that $a_{\lambda,i} = 0$ for every i thanks to Proposition IV.1.30 and thus $\Phi = 0$.

From the Fattorini-Hautus test (Theorem III.3.7), the claim is proved. ∎

V.3.3.3 Approximate controllability in 1D

In the 1D case (see [BO14]), we can give a more precise result which is a necessary and sufficient approximate controllability condition. Since, in that case, each eigenvalue of $\mathcal{A}$ is simple we can use Proposition V.3.7 with $n_\lambda = 1$ for any λ. As a consequence, we will drop the index i in the notation. To get a complete analysis we will need to introduce a function ψ_λ linearly independent from ϕ_λ and that solves the ODE

$$\mathcal{A}\psi_\lambda = \lambda\psi_\lambda.$$

Note that ψ_λ does not satisfy the Dirichlet boundary conditions.

Definition V.3.9

For any $\lambda \in \mathrm{Sp}(\mathcal{A})$, any interval $[a, b] \subset [0, 1]$, and any integrable function f, we define the following element of $\mathbb{R}^2$

$$M_\lambda(f, [a, b]) \stackrel{\text{def}}{=} \begin{cases} \begin{pmatrix} \int_a^b f\phi_\lambda \\ \int_a^b f\psi_\lambda \end{pmatrix}, & \text{if } [a,b]\cap\partial\Omega = \emptyset, \\ \begin{pmatrix} \int_a^b f\phi_\lambda \\ 0 \end{pmatrix}, & \text{if } [a,b]\cap\partial\Omega \neq \emptyset. \end{cases}$$

Theorem V.3.10

Assume that c_{21} identically vanishes in the control region ω.
Then the 2×2 cascade system (V.7) is approximately controllable if and only if, for any $\lambda \in \mathrm{Sp}(\mathcal{A})$, there exists a connected component $[a,b]$ of $\overline{\Omega \setminus \omega}$ such that

$$M_\lambda(c_{21}\phi_\lambda, [a,b]) \neq 0.$$

Remark V.3.11

If c_{21} does not identically vanish in ω, we already know by Theorem V.3.8 that the system is approximately controllable, in any dimension.

Proof :

- Let us show that the condition is sufficient. To this end, we assume that the system is not approximately controllable. By the Fattorini-Hautus test (see Theorem III.3.7) we know that it necessarily exists an eigenfunction Φ of $\mathcal{L}^*$ associated with the eigenvalue λ such that $\mathcal{B}^*\Phi = 0$.
  - If $I_\lambda(c_{21}) \neq 0$, then we know that Φ is necessarily a multiple of $\Phi_\lambda = \begin{pmatrix}\phi_\lambda \\ 0\end{pmatrix}$ and therefore $\mathcal{B}^*\Phi$ is a multiple of $1_\omega \phi_\lambda$ which cannot be identically zero.
  - We thus conclude that $I_\lambda(c_{21}) = 0$, and thus up to a multiplicative factor Φ is necessarily of the form

    $$\Phi = \begin{pmatrix}\tilde{\phi}_\lambda \\ \phi_\lambda\end{pmatrix},$$

    where $\tilde{\phi}_\lambda$ satisfies, along with the Dirichlet boundary conditions, the equation

    $$(\mathcal{A} - \lambda)\tilde{\phi}_\lambda = -c_{21}\phi_\lambda.$$

    By assumption we have $\mathcal{B}^*\Phi = 0$ which implies that $\tilde{\phi}_\lambda = 0$ on ω.
  - Let $[a,b]$ be a connected component of $\overline{\Omega \setminus \omega}$, and let us compute by integration by parts

    $$\begin{aligned}\int_a^b c_{21}|\phi_\lambda|^2\,dx &= -\int_a^b ((\mathcal{A}-\lambda)\tilde{\phi}_\lambda)\phi_\lambda\,dx \\ &= [\gamma\tilde{\phi}'_\lambda\phi_\lambda]_a^b - [\gamma\tilde{\phi}_\lambda\phi'_\lambda]_a^b.\end{aligned}$$

    Let us show that all the terms in this last formula vanish.
    - If $a \in \Omega$, we have $a \in \partial\omega$, and since we have assumed that $\tilde{\phi}_\lambda = 0$ in ω, we obtain $\tilde{\phi}_\lambda(a) = \tilde{\phi}'_\lambda(a) = 0$ and thus

      $$(\gamma\phi'_\lambda\tilde{\phi}_\lambda)(a) = (\gamma\tilde{\phi}'_\lambda\phi_\lambda)(a) = 0.$$
    - If $a \in \partial\Omega$ then $\phi_\lambda(a) = \tilde{\phi}_\lambda(a) = 0$ thanks to the boundary conditions and thus we also have

      $$(\gamma\phi'_\lambda\tilde{\phi}_\lambda)(a) = (\gamma\tilde{\phi}'_\lambda\phi_\lambda)(a) = 0.$$
    - A similar reasoning holds for the point b.

    It follows that we necessarily have

    $$\int_a^b c_{21}|\phi_\lambda|^2\,dx = 0.$$

- If, in addition, $[a,b]$ does not touch the boundary of Ω we can compute similarly

$$\begin{aligned}\int_a^b c_{21}\phi_\lambda\psi_\lambda\,dx &= -\int_a^b ((\mathcal{A}-\lambda)\tilde{\phi}_\lambda)\psi_\lambda\,dx\\ &= [\gamma\tilde{\phi}'_\lambda\psi_\lambda]_a^b - [\gamma\tilde{\phi}_\lambda\psi'_\lambda]_a^b\\ &= 0,\end{aligned}$$

  by the same argument as before.
- All in all, we have eventually shown that

$$M_\lambda(c_{21}\phi_\lambda,[a,b]) = 0,$$

  and the claim is proved.

- Let us now show that the proposed condition is necessary. Let us assume that for a given eigenvalue λ, we have $M_\lambda(c_{21}\phi_\lambda,[a,b]) = 0$ for any connected component $[a,b]$ of $\overline{\Omega\setminus\omega}$.

  This implies, in particular that for any such $[a,b]$ we have

$$\int_a^b c_{21}|\phi_\lambda|^2\,dx = 0,$$

  and since $c_{21}=0$ in ω, we eventually find by summation that

$$\int_\Omega c_{21}|\phi_\lambda|^2\,dx = 0.$$

  This exactly means that $I_\lambda(c_{21}) = 0$.

  By Proposition V.3.7 we conclude that there any function of the form

$$\Phi = \tilde{\Phi}_\lambda + \beta\Phi_\lambda,$$

  with $\beta\in\mathbb{R}$, is an eigenfunction of $\mathcal{L}^*$. In particular we have

$$\mathcal{B}^*\Phi = 1_\omega(\tilde{\phi}_\lambda + \beta\phi_\lambda).$$

  We set $\zeta = \tilde{\phi}_\lambda + \beta\phi_\lambda$ and we will determine β is such a way that ζ identically vanish in ω.

  - We will first find a value of β and a point $x_0\in\overline{\omega}$ such that $\zeta(x_0)=\zeta'(x_0)=0$.
    * If $\overline{\omega}\cap\partial\Omega\neq\emptyset$, then we take any $x_0\in\overline{\omega}\cap\partial\Omega$. We immediately have $\zeta(x_0)=0$ and $\zeta'(x_0) = \tilde{\phi}'_\lambda(x_0)+\beta\phi'_\lambda(x_0)$. Since $\phi'_\lambda(x_0)\neq 0$ we see that one can choose β such that $\zeta'(x_0)=0$.
    * If $\overline{\omega}\cap\partial\Omega=\emptyset$, we consider $[0,b]$ the connected component of $\overline{\Omega\setminus\omega}$ that contains 0. By assumption, we have

$$\int_0^b c_{21}|\phi_\lambda|^2 = 0.$$

      We can find a $\delta>0$ small enough such that $]b,b+\delta[\subset\omega$ and $\phi_\lambda(b+\delta)\neq 0$. We can then choose β such that

$$0 = \tilde{\phi}_\lambda(b+\delta)+\beta\phi_\lambda(b+\delta) = \zeta(b+\delta).$$

      Since $c_{21}=0$ in ω, we deduce that

$$\begin{aligned}0 &= \int_0^{b+\delta} c_{21}|\phi_\lambda|^2\,dx\\ &= -\int_0^{b+\delta}(\mathcal{A}\zeta-\lambda\zeta)\phi_\lambda\,dx\\ &= -(\gamma\zeta'\phi_\lambda)(b+\delta),\end{aligned}$$

where we have used that $\zeta(0) = \phi_\lambda(0) = \zeta(b+\delta) = 0$.
Since $\gamma(b+\delta)\phi_\lambda(b+\delta) \neq 0$, we necessarily have $\zeta'(b+\delta) = 0$ and therefore the point $x_0 = b+\delta$ fulfills our requirements.

– Let us show now that $\zeta(x_1) = 0$ for any point $x_1 \in \omega$. Assume for instance that $x_1 > x_0$. Since $[x_0, x_1] \cap \overline{\Omega\setminus\omega}$ is an union of connected components of $\overline{\Omega\setminus\omega}$ we have, by assumption

$$\int_{x_0}^{x_1} c_{21}|\phi_\lambda|^2\,dx = \int_{x_0}^{x_1} c_{21}\phi_\lambda\psi_\lambda\,dx = 0.$$

Using again an integration by parts, the equations satisfied by ζ, ϕ_λ and ψ_λ, and the fact that $\zeta(x_0) = \zeta'(x_0) = 0$, we obtain the two equations

$$\begin{cases} 0 = -\zeta'(x_1)\phi_\lambda(x_1) + \zeta(x_1)\phi_\lambda'(x_1), \\ 0 = -\zeta'(x_1)\psi_\lambda(x_1) + \zeta(x_1)\psi_\lambda'(x_1). \end{cases}$$

Since ϕ_λ and ψ_λ are two linearly independent solutions of the same second order linear ODE, we know that the Wronskian determinant satisfies

$$\begin{vmatrix} \phi_\lambda(x_1) & \psi_\lambda(x_1) \\ \phi_\lambda'(x_1) & \psi_\lambda'(x_1) \end{vmatrix} \neq 0,$$

and thus we conclude that

$$\zeta(x_1) = \zeta'(x_1) = 0.$$

The claim is proved.

We have thus found an eigenfunction $\Phi = \begin{pmatrix} \zeta \\ \phi_\lambda \end{pmatrix}$ of $\mathcal{L}^*$ such that $\mathcal{B}^*\Phi = 1_\omega\zeta = 0$ and thus (V.7) is not approximately controllable, thanks to the Fattorini-Hautus test.

■

Some examples. Let us analyze some particular examples of such systems. We will see that many different situations can occur.

- We consider the set $\mathcal{O} = (1/4, 3/4)$ and we take for some $a \in \mathbb{R}$

$$c_{21}(x) = (x-a)1_{\mathcal{O}}(x).$$

– Subcase 1 : Assume that $\omega \subset (3/4, 1)$. The only connected component of $\overline{\Omega\setminus\omega}$ that touches the coupling support $\mathcal{O}$ contains $(0, 3/4)$. In that case we know that the system is approximately controllable if and only if

$$\int_{\mathcal{O}} c_{21}|\phi_\lambda|^2\,dx \neq 0.$$

A simple computation thus shows that

$$\text{the system is approximately controllable} \iff a \notin \{a_\lambda\}_{\lambda\in\Lambda},$$

where

$$a_\lambda = \frac{\displaystyle\int_{\mathcal{O}} x|\phi_\lambda|^2}{\displaystyle\int_{\mathcal{O}} |\phi_\lambda|^2}, \quad \forall \lambda \in \Lambda.$$

– Subcase 2 : Assume now that $\omega \cap (3/4, 1) \neq \emptyset$ and $\omega \cap (0, 1/4) \neq \emptyset$. If $a \notin \{a_\lambda\}_{\lambda\in\Lambda}$, then it is clear that the system is approximately controllable from the previous analysis. However, since the concerned connected component of $\Omega \setminus \omega$ does not touch the boundary of Ω, we have to check whether or not we have

$$\int_{\mathcal{O}} c_{21}\phi_\lambda\psi_\lambda = 0.$$

This condition is not explicit in general but we can discuss a particular case where $\mathcal{A} = -\partial_x^2$. In this case we have $\Lambda = \{k^2\pi^2,\ k \geq 1\}$ and $\phi_\lambda(x) = \sin(\sqrt{\lambda}x)$ and $\psi_\lambda(x) = \cos(\sqrt{\lambda}x)$ and we can check that $a_\lambda = 1/2$ for any $\lambda \in \Lambda$.

It remains to compute, for $a = a_\lambda = 1/2$,

$$\int_{\mathcal{O}} c_{21}\phi_\lambda\psi_\lambda = \int_{1/4}^{3/4} (x - 1/2)\sin(\sqrt{\lambda}x)\cos(\sqrt{\lambda}x) = \begin{cases} \frac{-1}{8\sqrt{\lambda}}(-1)^{k/2}, & \text{if } \lambda = k^2\pi^2 \text{ with } k \text{ even}, \\ \frac{-1}{4\lambda}(-1)^{(k-1)/2}, & \text{if } \lambda = k^2\pi^2 \text{ with } k \text{ odd}. \end{cases}$$

Since those quantities never vanish, we deduce that our system, for this choice of ω, is always approximately controllable.

V.3.3.4 Null controllability in 1D

The main result in this direction proved in [KBGBdT16] is, in a simplified version, the following

Theorem V.3.12

Assume that ω in an interval that touches the boundary of Ω and that $c_{21} = 0$ in the control domain ω. Then there exists a time $T_0(c_{21}) \in [0, +\infty]$ such that

- *For $T > T_0(c_{21})$, the system (V.7) with (V.8) is null-controllable.*
- *For $T < T_0(c_{21})$, the system (V.7) with (V.8) is not null-controllable.*

Moreover, for any $T^* \in [0, \infty]$, there exists a coupling function c_{21} such that $T_0(c_{21}) = T^*$.

Note that in the above reference a more or less explicit formula for $T_0(c_{21})$ is given.
The proof strategy is the following

- Compute the eigenelements of the operator $\mathcal{L}^*$. We find that the eigenfunctions are the

$$\begin{pmatrix} \phi_\lambda \\ 0 \end{pmatrix},$$

with the associated generalized eigenfunctions given by

$$\begin{pmatrix} \psi_\lambda \\ \phi_\lambda \end{pmatrix},$$

for some explicit function ψ_λ.

- Case $T > T_0(c_{12})$: the positive controllability result is proved by using the moments method.

- Case $T < T_0(c_{12})$: the negative controllability result is proved by showing that the observability inequality does not hold for some well-chosen final data q_T built as a combination of the above two (generalized) eigenfunctions of $\mathcal{L}^*$.

V.4 Boundary controllability results for some 1D systems

We will only consider here the following constant coefficient system in the 1D interval $\Omega = (0,1)$

$$\begin{cases} \partial_t y + \mathcal{A}y + Cy = 0, & \text{in } \Omega = (0,1) \\ y = \mathbf{1}_{\{0\}} Bv, & \text{on } \partial\Omega. \end{cases} \tag{V.9}$$

We will point out the main differences with the distributed control problem for the same system.

V.4.1 Approximate controllability

Proposition V.4.13

A necessary condition for the null- or approximate- controllability for (V.9) is that the pair (C,B) is controllable.

Proof :

Let y be any solution of (V.9) and ϕ_λ an eigenfunction of $\mathcal{A}$ associated with an eigenvalue λ. Then, we obtain that the quantity $z(t) = \langle y(t), \phi_\lambda \rangle_{L^2} \in \mathbb{R}^n$, solves the following ordinary differential equation

$$\frac{d}{dt} z + \lambda z + Cz = \pm \phi'_\lambda(0) Bv(t). \tag{V.10}$$

Then the null-controllability (resp. approximate controllability) of (V.9), implies the null-controllability (resp. approximate controllability) of the reduced system (V.10). It implies that the pair $(C + \lambda \mathrm{Id}, \phi'_\lambda(0)B)$ is controllable and since $\phi'_\lambda(0) \neq 0$, this gives in turn that (C,B) satisfies the Kalman criterion. ∎

Theorem V.4.14

Assume that $m = 1 = \mathrm{Rank} B$ (the general case can be studied similarly).
System (V.9) is approximately controllable at time $T > 0$ if and only if the pair (C,B) is controllable and the following condition holds

$$\sigma + \mu = \sigma' + \mu' \quad \Longrightarrow \sigma = \sigma', \tag{V.11}$$

for any $\sigma, \sigma' \in \mathrm{Sp}(\mathcal{A})$ and $\mu, \mu' \in \mathrm{Sp}(C^*)$.

Proof :

Each eigenvalue of $\mathcal{L}^* = \mathcal{A} + C^*$ is of the form $\lambda = \sigma + \mu$ where $\sigma \in \mathrm{Sp}(\mathcal{A})$ and $\mu \in \mathrm{Sp}(C^*)$ and any element in Ker $(\mathcal{L}^* - \lambda)$ can be written

$$\Phi_\lambda = \sum_{\substack{\sigma \in \mathrm{Sp}(\mathcal{A}) \\ \mu \in \mathrm{Sp}(C^*) \\ \lambda = \sigma + \mu}} \phi_\sigma(x) V_\mu,$$

where each V_μ belongs to Ker $(C^* - \mu)$.

When applying the observation operator $\mathcal{B}^* = B^* \frac{\partial}{\partial x}_{|x=0}$ we obtain

$$\mathcal{B}^* \Phi_\lambda = - \sum_{\substack{\sigma \in \mathrm{Sp}(\mathcal{A}) \\ \mu \in \mathrm{Sp}(C^*) \\ \lambda = \sigma + \mu}} \phi'_\sigma(0) B^* V_\mu.$$

- Assume that Condition (V.11) holds. It implies that there is only one term in the sum above. It follows that

  $$\mathcal{B}^* \Phi_\lambda = -\phi'_\sigma(0) B^* V_\mu,$$

  for a given σ and a given μ. Since we have assumed that (C,B) is controllable the finite dimensional Fattorini-Hautus test proves that $B^* V_\mu \neq 0$, and since $\phi'_\sigma(0) \neq 0$ we deduce that $\mathcal{B}^* \Phi_\lambda \neq 0$.

  This proves the Fattorini-Hautus condition.

- Assume that (V.11) does not hold. Then there exist $\sigma, \sigma' \in \mathrm{Sp}(A)$ with $\sigma \neq \sigma'$ and $\mu, \mu' \in \mathrm{Sp}(C^*)$ such that $\sigma + \mu = \sigma' + \mu'$.

  We pick $V_\mu, V_{\mu'}$ two eigenvectors of C^* associated with μ and μ' respectively. Then, the function

$$\Phi(x) = \frac{\phi'_{\sigma'}(0)}{B^* V_\mu} \phi_\sigma(x) V_\mu - \frac{\phi'_\sigma(0)}{B^* V_{\mu'}} \phi_{\sigma'}(x) V_{\mu'},$$

  which is well-defined since, by the Fattorini-Hautus test applied to the pair (C, B), we have $B^* V_\mu \neq 0$ and $B^* V_{\mu'} \neq 0$. By construction, Φ is an eigenfunction of our adjoint operator $\mathcal{L}^*$. Moreover we have

$$\mathcal{B}^* \Phi = -\frac{\phi'_{\sigma'}(0)}{B^* V_\mu} \phi'_\sigma(0) B^* V_\mu + \frac{\phi'_\sigma(0)}{B^* V_{\mu'}} \phi'_{\sigma'}(0) B^* V_{\mu'} = 0.$$

  This shows that the Fattorini-Hautus test is not fulfilled by our system and thus it is not approximately controllable.

∎

Remark V.4.15

Observe that Condition (V.11) automatically holds when C^* has only one eigenvalue, which is the case for instance when C is a Jordan block, that is to say when our parabolic system has a **cascade** *structure.*

V.4.2 Null-controllability

Let us now study the null-controllability of (V.9). The usual Kalman matrix change of variable let us put the system in cascade form (observe that it is crucial here that the same diffusion operator appears in each equation.

To simplify the presentation we assume $n = 2$ and $m = 1$ and thus we consider the following cascade system

$$\begin{cases} \partial_t y_1 + \mathcal{A} y_1 = 0, & \text{in } (0,1) \\ \partial_t y_2 + \mathcal{A} y_2 + y_1 = 0, & \text{in } (0,1) \\ \quad y(t, x=1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \text{ and } y(t, x=0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} v(t). \end{cases} \tag{V.12}$$

The proof will rely on the moments method for which we need a generalized version of the results given in Section IV.1.2 and that we will present now.

V.4.2.1 More about biorthogonal families of real exponential type functions

We first generalize the definitions introduced in Section IV.1.1.3, following the formalism of generalized divided differences that we recall in Appendix, Section A.3

$$e[\lambda^{(j+1)}] \overset{\text{def}}{=} \left(t \mapsto \frac{(-t)^j}{j!} e^{-\lambda t} \right) \in L^2(0, \infty).$$

We can then formulate the suitable generalization of Theorem IV.1.9 for taking into account the multiplicity of the eigenvalues in our control problem. Its proof is postponed to Section V.4.2.3.

Theorem V.4.16 (Generalized biorthogonal families in finite horizon)

Let Λ be a family in $(0,+\infty)$ satisfying (IV.19)-(IV.20) and let $m \in \mathbb{N}^*$.
Then, for any $T > 0$, there exists a family $(q^l_{\lambda,T})_{\substack{\lambda\in\Lambda\\ l\in[\![0,m]\!]}}$ in $L^2(0,T)$ satisfying

$$(q^l_{\lambda,T}, e[\mu^{(j+1)}])_{L^2(0,T)} = \delta_{\lambda,\mu}\delta_{l,j}, \quad \forall \lambda,\mu \in \Lambda, l,j \in [\![0,m]\!],$$

and the estimate

$$\|q^l_{\lambda,T}\|_{L^2(0,T)} \leq K e^{\varepsilon(\lambda)\lambda}, \;\; \forall \lambda \in \Lambda, \forall l \in [\![0,m]\!],$$

where $\varepsilon : \mathbb{R} \to \mathbb{R}$ is a function such that $\lim_{s\to+\infty} \varepsilon(s) = 0$ that only depends on m, ρ and the remainder function R, and $K > 0$ is a constant depending only on T, m, ρ and R.
Moreover, if the counting function of Λ satisfies (IV.21) or (IV.22) then ε can be chosen as in IV.1.10.

V.4.2.2 Application to the null-controllability of (V.12)

Theorem V.4.17

For any initial data $y_0 \in (L^2(\Omega))^2$, and any $T > 0$, there exists a control $v \in L^2(0,T)$ such that the solution of (V.12) satisfies $y(T) = 0$. Moreover, we have the estimate

$$\|v\|_{L^2(0,T)} \leq C e^{\frac{C}{T}} \|y_0\|_{L^2},$$

where $C > 0$ does not depend on T.

We first give the proof of the controllability result. The control cost estimate will be discussed in Section V.4.2.4.
Proof (Existence of the control):

The spectrum of the adjoint operator $\mathcal{L}^* = \mathcal{A} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ is described in Proposition V.3.7 (with $c_{21} = 1$ here). Since, in the current setting we have $n_\lambda = 1$ and $I_\lambda(c_{21}) = I_\lambda(1) \neq 0$ for any λ, we deduce that for each $\lambda \in \Lambda$, there is, up to a constant, a single eigenfunction

$$\Phi_\lambda = \begin{pmatrix} \phi_\lambda \\ 0 \end{pmatrix},$$

and an associated generalized eigenvector

$$\tilde{\Phi}_\lambda = \begin{pmatrix} 0 \\ \phi_\lambda \end{pmatrix},$$

and we observe that

$$\mathcal{B}^*\Phi_\lambda = \phi'_\lambda(0), \qquad \mathcal{B}^*\tilde{\Phi}_\lambda = 0. \tag{V.13}$$

We can immediately compute

$$\begin{cases} e^{-t\mathcal{L}^*}\Phi_\lambda = e^{-t\lambda}\Phi_\lambda, \\ e^{-t\mathcal{L}^*}\tilde{\Phi}_\lambda = e^{-t\lambda}(\tilde{\Phi}_\lambda - t\Phi_\lambda). \end{cases}$$

In that case it is cleat that the family $\{\Phi_\lambda, \tilde{\Phi}_\lambda, \lambda \in \Lambda\}$ is an Hilbert basis of $(L^2(\Omega))^2$ (we actually only need that it is complete) and therefore a function $v \in L^2(0,T)$ is a null-control for our problem if and only if it satisfies the following moments equations

$$\begin{cases} e^{-T\lambda}\langle y_0, \Phi_\lambda\rangle_E = \displaystyle\int_0^T v(s)e^{-\lambda(T-s)}\mathcal{B}^*\Phi_\lambda\, ds \\ e^{-T\lambda}\langle y_0, \tilde{\Phi}_\lambda - T\Phi_\lambda\rangle_E = \displaystyle\int_0^T v(s)e^{-\lambda(T-s)}\mathcal{B}^*(\tilde{\Phi}_\lambda - (T-s)\Phi_\lambda) ds. \end{cases}$$

Those equations can be simplified using the definitions of Φ_λ, $\tilde{\Phi}_\lambda$ and (V.13) as follows

$$\begin{cases} \dfrac{e^{-T\lambda}}{\phi_\lambda'(0)}\langle y_{0,1},\phi_\lambda\rangle_{L^2} = \displaystyle\int_0^T v(s)e^{-\lambda(T-s)}\,ds, \\ \dfrac{e^{-T\lambda}}{\phi_\lambda'(0)}\Big(\langle y_{0,1},\phi_\lambda\rangle_{L^2} - T\langle y_{0,2},\phi_\lambda\rangle_{L^2}\Big) = \displaystyle\int_0^T v(s)[-(T-s)]e^{-\lambda(T-s)}ds \end{cases}$$

Setting $u(t) := v(T-t)$, we are now looking for a function u that solves the following moment problem

$$\begin{cases} \displaystyle\int_0^T u(t)e_t[\lambda]\,dt = \omega^0_{\lambda,T,y_0}, \\ \displaystyle\int_0^T u(t)e_t[\lambda^{(2)}]dt = \omega^1_{\lambda,T,y_0}, \end{cases}$$

where

$$\omega^0_{\lambda,T,y_0} \overset{\text{def}}{=} \frac{e^{-T\lambda}}{\phi_\lambda'(0)}\langle y_{0,1},\phi_\lambda\rangle_{L^2}, \quad \text{and} \quad \omega^1_{\lambda,T,y_0} \overset{\text{def}}{=} \frac{e^{-T\lambda}}{\phi_\lambda'(0)}\Big(\langle y_{0,2},\phi_\lambda\rangle_{L^2} - T\langle y_{0,1},\phi_\lambda\rangle_{L^2}\Big).$$

This moment problem can now be solved by using the generalized biorthogonal family given by Theorem V.4.16 (with $m = 1$ in the present case) as follows

$$u(t) = \sum_{\lambda\in\Lambda}\left(\omega^0_{\lambda,T,y_0}q^0_{\lambda,T}(t) + \omega^1_{\lambda,T,y_0}q^1_{\lambda,T}(t)\right).$$

Indeed, by the estimates given in the Theorem and the definition of the terms $\omega_\bullet$ we find the convergence of the series in $L^2(0,T)$, exactly as we did in the proof of Theorem IV.1.11. Moreover, it clearly satisfies the required moment problem by construction of the biorthogonal family.

■

V.4.2.3 Proof of Theorem V.4.16

As we did in Section IV.1.2, we will start by proving the result with $T = +\infty$, then we will present the restruction argument to justify the construction in the case $T < +\infty$.

Infinite time horizon. Assume that Λ satisfies (IV.20) and (IV.19) and let $\lambda \in \Lambda$. For any $h > 0$ we introduce the new family

$$\Lambda_h \overset{\text{def}}{=} \bigcup_{j=0}^{m}(\Lambda + jh),$$

and the subset

$$L_h \overset{\text{def}}{=} \bigcup_{j=0}^{m}(\Lambda\setminus\{\lambda\} + jh).$$

Lemma V.4.18

Assume that $h < \frac{\rho}{2m}$, then the family Λ_h satisfies the weak gap condition (IV.25) with the gap $\rho/2$ and $p = m+1$.
Moreover, Λ_h has a remainder function $\tilde{R}$ which only depends on R and m.

Proof :

- Assume that (IV.25) does not hold for Λ_h with the given parameters. Then, for some $\mu > 0$ we have
$$\#\big(\Lambda \cap [\mu, \mu+\rho/2)\big) > m+1.$$
In particular there are two elements in $[\mu,\mu+\rho)$ that are of the form $\lambda + ih$ and $\lambda' + jh$ with $\lambda \neq \lambda'$ and $i, j \in [\![0,m]\!]$. In particular we have
$$|(\lambda+ih)-(\lambda'+jh)| < \rho/2,$$
and thus
$$|\lambda-\lambda'| < \rho/2 + |i-j|h \leq \rho/2 + mh < \rho.$$
This is a contradiction with (IV.20).

- Let $r > 0$, we have
$$\begin{aligned}\sum_{\substack{\sigma\in\Lambda_h\\ \sigma>r}} \frac{1}{\sigma} &= \sum_{i=0}^m \sum_{\substack{\lambda\in\Lambda\\ \lambda+ih>r}} \frac{1}{\lambda+ih}\\ &= \sum_{i=0}^m \sum_{\substack{\lambda\in\Lambda\\ \lambda>r}} \frac{1}{\lambda+ih} + \sum_{i=0}^m \sum_{\substack{\lambda\in\Lambda\\ r\geq\lambda+ih>r}} \frac{1}{\lambda+ih}\\ &\leq \sum_{i=0}^m \sum_{\substack{\lambda\in\Lambda\\ \lambda>r}} \frac{1}{\lambda} + \sum_{i=0}^m \sum_{\substack{\lambda\in\Lambda\\ r\geq\lambda>r-ih}} \frac{1}{\lambda+ih}\\ &\leq (m+1)R(r) + \sum_{i=0}^m \sum_{\substack{\lambda\in\Lambda\\ r\geq\lambda>r-ih}} \frac{1}{\lambda+ih}.\end{aligned}$$
  - If $r < \frac{1}{R(0)}$ then in particular (A.14) we have $r < \inf \Lambda$, so that the second term above is 0.
  - If $r \geq \frac{1}{R(0)}$ then we can simply bound the second term by $\sum_{i=0}^m \frac{i}{r}$.

All in all, we got that the function
$$\tilde{R}(r) \stackrel{\text{def}}{=} (m+1)R(r) + \frac{m^2}{2}\min\left(R(0), \frac{1}{r}\right),$$
if a remainder function for Λ_h, which proves the claim.

■

For any $\sigma \in (0,+\infty)$ with $\sigma \notin L_h$, we define now
$$p_h[\sigma] \stackrel{\text{def}}{=} e[\sigma] - \pi_{L_h} e[\sigma],$$
and we set
$$P_{\lambda,h} \stackrel{\text{def}}{=} \Big\{p_h[\lambda], p_h[\lambda,\lambda+h], \ldots, p_h[\lambda, ..., \lambda+mh]\Big\}.$$

Proposition V.4.19

The minimal biorthogonal family in $L^2(0,+\infty)$ to the family $P_{\lambda,h}$, denoted by $(q^l_{\lambda,h})_{l\in[\![0,m]\!]}$, satisfies
$$\|q^l_{\lambda,h}\|_{L^2(0,+\infty)} \leq C\lambda^{m+\frac{1}{2}} e^{\varepsilon(\lambda-m\rho)\lambda}, \quad \forall h < h_0(\lambda), \tag{V.14}$$
for some $h_0(\lambda)$ depending only on λ, $C > 0$ depending only on m and ε a decreasing function such that $\lim_{r\to+\infty} \varepsilon(r) = 0$ *depending only on R, ρ and m.*

Proof :

Let us introduce the functions

$$f_h[\sigma] \stackrel{\text{def}}{=} \frac{p_h[\sigma]}{W_{L_h}(\sigma)}.$$

Using Propositions A.4.13 and A.4.14 we obtain that for any $\sigma, \sigma' \notin L_h$, we have

$$(p_h[\sigma], p_h[\sigma'])_{L^2(0,+\infty)} = \frac{W_{L_h}(\sigma)W_{L_h}(\sigma')}{\sigma+\sigma'},$$

and thus

$$(f_h[\sigma], f_h[\sigma'])_{L^2(0,+\infty)} = \frac{1}{\sigma+\sigma'}.$$

In particular, it appears that

$$(f_h[\sigma], f_h[\sigma'])_{L^2(0,+\infty)} = (e[\sigma], e[\sigma'])_{L^2(0,\infty)}. \tag{V.15}$$

We consider the (linearly independent) family

$$F_{\lambda,h} \stackrel{\text{def}}{=} \left\{(2\lambda)^{1/2} f_h[\lambda], (2\lambda)^{1+1/2} f_h[\lambda,\lambda+h], \dots, (2\lambda)^{m+1/2} f_h[\lambda, ..., \lambda+mh]\right\},$$

that spans the same space as $P_{\lambda,h}$.

By using (V.15) we get for any $k, l \in [\![0, m]\!]$ that

$$\begin{aligned}\Big((2\lambda)^{k+1/2} f_h[\lambda, ..., \lambda+kh], &(2\lambda)^{l+1/2} f_h[\lambda, ..., \lambda+lh]\Big)_{L^2(0,+\infty)} \\ &= (2\lambda)^{k+l+1} \left(e[\lambda, ..., \lambda+kh], e[\lambda, ..., \lambda+lh]\right)_{L^2(0,+\infty)} \\ &\xrightarrow[h\to 0]{} (2\lambda)^{k+l+1} \left(e[\lambda^{(k+1)}], e[\lambda^{(l+1)}]\right)_{L^2(0,+\infty)} \\ &= (2\lambda)^{k+l+1} \int_0^{+\infty} \frac{(-t)^k}{k!} e^{-\lambda t} \frac{(-t)^l}{l!} e^{-\lambda t}\, dt \\ &= \int_0^{+\infty} \frac{(-t)^{k+l}}{k!l!} e^{-t}\, dt.\end{aligned}$$

It follows that the Gram matrix of $F_{\lambda,h}$ converges, when $h \to 0$ towards a matrix which is independent of λ and which is, in fact, nothing but the Gram matrix of the family $t \mapsto (-t)^k/k!$ in the weighted space $L^2(0, +\infty, e^{-t}\, dt)$.

Therefore, by Propositions A.4.10 and A.4.9, there exists $h_0(\lambda) > 0$, such that for any $h < h_0$, the minimal biorthogonal family of $F_{\lambda,h}$, denoted by $(g_{\lambda,h,i})_{i\in[\![0,m]\!]}$ satisfies the uniform bound

$$\|g_{\lambda,h,i}\|_{L^2(0,+\infty)} \leq C, \tag{V.16}$$

where $C > 0$ depends only on m.

We set now

$$q^l_{\lambda,h} \stackrel{\text{def}}{=} \sum_{j=l}^{m} \left(\frac{1}{W_{L_h}}\right)[\lambda+lh, ..., \lambda+jh](2\lambda)^{j+1/2} g_{\lambda,h,j}.$$

It is clear that $q^l_{\lambda,h} \in \mathrm{Span}(P_{\lambda,h})$ and we compute the following inner product

$$\begin{aligned}(p_h[\lambda, ..., \lambda+kh], &q^l_{\lambda,h})_{L^2(0,+\infty)} \\ &= \sum_{i=0}^{k} W_{L_h}[\lambda+ih, ..., \lambda+kh]\left(f_h[\lambda, ..., \lambda+ih], q^l_{\lambda,h}\right)_{L^2(0,+\infty)} \\ &= \sum_{i=0}^{k}\sum_{j=l}^{m} W_{L_h}[\lambda+ih, ..., \lambda+kh]\left(\frac{1}{W_{L_h}}\right)[\lambda+lh, ..., \lambda+jh] \\ &\qquad\qquad \times (2\lambda)^{j-i} \underbrace{\left((2\lambda)^{i+1/2} f_h[\lambda, ..., \lambda+ih], g_{\lambda,h,j}\right)_{L^2(0,+\infty)}}_{=\delta_{i,j}}.\end{aligned}$$

In the case where $k < l$, the sum above is zero since it is not possible that $i = j$. Assume now that $k \geq l$, thanks to the Leibniz formula (Proposition A.3.4), the sum reduces to

$$\begin{aligned}(p_h[\lambda, ..., \lambda + kh], q^l_{\lambda,h})_{L^2(0,+\infty)} &= \sum_{i=l}^{k} W_{L_h}[\lambda + ih, ..., \lambda + kh] \left(\frac{1}{W_{L_h}}\right)[\lambda + lh, ..., \lambda + ih] \\ &= \left(\frac{1}{W_{L_h}} W_{L_h}\right)[\lambda + lh, ..., \lambda + kh] \\ &= 1[\lambda + lh, ..., \lambda + kh] \\ &= \delta_{k,l}.\end{aligned}$$

This proves that $(q^l_{\lambda,h})_{l\in[\![0,m]\!]}$ is indeed the minimal biorthogonal family to $P_{\lambda,h}$.

Moreover, thanks to (V.16) , we have the explicit bound

$$\|q^l_{\lambda,h}\|_{L^2(0,+\infty)} \leq C\lambda^{m+\frac{1}{2}} \max_{j\in[\![l,m]\!]} \left| \left(\frac{1}{W_{L_h}}\right)[\lambda + lh, ..., \lambda + jh] \right|.$$

Thanks to the Lagrange theorem (Proposition A.3.3) and to the estimates given in Corollary IV.1.16, we finally get the uniform bound (V.14). Here we have used that L_h satisfies the assumptions (IV.25) and (IV.19) uniformly with respect to h, thanks to Lemma V.4.18.

The proof is complete. ■

We can now terminate the proof. Let $\mu \in \Lambda$.

- If $\mu \neq \lambda$, then $\mu + ih \in L_h$ for any $i \in [\![0, m]\!]$, and thus by construction we have

$$(q^l_{\lambda,h}, e[\mu + ih])_{L^2(0,+\infty)} = 0,$$

  which gives, by linear combinations,

$$(q^l_{\lambda,h}, e[\mu, ..., \mu + kh])_{L^2(0,+\infty)} = 0, \quad \forall k \in [\![0, m]\!].$$

- If $\mu = \lambda$, still by construction, we have

$$\begin{aligned}(q^l_{\lambda,h}, e[\lambda, ..., \lambda + kh])_{L^2(0,+\infty)} &= (q^l_{\lambda,h}, e[\lambda, ..., \lambda + kh] - \pi_{L_h} e[\lambda, ..., \lambda + kh])_{L^2(0,+\infty)} \\ &= (q^l_{\lambda,h}, p[\lambda, ..., \lambda + kh])_{L^2(0,+\infty)} \\ &= \delta_{k,l}.\end{aligned}$$

We have thus proved that

$$(q^l_{\lambda,h}, e[\mu, ..., \mu + kh])_{L^2(0,+\infty)} = \delta_{\lambda,\mu}\delta_{k,l}, \ \ \forall \mu \in \Lambda, \forall k, l \in [\![0, m]\!]. \tag{V.17}$$

Moreover, by Propositions A.3.3 and A.3.6 and the Lebesgue theorem, we easily see that for any $\mu > 0$ and any integer k, we have

$$e[\mu, ..., \mu + kh] \xrightarrow[h\to 0]{} e[\mu^{(k+1)}], \text{ strongly in } L^2(0, +\infty).$$

By (V.14) we see that, up to a subsequence, we can find a $q^l_\lambda \in L^2(0, +\infty)$ such that

$$q^l_{\lambda,h} \xrightharpoonup[h\to 0]{} q^l_\lambda, \quad \text{weakly in } L^2(0, \infty),$$

and that satisfies the same bound as in (V.14).

The claim is finally proved by performing a weak-strong limit in (V.17).

Restriction argument on $(0,T)$. The estimate of the restriction operator obtained in Proposition IV.1.17 can be easily extended to the present case by replacing $\mathcal{E}_\Lambda$ by

$$\mathcal{E}_\Lambda^m \stackrel{\text{def}}{=} \{e[\lambda^{(k+1)}], \quad \lambda \in \Lambda,\ k \in [\![0,m]\!]\},$$

and E_Λ (resp. $F_{\Lambda,T}$ and $F_{\Lambda,\infty}$) by E_Λ^m (resp. $F_{\Lambda,T}^m$ and $E_{\Lambda,\infty}^m$) accordingly.

V.4.2.4 Control cost estimate.

Refined estimate. Assuming that the counting function of Λ satisfies (IV.21) (in addition to the gap assumption (IV.20)) we can also extend Theorem IV.1.18 to obtain a sharp estimate of the restriction operator as a function of time.

Theorem V.4.20

There exists $C_{11} > 0$ (depending only on $\alpha, \bar{N}$ and on m) such that

$$\|P\|_{L^2(0,+\infty)} \le C_{11} e^{C_{11} T^{-\frac{\alpha}{1-\alpha}}} \|P\|_{L^2(0,T)}, \quad \forall P \in \mathcal{E}_\Lambda^m.$$

Proof :

Let $P \in \mathcal{E}_\Lambda^m$ that we write

$$P = \sum_{j=0}^m \sum_{\lambda\in\Lambda} a_\lambda^j e[\lambda^{(j+1)}],$$

where only a finite number of coefficients $(a_\lambda^j)_{j,\lambda}$ are non zero. For $h > 0$ we define

$$P_h = \sum_{j=0}^m \sum_{\lambda\in\Lambda} a_\lambda^j e[\lambda, ..., \lambda + jh] \in \mathcal{E}_{\Lambda_h}.$$

It is straightforward to see that the counting function N_h of Λ_h satisfies

$$N_h(r) \le mN(r), \ \forall r > 0,$$

and thus

$$N_h(r) \le m\bar{N} r^\alpha, \ \forall r > 0.$$

This estimate being uniform in h we can apply Theorem IV.1.18 to P_h so that for a $C > 0$, independent of h, we have

$$\|P_h\|_{L^2(0,+\infty)} \le C e^{CT^{-\frac{\alpha}{1-\alpha}}} \|P_h\|_{L^2(0,T)}. \tag{V.18}$$

The conclusion follows by passing to the limit as $h \to 0$ in this estimate since, as we have already seen, we have

$$P_h \xrightarrow[h\to 0]{} P, \quad \text{in } L^2(0,\infty).$$

■

Using the same arguments as in Section IV.1.2.5 based on Theorems V.4.16 and V.4.20, the control cost estimate given in Theorem V.4.17 is now straightforward.

Appendix A

Appendices

A.1 Non-autonomous linear ODEs. Resolvant

We consider a linear, non autonomous and homogeneous ODE of dimension n as follows

$$\begin{cases} y'(t) + A(t)y(t) = f(t), \\ y(0) = y_0, \end{cases} \tag{A.1}$$

It can be proved that there exists a unique map $(t,s) \in \mathbb{R}\times\mathbb{R} \mapsto R(t,s) \in M_n(\mathbb{R})$ called the resolvant that satisfies

$$\begin{cases} \frac{d}{dt}R(t,t_0) + A(t)R(t,t_0) = 0, \\ R(t_0,t_0) = \mathrm{Id}. \end{cases}$$

This maps satisfies the group property

$$R(t_1,t_2)R(t_2,t_3) = R(t_1,t_3), \quad \forall t_1,t_2,t_3 \in \mathbb{R}.$$

With this definition, the unique solution to the problem (A.1), is given by the Duhamel formula

$$y(t) = R(t,0)y_0 + \int_0^t R(t,s)f(s)\,ds.$$

Example A.1.1 (Autonomous case)

When $A(t) = A$ does not depend on time, we can check that

$$R(t,s) = e^{-(t-s)A},$$

and the above formula becomes

$$y(t) = e^{-tA}y_0 + \int_0^t e^{-(t-s)A}f(s)\,ds.$$

A.2 Linear ODEs with integrable data

Consider the following system of ODEs, with $A \in M_n(\mathbb{R})$ independent of time and $f \in L^1(0,T,\mathbb{R}^n)$,

$$\begin{cases} y'(t) + Ay(t) = f(t), \\ y(0) = y_0, \end{cases}$$

The usual Cauchy theorem applies (with minor adaptation related to the fact that, because of the non regularity of f, the solution y may not be of class $\mathcal{C}^1$) and gives a unique solution y.

Let us prove that the linear solution map

$$\Phi : (y_0, f) \in \mathbb{R}^n \times L^1(0,T,\mathbb{R}^n) \mapsto y \in \mathcal{C}^0([0,T],\mathbb{R}^n),$$

is continuous. The Duhamel formula gives

$$y(t) = e^{-tA}y_0 + \int_0^t e^{-(t-s)A} f(s)\,ds,$$

and by taking the norm, for a given $t \in [0,T]$, we get

$$\|y(t)\| \leq e^{t\|A\|}\|y_0\| + \int_0^t e^{(t-s)\|A\|}\|f(s)\|\,dt \leq C_T(\|y_0\| + \int_0^T \|f(s)\|\,ds).$$

Which proves that

$$\|y\|_{\mathcal{C}^0([0,T],\mathbb{R}^n)} \leq C_T(\|y_0\| + \|f\|_{L^1(0,T,\mathbb{R}^n)}).$$

A.3 Divided differences

A.3.1 Definition and basic properties

Let V be a real vector space, $n \in \mathbb{N}$ and $x_1, \ldots, x_n \in \mathbb{R}$. Assume that $x_1, \ldots, x_n$ are pairwise distinct (see Section A.3.2 for a generalization). Let $f_1, \ldots, f_n \in V$ be given.

Definition A.3.2

The divided differences are defined by

$$f[x_i] \stackrel{\text{def}}{=} f_i,\ \forall i \in [\![1,n]\!],$$

and then recursively for any $k \in [\![2,n]\!]$, for any pairwise distinct $i_1, \ldots, i_k \in [\![1,n]\!]$, by

$$f[x_{i_1}, \ldots, x_{i_k}] \stackrel{\text{def}}{=} \frac{f[x_{i_1}, \ldots, x_{i_{k-1}}] - f[x_{i_2}, \ldots, x_{i_k}]}{x_{i_1} - x_{i_k}}.$$

If $f : \mathbb{R} \to V$ is a given function it will be implicitely assumed that $f_i = f[x_i] = f(x_i)$. It can be proved that the divided differences are symmetric functions of their arguments.

Proposition A.3.3 (Lagrange theorem)

Assume that $V = \mathbb{R}$ and that $f \in C^{n-1}(\mathrm{Conv}\{x_1, \ldots, x_n\})$. For any $k \in [\![1,n]\!]$, for any pairwise distinct $i_1, \ldots, i_k \in [\![1,n]\!]$, there exists a $z \in \mathrm{Conv}\{x_{i_1}, \ldots, x_{i_k}\}$ such that

$$f[x_{i_1}, \ldots, x_{i_k}] = \frac{f^{(k-1)}(z)}{(k-1)!}.$$

We recall a simple way to compute divided differences of a product which is known as the Leibniz rule.

Proposition A.3.4

Let $g : \mathbb{R} \to \mathbb{R}$ and $(gf)[x] \stackrel{\text{def}}{=} g(x)f[x]$. For any $k \in [\![1,n]\!]$, for any pairwise distinct $i_1, \ldots, i_k \in [\![1,n]\!]$,

$$(gf)[x_{i_1}, \ldots, x_{i_k}] = \sum_{j=1}^{k} g[x_{i_1}, \ldots, x_{i_j}]f[x_{i_j}, \ldots, x_{i_k}].$$

A.3.2 Generalized divided differences

Assume that V is a normed vector space. Let $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ be pairwise distinct real numbers and let $\alpha \in \mathbb{N}^n$ a multi-index such that $\alpha > 0$. To such a multi-index we associate elements of V that we gather in a $f_\alpha \in V^{|\alpha|}$ and that are indexed as follows

$$f_j^l, \;\; j \in [\![1, n]\!], l \in [\![0, \alpha_j - 1]\!].$$

Definition A.3.5

For any $\mu \in \mathbb{N}^n$ such that $\mu \leq \alpha$, we can define $f[x_1^{(\mu_1)}, \dots, x_n^{(\mu_n)}] \in V$ by using the following rules

$$f[x_1^{(\mu_1)}, \dots, x_n^{(\mu_n)}] = f_j^{\mu_j - 1}, \;\; \text{if } \mu_{j'} = 0 \text{ for all } j' \neq j, \tag{A.2}$$

and for all $j_1 \neq j_2$ and $\mu_{j_1} > 0, \mu_{j_2} > 0$

$$f[x_1^{(\mu_1)}, \dots, x_n^{(\mu_n)}] = \frac{f[\dots, x_{j_1}^{(\mu_{j_1}-1)}, \dots, x_{j_2}^{(\mu_{j_2})}, \dots] - f[\dots, x_{j_1}^{(\mu_{j_1})}, \dots, x_{j_2}^{(\mu_{j_2}-1)}, \dots]}{x_{j_1} - x_{j_2}}. \tag{A.3}$$

The above definition does not depend on the order in which we apply the second rule (A.3), moreover it is consistent with the standard divided differences definition in the following sense:

Proposition A.3.6

If $f : \mathbb{R} \to V$ is a smooth function, and if we set

$$f_j^l = \frac{f^{(l)}(x_j)}{l!}, \forall j \in [\![1, n]\!], \forall l \in [\![0, \alpha_j - 1]\!],$$

then for any $\mu \in \mathbb{N}^d$, $\mu \leq \alpha$, the associated generalized divided difference satisfies

$$f[x_1^{(\mu_1)}, \dots, x_n^{(\mu_n)}] = \lim_{h \to 0} f[y_1^h, ..., y_{|\mu|}^h],$$

where:

- *for each $p \in [\![1, |\mu|]\!]$, $\lim_{h\to 0} y_p^h$ exists and belongs to $\{x_1, ..., x_n\}$,*
- *for each $j \in [\![1, n]\!]$, there is exactly μ_j values of p such that $\lim_{h\to 0} y_p^h = x_j$.*

A.4 Biorthogonal families in a Hilbert space

A.4.1 Notation and basic result

Let A be any subset of H. We denote by π_A the orthogonal projection onto $\overline{\mathrm{Span}(A)}$ and we introduce the quantity

$$\delta(x, A) \stackrel{\text{def}}{=} d(x, \mathrm{Span}(A)) = d(x, \overline{\mathrm{Span}(A)}) = \|x - \pi_A x\|_H, \quad \forall x \in H. \tag{A.4}$$

We will see below a systematic way, based on linear algebra, to compute $\delta(x, A)$ when A is finite. The following elementary result gives us a way to compute $\delta(x, A)$ when A is countable by approaching A by a sequence of finite sets A_n.

Lemma A.4.7

Let A be any subset of H and $(A_n)_n$ an increasing sequence of subsets such that

$$A = \bigcup_{n\geq 1} A_n. \tag{A.5}$$

For any $x \in H$, we have

$$\pi_{A_n} x \xrightarrow{n\to\infty} \pi_A x,$$

and in particular

$$\delta(x, A_n) \xrightarrow{n\to\infty} \delta(x, A).$$

Proof :

Let us define the operators $T_n \stackrel{\text{def}}{=} \pi_{A_n} - \pi_A$.

We have the standard estimate $\|T_n\| \leq 2$ from the properties of orthogonal projections. Moreover, thanks to (A.5) we know that for any $x \in \text{Span}(A)$ there exists a n_0 such that $x \in \text{Span}(A_n)$ for any $n \geq n_0$ so that

$$T_n x = 0, \quad \forall n \geq n_0,$$

and in particular

$$\lim_{n\to\infty} T_n x = 0, \quad \forall x \in \text{Span}(A), \tag{A.6}$$

For any $x \in H$, and $y \in \text{Span}(A)$ we can write

$$\|T_n x\|_H \leq \|T_n(x-y)\|_H + \|T_n y\|_H \leq 2\|x-y\|_H + \|T_n y\|_H,$$

and thus by (A.6), we get

$$\limsup_{n\to\infty} \|T_n x\|_H \leq 2\|x-y\|_H.$$

By density of $\text{Span}(A)$ into $\overline{\text{Span}(A)}$, we deduce that

$$\lim_{n\to\infty} T_n x = 0, \quad \forall x \in \overline{\text{Span}(A)}.$$

Moreover, by construction, for any $x \in \text{Span}(A)^\perp$ we have

$$\pi_{A_n} x = \pi_A x = 0,$$

and thus $T_n x = 0$ for any n. The claim is proved since

$$H = \overline{\text{Span}(A)} \oplus \text{Span}(A)^\perp.$$

■

A.4.2 Gram matrices. Gram determinants

For any finite subset $E = \{e_1, ..., e_n\} \subset H$, the Gram matrix of E is defined by

$$G_E \stackrel{\text{def}}{=} \Big((e_i, e_j)_H\Big)_{i,j\in[\![1,n]\!]},$$

and the associated (Gram) determinant is denotes by $\Delta_E \stackrel{\text{def}}{=} \det G_E$. Note that G_E depends on the numbering of the elements of E but not Δ_E.

Lemma A.4.8 (Linear independence characterization)

We have the following two properties.

1. *The family E is linearly independent if and only if*

$$\delta(e_i, E \setminus \{e_i\}) > 0, \quad \forall i \in [\![1, n]\!].$$

2. *The family E is linearly independent if and only if $\Delta_E \neq 0$.*

Proof :

1. Since E is finite, $\mathrm{Span}(E \setminus \{e_i\})$ is closed and it follows that

$$\delta(e_i, E \setminus \{e_i\}) > 0 \Longleftrightarrow e_i \notin \mathrm{Span}(E \setminus \{e_i\}),$$

which proves the claim.

2. Let $X = {}^t\big(x_1, \dots, x_n\big) \in \mathbb{R}^n$ and $x = \sum_{i=1}^{n} x_i e_i \in H$. By definition we have

$$(X, G_E X) = \sum_{i,j=1}^{n} x_i x_j (e_i, e_j)_H = \|x\|_H^2.$$

It follows that G_E is a positive symmetric matrix and that $\det G_E = 0$ if and only if 0 is an eigenvalue of G_E. Moreover, with the above notation, we have $G_E X = 0$ if and only if $x = 0$ and the claim is proved.

■

Proposition A.4.9

With the notation above, for any $x \in H \setminus E$, we have

$$\delta(x, E)^2 = \frac{\Delta_{E\cup\{x\}}}{\Delta_E}.$$

Note that for $x \in E$ we have $\delta(x, E) = 0$.

Proof :

We observe, by elementary operations on rows and columns, that $\Delta_{E\cup\{x\}} = \Delta_{E\cup\{x-\pi_E x\}}$. Moreover, since $x - \pi_E x$ is orthogonal to all the vectors $(e_i)_i$, this last Gram matrix has the following block-by-block form

$$G_{E\cup\{x-\pi_E x\}} = \begin{pmatrix} G_E & 0 \\ 0 & \|x - \pi_E x\|_H^2 \end{pmatrix},$$

and therefore we have

$$\Delta_{E\cup\{x-\pi_E x\}} = \|x - \pi_E x\|_H^2 \Delta_E,$$

which is the claimed formula. ■

Proposition A.4.10 (Bi-orthogonal family. Finite case)

Let $E = \{e_1, ..., e_n\} \subset H$ be a finite family in H.
The following two properties are equivalent.

1. *The family E is linearly independent.*
2. *There exists a finite family $F = \{f_1, ..., f_n\}$ of cardinal n such that*

$$(e_i, f_j)_H = \delta_{i,j}, \quad \forall i, j \in [\![1, n]\!]. \tag{A.7}$$

We say that F is a **biorthogonal family** *of E.*

If those two properties hold then there exists a unique such biorthogonal family such that $F \subset \mathrm{SpanE}$. It satisfies moreover the matrix equality

$$G_E G_F = \mathrm{Id},$$

and in particular we have

$$\|f_i\|_H = \frac{1}{\delta(e_i, E \setminus \{e_i\})}, \quad \forall i \in [\![1, n]\!]. \tag{A.8}$$

Remark A.4.11

If $\tilde{F}$ is any biorthogonal family of E in H, then the orthogonal projections $f_i = \pi_E \tilde{f}_i$ still satisfy (A.7) and belong to $\mathrm{Span}(E)$. Therefore it is the unique family F given in the proposition.
It follows that F is the minimal biorthogonal family to E in the sense that

$$\|f_i\|_H \leq \|\tilde{f}_i\|_H, \quad \forall i \in [\![1, n]\!].$$

Proof :

- Assume that F is a biorthogonal family of E and let $(\alpha_i)_{i \in [\![1,n]\!]} \subset \mathbb{R}$ such that

$$0 = \sum_{i=1}^{n} \alpha_i e_i.$$

For any $j \in [\![1, n]\!]$ we take the inner product of this equality with f_j and we get

$$0 = \sum_{i=1}^{n} \alpha_i (e_i, f_j)_H = \alpha_j.$$

This proves that E is linearly independent.

- Assume now that E is linearly independent. We will look for a family F in the following form

$$f_j = \sum_{k=1}^{n} a_{jk} e_k,$$

where the matrix $A = (a_{jk})_{j,k} \in M_n(\mathbb{R})$ has to be determined.
The equations (A.7) can be written for any $i, j \in [\![1, n]\!]$,

$$\begin{aligned}\delta_{ij} &= \sum_{k=1}^{n} a_{jk} (e_k, e_i)_H \\ &= (AG_E)_{ji}.\end{aligned}$$

This reduces to the matrix equation $AG_E = \mathrm{Id}$. Since E is linearly independent, we know that G_E is invertible and thus that there exists an unique matrix A (which appears to be symmetric) that satisfies our requirements. This proves existence and uniqueness of the biorthogonal family F. We can then compute

$$(f_i, f_j)_H = \sum_{k,l} a_{ik} a_{jl} (e_k, e_l)_H = (AG_E A)_{ij},$$

and since $AG_E = \mathrm{Id}$, we deduce that $G_F = A$.

Since $G_F = G_E^{-1}$ we can express G_F thanks to the cofactor matrix of G_E and in particular, for the diagonal coefficient $\|f_i\|_H^2$ of G_F, using that the associated cofactor of G_E is nothing but the Gram determinant $\Delta_{E\setminus\{e_i\}}$ we obtain

$$\|f_i\|_H^2 = \frac{\Delta_{E\setminus\{e_i\}}}{\Delta_E},$$

and thus (A.8) by Proposition A.4.9.

■

When E is an infinite family, the existence of a biorthogonal family is no more equivalent to the linear independence of E, and we need a slightly stronger assumption.

Proposition A.4.12 (Bi-orthogonal family. Infinite case)

Let E be any family of elements of H.
The following two propositions are equivalent.

1. *There exists a family $F = (f_e)_{e\in E} \subset H$ such that*

$$(\tilde{e}, f_e)_H = \delta_{e,\tilde{e}}, \quad \forall e, \tilde{e} \in E.$$

   Such a family is called a biorthogonal family to E.

2. *We have*

$$\delta(e, E \setminus \{e\}) > 0, \quad \forall e \in E. \tag{A.9}$$

If those properties hold, there is a unique such family F such that $F \subset \overline{\mathrm{Span}(E)}$ and it satisfies

$$\|f_e\|_H = \frac{1}{\delta(e, E \setminus \{e\})}, \quad \forall e \in E.$$

Proof :

- Assume that there exists a biorthogonal family F then for any $y \in \mathrm{Span}(E \setminus \{e\})$ we have

$$1 = (e, f_e)_H = (e - y, f_e)_H \le \|e - y\|_H \|f_e\|_H.$$

  Taking the infimum with respect to y, we get

$$1 \le \delta(e, E \setminus \{e\}) \|f_e\|_H.$$

- Conversely, assume (A.9) and define

$$f_e = \frac{1}{\delta(e, E \setminus \{e\})^2} (e - \pi_{E\setminus\{e\}} e).$$

By construction, if $\tilde{e} \in E \setminus \{e\}$ we have

$$(f_e, \tilde{e})_H = \frac{1}{\delta(e, E \setminus \{e\})^2}(e - \pi_{E\setminus\{e\}}e, \tilde{e})_H = 0,$$

and

$$(f_e, e)_H = \frac{1}{\delta(e, E \setminus \{e\})^2}(e - \pi_{E\setminus\{e\}}e, e)_H = \frac{1}{\delta(e, E \setminus \{e\})^2}(e - \pi_{E\setminus\{e\}}e, e - \pi_{E\setminus\{e\}}e)_H = 1.$$

The claim is proved.

■

A.4.3 Generalized Gram determinants

Let $E = \{e_1, ..., e_n\}$ and $F = \{f_1, ..., f_n\}$ two finite families of elements of H. We introduce the generalized Gram matrix

$$G_{E,F} \stackrel{\text{def}}{=} \Big((e_i, f_j)_H\Big)_{i,j\in[\![1,n]\!]},$$

and the associated Gram determinant is denoted $\Delta_{E,F} = \det G_{E,F}$.

With this definition we can find a useful generalization of Proposition A.4.9.

Proposition A.4.13

Let $E = (e_i)_{1\leq i\leq n}$ be a linearly independent family in H.
For any $x, y \in H \setminus E$ we have

$$(x - \pi_E x, y - \pi_E y)_H = \frac{\Delta_{E\cup\{x\},E\cup\{y\}}}{\Delta_E}.$$

Proof :

The proof is very similar to the one of Proposition A.4.9. We first use elementary operations on the columns of $G_{E\cup\{x\},E\cup\{y\}}$ to prove that

$$\Delta_{E\cup\{x\},E\cup\{y\}} = \Delta_{E\cup\{x\},E\cup\{y-\pi_E y\}},$$

then we use elementary operations on the rows of this matrix to get

$$\Delta_{E\cup\{x\},E\cup\{y\}} = \Delta_{E\cup\{x-\pi_E x\},E\cup\{y-\pi_E y\}}.$$

Since $x - \pi_E x$ and $y - \pi_E y$ are orthogonal to E, this generalized Gram matrix is block diagonal

$$G_{E\cup\{x-\pi_E x\},E\cup\{y-\pi_E y\}} = \begin{pmatrix} G_E & 0 \\ 0 & (x - \pi_E x, y - \pi_E y)_H \end{pmatrix}.$$

The claim is proved by computing the determinant. ■

A.4.4 Cauchy determinants

As an example of Gram determinant we will need to compute the Cauchy determinant, which is by definition the determinant of the Cauchy matrix

$$C_A \stackrel{\text{def}}{=} \left(\frac{1}{a_i+a_j}\right)_{i,j\in[\![1,n]\!]},$$

for any family $A = \{a_1, ..., a_n\}$ of n positive real numbers.

We will actually need to introduce a generalized version of that by considering two families $A = \{a_1, ..., a_n\}$ and $B = \{b_1, ..., b_n\}$ of n positive real numbers and the associated Cauchy matrix

$$C_{A,B} \stackrel{\text{def}}{=} \left(\frac{1}{a_i+b_j}\right)_{i,j\in[\![1,n]\!]}.$$

Let us recall the following explicit formula for this determinant.

Proposition A.4.14

For any n and any families A, B we have

$$\det C_{A,B} = \left(\prod_{i=1}^{n} \frac{1}{a_i+b_i}\right) \times \prod_{\substack{i,j\in[\![1,n]\!]\\ i<j}} \frac{(a_i-a_j)(b_i-b_j)}{(a_i+b_j)(a_j+b_i)}.$$

Proof :

Let us perform the proof by induction. For $n = 1$, the result is clear. Let us now assume $n \geq 2$ and we write

$$A = \tilde{A} \cup \{a_n\}, \text{ with } \tilde{A} = \{a_1, ..., a_{n-1}\},$$

$$B = \tilde{B} \cup \{b_n\}, \text{ with } \tilde{B} = \{b_1, ..., b_{n-1}\}.$$

In the definition of $\det C_{A,B}$ we perform row manipulations to cancel all the upper diagonal entries in the last column. We obtain that

$$\det C_{A,B} \stackrel{\text{def}}{=} \det \begin{pmatrix} M & 0 \\ \star & \frac{1}{a_n+b_n} \end{pmatrix}_{i,j\in[\![1,n]\!]},$$

where M is a $(n-1)\times(n-1)$ matrix whose entries are

$$m_{ij} = \frac{a_n-a_i}{b_n+a_i}\frac{b_n-b_j}{a_n+b_j}\frac{1}{a_i+b_j}, \quad \forall i,j \in [\![1,n-1]\!].$$

In other words we have

$$M = D_{A,B,1} C_{\tilde{A},\tilde{B}} D_{A,B,2},$$

where $D_{A,B,1}$ (resp. $D_{A,B,2}$) is a $(n-1)\times(n-1)$ diagonal matrices whose entries are $\frac{a_n-a_i}{b_n+a_i}$ (resp. $\frac{b_n-b_j}{a_n+b_j}$). Computing the determinant, it follows that

$$\det M = (\det C_{\tilde{A},\tilde{B}}) \prod_{i=1}^{n-1} \frac{(a_n-a_i)(b_n-b_i)}{(a_n+b_i)(b_n+a_i)},$$

and finally

$$\det C_{A,B} = (\det C_{\tilde{A},\tilde{B}}) \times \frac{1}{a_n+b_n} \prod_{i=1}^{n-1} \frac{(a_n-a_i)(b_n-b_i)}{(a_n+b_i)(b_n+a_i)}.$$

The claim follows by using the induction hypothesis. ∎

A.5 Sturm comparison theorem

Theorem A.5.15

Let I be an interval of $\mathbb{R}$, $\gamma \in \mathcal{C}^1(I)$, with $\gamma > 0$ and $q_1, q_2 \in \mathcal{C}^0(I)$. Let u_1 and u_2 be non trivial solutions to the differential equations

$$-\partial_x(\gamma(x)\partial_x u_1) + q_1(x)u_1 = 0, \text{ on } I,$$
$$-\partial_x(\gamma(x)\partial_x u_2) + q_2(x)u_2 = 0, \text{ on } I.$$

We assume that $q_1 \geq q_2$ in I. Then for any distinct zeros $\alpha < \beta$ of u_1 one the two following proposition holds

- *Either, there exists one zero of u_2 in the open interval (α, β).*
- *Or, u_1 and u_2 are proportional in $[\alpha, \beta]$, which implies in particular that $q_1 = q_2$ on $[\alpha, \beta]$.*

Proof :

The main needed ingredient is the Wronskian of u_1, u_2 defined as follows

$$W(x) = (\gamma\partial_x u_1)u_2 - u_1(\gamma\partial_x u_2),$$

whose derivative has the following expression, using the two equations satisfied by u_1 and u_2

$$W'(x) = (q_1 - q_2)u_1u_2. \tag{A.10}$$

Let $\alpha < \beta$ be two zeros of u_1 in I and assume that there is no zero of u_2 in (α, β). Without loss of generality we can assume that α and β are consecutive zeros of u_1. This means that we can change the sign of u_1 and u_2 in such a way that

$$u_1 > 0 \text{ and } u_2 > 0, \quad \text{in } (\alpha, \beta).$$

And since $u_1(\alpha) = u_1(\beta) = 0$, we necessarily have $\partial_x u_1(\alpha) > 0$ and $\partial_x u_1(\beta) < 0$.

We can now collect the following facts:

- We have $W(\alpha) = (\gamma\partial_x u_1(\alpha))u_2(\alpha) \geq 0$ and $W(\alpha) = 0$ if and only if $u_2(\alpha) = 0$.
- We have $W(\beta) = (\gamma\partial_x u_1(\beta))u_2(\beta) \leq 0$ and $W(\beta) = 0$ if and only if $u_2(\beta) = 0$.
- Since $q_1 \geq q_2$, and u_1, u_2 are positive in (α, β), we deduce from (A.10) that $W' \geq 0$ in (α, β) and in particular that W is non decreasing in $[\alpha, \beta]$.

The above three properties are only possible if W is identically zero in (α, β), and in particular $u_2(\alpha) = u_2(\beta) = 0$. It follows that we necessarily have $W' = 0$ in (α, β) which implies, from (A.10), that $q_1 = q_2$ on $[\alpha, \beta]$.

Therefore, u_1 and u_2 are solutions to the same equation on $[\alpha, \beta]$ and both vanish at α. It follows that u_1 and

$v = u_2 \frac{u_1'(\alpha)}{u_2'(\alpha)}$ solve the same linear Cauchy problem in $[\alpha, \beta]$ and thus are equal. The claim is proved. ■

Corollary A.5.16

Let I be an interval of $\mathbb{R}$, $\gamma \in \mathcal{C}^2(I)$, with $\gamma > 0$, $q \in \mathcal{C}^0(I)$ and $\lambda > 0$. Let u be a non trivial solutions to the differential equation

$$-\partial_x(\gamma(x)\partial_x u) + q(x)u = \lambda u, \text{ on } I.$$

Let $a < b$ two points in I. Then, if

$$\lambda \geq \|q\|_\infty + \left(\frac{4\pi}{b-a}\right)^2 \|\gamma\|_\infty + \frac{1}{2}\|\gamma''\|_\infty, \tag{A.11}$$

there exists two distinct zeros of u in $[a, b]$ denoted by α, β such that

$$|\alpha - \beta| \geq |a - b|/2.$$

Proof :

Let us introduce the function

$$w(x) = \sin\left((x-a)\frac{4\pi}{b-a}\right),$$

which satisfies the equation

$$-w'' = \left(\frac{4\pi}{b-a}\right)^2 w,$$

and that have the following two explicit zeros

$$w(a) = 0, \quad w\left(a + \frac{b-a}{4}\right) = 0.$$

Let us set $v = \sqrt{\gamma} w$ and observe that v has the same zeros as w. Moreover, a straightforward computation shows that v solves the equation

$$-\partial_x(\gamma \partial_x v) + \tilde{q} v = 0,$$

where we have defined

$$\tilde{q}(x) = \left[-\left(\frac{4\pi}{b-a}\right)^2 \gamma - \frac{\gamma''}{2} + \frac{1}{4}\frac{(\gamma')^2}{\gamma}\right].$$

By the assumption (A.11) on λ, we have for any $x \in [a, b]$

$$\begin{aligned}\tilde{q}(x) &\geq -\left(\frac{4\pi}{b-a}\right)^2 \|\gamma\|_\infty - \frac{1}{2}\|\gamma''\|_\infty \\ &\geq \|q\|_\infty - \lambda \\ &\geq q(x) - \lambda.\end{aligned}$$

Therefore, we can apply the comparison principle (Theorem A.5.15) to u and w and deduce that between any two zeros of w there is a zero of u. In particular, there exists a zero of u, in the interval $\left[a, a + \frac{b-a}{4}\right]$, that we call α.

By the exact same reasoning we find a zero of u in the interval $\left[b - \frac{b-a}{4}, b\right]$ that we call β and it is straightforward to check that $|\alpha - \beta| \geq |a - b|/2$. ■

A.6 Counting function and summation formulas

Let $\Lambda \subset (0,+\infty)$ be a locally finite family of positive numbers.

Definition A.6.17 (Counting function)

The counting function associated with the family Λ is defined by

$$N_\Lambda(r) \stackrel{\text{def}}{=} \#\big(\Lambda \cap (-\infty, r]\big).$$

If there is no ambiguity we shall simply call it N.

We will make use of the following summation formulas.

Proposition A.6.18

Let $f : (0,+\infty) \to \mathbb{R}$ be a $\mathcal{C}^1$ function. For any $0 \leq s < r$ we have the following formulas

$$\sum_{\substack{\lambda\in\Lambda \\ \lambda\leq r}} f(\lambda) = f(r)N(r) - \int_0^r f'(t)N(t)\,dt,$$

$$\sum_{\substack{\lambda\in\Lambda \\ s<\lambda\leq r}} f(\lambda) = f(r)N(r) - f(s)N(s) - \int_s^r f'(t)N(t)\,dt,$$

and

$$\sum_{\substack{\lambda\in\Lambda \\ \lambda>r}} f(\lambda) = -f(r)N(r) - \int_r^{+\infty} f'(t)N(t)\,dt,$$

provided that the sum or the integral converges.

We assume now that

$$\sum_{\lambda\in\Lambda} \frac{1}{\lambda} < +\infty, \tag{A.12}$$

and we define the following notion.

Definition A.6.19 (Remainder function)

A function $R : \mathbb{R} \to [0,+\infty)$ is called **a remainder function** *for the family Λ, if it satisfies*

$$\lim_{r\to\infty} R(r) = 0,$$

and

$$\sum_{\substack{\lambda\in\Lambda \\ \lambda>r}} \frac{1}{\lambda} \leq R(r), \quad \forall r \in \mathbb{R}.$$

Proposition A.6.20

Assume (A.12) and let R be a remainder function for Λ.

1. *For any $s < r$ we have*
$$N(r) - N(s) \leq rR(s). \tag{A.13}$$
In particular, we have
$$\inf \Lambda \geq \frac{1}{R(0)}, \tag{A.14}$$
$$N(r)/r \xrightarrow[r\to\infty]{} 0.$$

2. *For any $t > 0$ we have*
$$\sum_{\lambda\in\Lambda} e^{-\lambda t} \leq \frac{R(0)}{t}. \tag{A.15}$$

Proof :

1. The following quantity
$$\sum_{s<\lambda\leq r} \frac{1}{\lambda},$$
can bound from below by $1/r$ multiplied by the number of terms which is exactly $N(r) - N(s)$ and can be bounded from above by $R(s)$. This proves the first claim.

   Taking $s = 0$ and $r = \inf \Lambda$ in (A.13), we get
$$1 \leq (\inf \Lambda)R(0),$$
   since $N(0) = 0$ and $N(\inf \Lambda) \geq 1$.

   Now for any given s, the inequality (A.13) gives
$$\frac{N(r)}{r} \leq R(s) + \frac{N(s)}{r}, \quad \forall r > s.$$
   Taking the superior limit when $r \to \infty$, it follows
$$\limsup_{r\to\infty} \frac{N(r)}{r} \leq R(s).$$
   This being true for any s, we can take the limit as $s \to \infty$ to get the claim
$$\limsup_{r\to\infty} \frac{N(r)}{r} \leq 0.$$

2. We use Proposition A.6.18 and (A.13) to get the estimate
$$\begin{aligned}\sum_{\lambda\in\Lambda} e^{-\lambda t} &= \int_0^{+\infty} te^{-tr} N(r)\,dr \\ &= \frac{1}{t}\int_0^{+\infty} tre^{-tr}\frac{N(r)}{r} t\,dr \\ &\leq \frac{R(0)}{t}\int_0^{+\infty} e^{-r} r\,dr \\ &= \frac{R(0)}{t}.\end{aligned}$$

   The claim is proved.

■

Proposition A.6.21

Assume that, for some $0 < \alpha < 1$, and some $\bar{N} > 0$ we have

$$N(r) \leq \bar{N} r^\alpha, \ \ \forall r > 0.$$

Then, we have the following bound from below

$$\inf \Lambda \geq \bar{N}^{-\frac{1}{\alpha}}, \tag{A.16}$$

and the function

$$R(r) = \begin{cases} \dfrac{\bar{N}}{1-\alpha} r^{\alpha-1}, & \forall r > \bar{N}^{-\frac{1}{\alpha}}, \\ \dfrac{\bar{N}^{\frac{1}{\alpha}}}{1-\alpha}, & \forall r \leq \bar{N}^{-\frac{1}{\alpha}}. \end{cases}$$

is a remainder function for Λ.

Proof :

Let us now prove (A.16). Since $\inf \Lambda \in \Lambda$, we obviously have

$$N(\inf \Lambda) \geq 1,$$

and therefore, with the assumption on N, we deduce

$$1 \leq \bar{N} (\inf \Lambda)^\alpha,$$

and the claim follows.

Note now that the assumption on N implies that (IV.19) holds necessarily. We apply the summation results of Proposition A.6.18 with $f(r) = \frac{1}{r}$ to obtain, for $r > \bar{N}^{-\frac{1}{\alpha}}$,

$$\begin{aligned} \sum_{\substack{\lambda \in \Lambda \\ \lambda > r}} \frac{1}{\lambda} &= -\frac{N(r)}{r} + \int_r^\infty \frac{1}{t^2} N(t)\, dt \\ &\leq \bar{N} \int_r^\infty t^{\alpha-2}\, dt \\ &\leq \frac{\bar{N}}{1-\alpha} r^{\alpha-1}. \end{aligned}$$

■

A.7 Generalized Tchebychev polynomials

Most of the material in this section is taken and adapted from [BE95, BE97]. We will only give here the results we need and we let the interested reader have a look at those references for a much more complete study of those properties.

Our main objective is to establish a Remez-type inequality

$$\|p\|_{L^\infty(0,\inf A)} \leq C \|p\|_{L^\infty(A)},$$

for any generalized polynomial

$$p(x) = \sum_{k=0}^{N-1} p_k x^{\lambda_k},$$

with $\lambda_k \in (0, +\infty)$, and any compact set A in $(0, +\infty)$. More precisely, we will identify the best constant C in this inequality and how it depends on A and on the set $\{0, \lambda_1, ..., \lambda_{N-1}\}$. The precise result will be given in Theorem A.7.31.

A.7.1 Interpolation in Müntz spaces

Let $L \subset [0, +\infty)$ be a finite subset of non negative numbers. In all this section we assume that

$$0 \in L,$$

and we set $N \overset{\text{def}}{=} \#L$. If $N \geq 2$ we define

$$\mu_L \overset{\text{def}}{=} \inf\left(L \setminus \{0\}\right),$$

to be the firt non zero element in L.

Let us define the following subset of $\mathcal{C}^0([0, +\infty), \mathbb{R})$ called, Müntz space,

$$M(L) \overset{\text{def}}{=} \text{Span}\{x \mapsto x^\lambda, \quad \lambda \in L\}.$$

We plot in Figure A.1 an example of such set

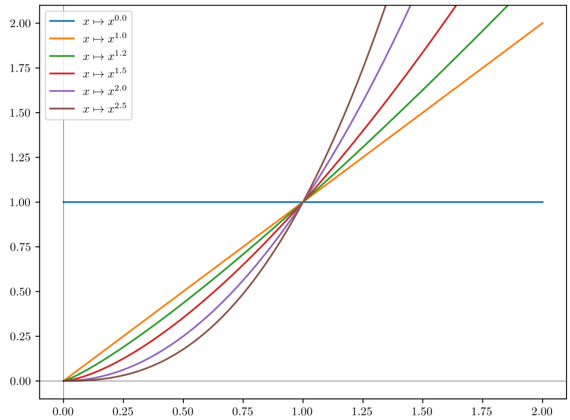

Figure A.1: Muntz space associated to the family $L = \{0, 1, 1.2, 1.5, 2, 2.5\}$.

Proposition A.7.22 (Interpolation properties)

The following properties hold

1. *0 is the only element of $M(L)$ that has at least N distinct zeros in $[0, +\infty)$.*
2. *If $f \in M(L)$ has exactly $N-1$ distinct zeros in $[0, +\infty)$, then the sign of f changes in the neighborhood of each of its zeros.*
3. *For any distinct points $x_1 < ... < x_N$ in $[0, +\infty)$, and any values $y_1, ..., y_N \in \mathbb{R}$, there exists a unique $f \in M(L)$ such that*

$$f(x_i) = y_i, \quad \forall i \in [\![1, N]\!].$$

We say that the set $M(L)$ is a Tchebychev system on $[0, +\infty)$.

Proof :

1. We prove the result by induction on N.
   - Let assume that $N = 1$, that is $L = \{0\}$. In that case, the functions in $M(L)$ are simply constants, and the claim is clear.
   - Assume that the result holds at rank N and let us consider a set L of cardinal $N + 1$.

     We assume that there exists a function $f \in M(L)$ that vanish at $N + 1$ distinct points $x_1 < ... < x_{N+1}$ in $[0, +\infty)$.

     We observe that $(x \mapsto xf'(x)) \in M(L \setminus \{0\})$ and that by the Rolle Theorem, f' has at least N distinct zeros in $[0, \infty)$. Thus, the function $x \mapsto g(x) \stackrel{\text{def}}{=} (xf'(x))/x^{\mu_L}$ belongs to $M(L \setminus \{0\} - \mu_L)$ and has at leat N distinct zeros. Since $L \setminus \{0\} - \mu_L$ contains 0 and has a cardinal N, the induction assumption shows that $g = 0$, which implies $f' = 0$ and thus $f = 0$.

2. We apply again the Rolle theorem that proves that f' has at least $N - 2$ zeros in $(0, +\infty)$ that are distinct from the zeros of f.

   We set $g(x) \stackrel{\text{def}}{=} (xf'(x))/x^{\lambda_2}$ and we observe that g is not identically 0, that it belongs to $M(L \setminus \{0\} - \lambda_2)$ and has at least $N - 2$ zeros in $(0, +\infty)$ that are distinct from the zeros of f. Therefore, g cannot have any other zero and in particular g cannot vanish at the zeros of f. This implies the f' cannot vanish at the zeros of f. In particular, f changes of sign in the neighborhood of each of its zero.

3. The linear map
$$\Phi : f \in M(L) \longmapsto (f(x_i))_i \in \mathbb{R}^N,$$
   is injective thanks to the first point and maps a space of dimension N into another space of dimension N. Therefore, Φ is a bijection, and the claim is proved.

■

Proposition A.7.23

Let $L = \{\lambda_1, ..., \lambda_N\}$ with $0 = \lambda_1 < ... < \lambda_N$.

1. *For any $0 \le x_1 < ... < x_N$ we have*
$$V_L(x_1, ..., x_N) \stackrel{\text{def}}{=} \det\left(x_i^{\lambda_j}\right)_{1 \le i,j \le N} > 0. \tag{A.17}$$
   If the points $x_1, ..., x_N$ are not ordered, the sign of the determinant is the signature of the corresponding ordering permutation.
2. *For any $k \le N - 1$ and any points $0 < w_1 < ... < w_k < +\infty$, there exists a $p \in M(L)$ such that*
$$\begin{cases} p(w_i) = 0, & \forall i \in [\![1, k]\!], \\ (-1)^i p(w) > 0, & \forall w \in (w_i, w_{i+1}),\ \forall i \in [\![0, k]\!], \end{cases}$$
   where, for convenience, we have set $w_0 \stackrel{\text{def}}{=} 0$ and $w_{k+1} \stackrel{\text{def}}{=} +\infty$.

Proof :

1. Les $0 \le y_1 < ... < y_N$ be another ordered set of points. For any $t \in [0, 1]$ we have $V_L(tx_1 + (1-t)y_1, ..., tx_N + (1-t)y_N) \neq 0$ by the previous proposition. By continuity, we deduce that $V_L(x_1, ..., x_N)$ and $V_L(y_1, ..., y_N)$ have the same sign. We fix the first $N - 1$ points and we let x_N go to $+\infty$. By developing the determinant along the last column, we see that
$$V_L(x_1, ..., x_N) \sim_{x_N \to \infty} V_{L'}(x_1, ..., x_{N-1})\, x_N^{(\max L)},$$

with $L' = L \setminus \{\max L\}$. This implies that $V_L(x_1, ..., x_N)$ has the same sign as $V_{L'}(x_1, ..., x_{N-1})$ and we conclude by induction.

2. We first remark that it is enough to consider the case $k = N - 1$. Indeed, if $k < N - 1$, we replace L by any subset $L' \subset L$ of cardinal $k + 1$ and containing 0, for which $M(L') \subset M(L)$.

   That being said, for a given sign $s \in \{-1, 1\}$ to be determined later, we define the function p as the following determinant
$$p(w) \stackrel{\text{def}}{=} s\, V_L(w, w_1, ..., w_{N-1}), \quad \forall w \in [0, +\infty).$$
   By developing the determinant along the first column we get that $p \in M(L)$ and moreover it is clear that $p(w_i) = 0$ for any $1 \le i \le N - 1$.

   The sign properties come from (A.17) and the choice of s.

■

Proposition A.7.24 (Elementary Lagrange interpolants)

For any set $X = \{x_1 < ... < x_N\} \subset (0, +\infty)$ of N distinct points there exists a unique family $(\Phi_{L,X,k})_{k \in [\![1,N]\!]} \subset M(L)$ such that
$$\Phi_{L,X,k}(x_j) = \delta_{j,k}, \quad \forall j, k \in [\![1, N]\!].$$
Moreover, if we set $x_0 = 0$ and $x_{N+1} = +\infty$, the sign of $\Phi_{L,X,k}$ is as follows

- $\Phi_{L,X,k} > 0$ *on* (x_{k-1}, x_{k+1}).
- $(-1)^{j+k+1}\, \Phi_{L,X,k} > 0$ *on* (x_j, x_{j+1}) *for* $j \in [\![0, k - 1]\!]$.
- $(-1)^{j+k}\, \Phi_{L,X,k} > 0$ *on* (x_j, x_{j+1}) *for* $j \in [\![k, N]\!]$.

Finally, we have
$$(-1)^{k+1} \Phi_{L,X,k}(0) > 0.$$

Let us show an example of such elementary Lagrange interpolants in Figure A.2

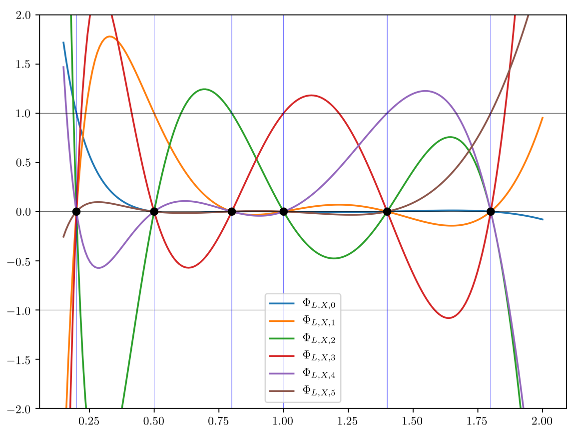

Figure A.2: Muntz space associated to the family $L = \{0, 1, 1.2, 1.5, 2, 2.5\}$ and the points $X = \{0.2, 0.5, 0.8, 1, 1.4, 1.8\}$.

Proof :

The existence and uniqueness of such a family of functions is just a consequence of the third point of Proposition A.7.22. It cannot have another zero in $[0,+\infty)$ since in that case we would have $\Phi_{L,X,k}=0$ everywhere by the first point of the same proposition.

From the second point of Proposition A.7.22, we know that $\Phi_{L,X,k}$ has a constant sign between two consecutive zeros and it changes of sign at each of those points. It is then straightforward to compute its sign by induction on each given interval starting from the fact that $\Phi_{L,X,k}(x_k)=1>0$.

We have seen above that $\Phi_{L,X,k}(0)\neq 0$ and therefore it has the same sign as $\Phi_{L,X,k}$ on $(0,x_1)$, which is $(-1)^{k+1}$.
■

Proposition A.7.25 (Comparison principle)

Let $X=\{x_1<...<x_N\},\tilde{X}=\{\tilde{x}_1<...<\tilde{x}_N\}$ *be two subsets of* $(0,+\infty)$ *made of* N *distinct points. Let* $k\in [\![1,N]\!]$ *and assume that*

$$\begin{cases} x_k\leq \tilde{x}_k, \\ |x_j-x_k|\geq |\tilde{x}_j-\tilde{x}_k|, \quad \forall j\in [\![1,N]\!], \end{cases}$$

then

$$|\Phi_{L,X,k}(0)|\leq |\Phi_{L,\tilde{X},k}(0)|,$$

with equality if and only if $X=\beta\tilde{X}$ *for some* $0<\beta\leq 1$.

Proof :

- Let us first define $\beta=\frac{x_k}{\tilde{x}_k}$, which is less than or equal to 1 by assumption. We define the set $\hat{X}=\beta\tilde{X}$. By construction, we have $\hat{x}_k=x_k$ and

$$|x_j-x_k|\geq |\hat{x}_j-\hat{x}_k|,\ \ \forall j\in [\![1,N]\!]. \tag{A.18}$$

Let us set $g(x)\stackrel{\text{def}}{=}\Phi_{L,\hat{X},k}(\beta x)$, for all $x\in[0,+\infty)$. By homogeneity we have that $g\in M(L)$ and satisfies

$$g(\tilde{x}_i)=\Phi_{L,\hat{X},k}(\beta\tilde{x}_i)=\Phi_{L,\hat{X},k}(\hat{x}_i)=\delta_{ik}.$$

Therefore $g=\Phi_{L,\tilde{X},k}$. In particular, we have

$$\Phi_{L,\tilde{X},k}(0)=\Phi_{L,\hat{X},k}(0).$$

The problem is thus reduced to proving that

$$|\Phi_{L,X,k}(0)|\leq |\Phi_{L,\hat{X},k}(0)|,$$

with equality if and only if $X=\hat{X}$. This will take several steps.

- We define the following sets:
  - For $i=[\![0,k]\!]$, we set $X^i\stackrel{\text{def}}{=}\{x_1,...,x_i,\hat{x}_{i+1},...,\hat{x}_N\}$.
    Note that $X^0=\hat{X}$ and that for $i\in[\![1,k]\!]$, we have $x_i\leq x_k$ and $\hat{x}_i\leq \hat{x}_k=x_k$ so that (A.18) gives

$$x_i\leq \hat{x}_i,$$

which implies

$$x_i<\hat{x}_{i+1}.$$

Therefore the points in X^i are distinct and well ordered.

- For $i = [\![k, N]\!]$, we set $X^i \stackrel{\text{def}}{=} \{x_1, ..., x_{k-1}, \hat{x}_k, ..., \hat{x}_{N+k-i}, x_{N+k-i+1}, ..., x_N\}$.
  Note that $X^N = X$ and that for $i \in [\![k, N-1]\!]$ we have $x_{N+k-i} \geq x_k$ and $\hat{x}_{N+k-i} \geq \hat{x}_k = x_k$ so that (A.18) gives
  $$x_{N+k-i} \geq \hat{x}_{N+k-i},$$
  so that
  $$\hat{x}_{N+k-i} < x_{N+k-i+1},$$
  and here also the points in X^i are distinct and well ordered.

Observe finally that both definition coincide for $i = k$ since $x_k = \hat{x}_k$ and that $X^k = X^{k-1}$. Moreover, by construction, for any i, X^i and X^{i+1} differ at most by one single point.

It thus remains to show that
$$|\Phi_{L,X^{i+1},k}(0)| \leq |\Phi_{L,X^i,k}(0)|, \quad \forall i \in [\![0, N-1]\!],$$
with equality if and only if $X^i = X^{i+1}$.

- Assume that $X^i \neq X^{i+1}$ for some i. We set $g \stackrel{\text{def}}{=} \Phi_{L,X^i,k} - \Phi_{L,X^{i+1},k}$, which is a function in $M(L)$, and we see that g cancels at the $N-1$ distinct points that are common to X^i and X^{i+1}. Let us analyse the sign of g at 0.

  - The function g cannot have any other zero. Indeed, in that case it would have N distinct zeros, and thus it would identically vanish. This would imply that $X^i = X^{i+1}$, a contradiction.
    This gives the equality case in our claim since $\Phi_{L,X^i,k}(0)$ and $\Phi_{L,X^{i+1},k}(0)$ have the same sign, which is $(-1)^{k+1}$ (see Proposition A.7.24).
  - By the second point of Proposition A.7.22 we know that g changes it sign at the neighborhood of each of its zeros. We are going to prove that
    $$(-1)^{k+1} g(0) > 0. \tag{A.19}$$
    We separate the analysis into two cases depending on the position of i with respect to $k-1$ (we recall that $i = k-1$ is not possible since in that case we would have $X^i = X^{i+1}$).
    * Case 1 : $i \in [\![0, k-1]\!]$:
      We compute
      $$g(x_{i+1}) = \Phi_{L,X^i,k}(x_{i+1}) - \Phi_{L,X^{i+1},k}(x_{i+1}) = \Phi_{L,X^i,k}(x_{i+1}), \tag{A.20}$$
      since x_{i+1} is a zero of $\Phi_{L,X^{i+1},k}$.
      By assumption on i we have $x_{i+1} < x_k$ and $\hat{x}_{i+1} < \hat{x}_k = x_k$, and we know that $x_{i+1} \neq \hat{x}_{i+1}$, so that (A.18) gives
      $$x_{i+1} < \hat{x}_{i+1},$$
      and thus $x_{i+1} \in (x_i, \hat{x}_{i+1})$. By (A.20), and Proposition A.7.24, we know that the sign of $g(x_{i+1})$ is such that
      $$(-1)^{i+k+1} g(x_{i+1}) > 0.$$
      Using that g changes it sign in the neighborhood of each of its zeros, we know that it changes it sign exactly i times in $[0, x_{i+1}]$ and we get (A.19).
    * Case 2 : $i \in [\![k-1, N]\!]$:
      We compute
      $$g(\hat{x}_{N+k-i}) = \Phi_{L,X^i,k}(\hat{x}_{N+k-i}) - \Phi_{L,X^{i+1},k}(\hat{x}_{N+k-i}) = -\Phi_{L,X^{i+1},k}(\hat{x}_{N+k-i}), \tag{A.21}$$
      since $\hat{x}_{N+k-i}$ is a zero of $\Phi_{L,X^i,k}$.

By assumption on i, we have $x_{N+k-i} > x_k$ and $\hat{x}_{N+k-i} > \hat{x}_k = x_k$, and we know that $x_{N+k-i} \neq \hat{x}_{N+k-i}$ so that (A.18) gives

$$\hat{x}_{N+k-i} < x_{N+k-i},$$

and thus $\hat{x}_{N+k-i} \in (\hat{x}_{N+k-i-1}, x_{N+k-i})$. By (A.21), and Proposition A.7.24, we know that the sign of $g(\hat{x}_{N+k-i})$ is such that

$$(-1)^{N-i} g(\hat{x}_{N+k-i}) > 0.$$

Using that g changes it sign in the neighborhood of each of its zeros, we know that it changes it sign exactly $N + k - i - 1$ times in $[0, \hat{x}_{N+k-i}]$ and we also get (A.19).

To conclude the proof, we write

$$|l_k^{L,X^i}(0)| - |l_k^{L,X^{i+1}}(0)| = (-1)^{k+1}(l_k^{L,X^i}(0) - l_k^{L,X^{i+1}}(0)) = (-1)^{k+1} g(0) > 0.$$

■

A.7.2 Best uniform approximation in Müntz spaces

Theorem A.7.26 (Best uniform approximation in Müntz spaces)

Let A be a (possibly infinite) compact subset of $[0, +\infty[$. We assume that $\#A \geq N + 1$. For any function $f \in \mathcal{C}^0(A)$, there is a unique $p \in M(L)$ such that

$$\|f - p\|_{L^\infty(A)} = \inf_{q \in M(L)} \|f - q\|_{L^\infty(A)}. \tag{A.22}$$

Moreover, p is the unique element in $M(L)$ such that $f - p$ equi-oscillates in at least $N + 1$ points of A. This means that there exists $x_1 < ... < x_{N+1}$, $x_i \in A$, and a sign $s = \pm 1$, such that

$$f(x_i) - p(x_i) = s(-1)^i \|f - p\|_{L^\infty(A)}, \quad \forall i \in [\![1, N+1]\!]. \tag{A.23}$$

Remark A.7.27

In the case where $\#A \leq N$, then by the interpolation property (Proposition A.7.22) shows that there exists $p \in M(L)$ such that $f = p$. Therefore, the best uniform approximation property is straightforward in that case.

Proof :

- Existence of at least one such best approximation is just a compactness argument related to the fact that, $M(L)$ is finite dimensional.
- Let us first show that any such best approximation p satisfies the claimed equi-oscillation property. We set $g \stackrel{\text{def}}{=} f - p$ and we assume that there exists a maximal equi-oscillating sequence for g in A of length $k < N + 1$ denoted by $x_1 < ... < x_k$ and we will obtain a contradiction.

  For any $i \in [\![1, k]\!]$ we introduce $C_i \stackrel{\text{def}}{=} \{x \in A, x_{i-1} \leq x \leq x_{i+1},\ \ g(x) = g(x_i)\}$, where we have conventionally set $x_0 = -\infty$ and $x_{k+1} = +\infty$. Since g is continuous on A, C_i is a closed subset of the compact set A, and in particular it's a compact set itself.

  We define the convex hull of C_i to be

  $$D_i \stackrel{\text{def}}{=} \text{conv} C_i = [x_i^-, x_i^+].$$

  We observe, by compacity, that $x_i^-, x_i^+ \in C_i$.

- We claim that the intervals D_i are disjoint. We are thus going to show that
$$x_i^+ < x_{i+1}^-, \ \ \forall i \in [\![1, k-1]\!].$$
By construction we know that $x_i \in C_i$ and $x_{i+1} \in C_{i+1}$ thus, we clearly get that
$$x_i^+, x_{i+1}^- \in [x_i, x_{i+1}],$$
and that
$$g(x_i^+) = g(x_i), \ \ g(x_{i+1}^-) = g(x_{i+1}),$$
that have two different signs. Hence, we deduce that $x_i^+ \neq x_{i-1}^-$.
Assume that for some i, we have $x_{i+1}^- < x_i^+$. It would imply that the sequence
$$x_1 < ... < x_i < x_{i+1}^- < x_i^+ < x_{i+1} < ... < x_k,$$
is an equi-oscillating sequence of length $k+2$, which is a contradiction with the maximality assumption for the original sequence. The claim is proved.
- We have thus built compact disjoint intervals $D_i = [x_i^-, x_i^+]$ surrounding each x_i such that
$$\|g\|_{L^\infty(A)} \geq s(-1)^i g(x) > -\|g\|_{L^\infty(A)}, \ \ \forall x \in A \cap D_i.$$
By continuity of g, we can find $\delta, \eta_1 > 0$ small enough such that
$$\|g\|_{L^\infty(A)} \geq s(-1)^i g(x) > -(1-\eta_1)\|g\|_{L^\infty(A)}, \ \ \forall x \in A \cap D_{i,\delta},$$
where $D_{i,\delta} =]x_i^- - \delta, x_i^+ + \delta[$ is the open δ-neighborhood of D_i.
- Introducing $D = \bigcup_{i=1}^{k} D_{i,\delta}$, we observe that, by construction, D contains all the points $x \in A$, where $|g(x)| = \|g\|_{L^\infty(A)}$. Therefore, for some $\eta_2 > 0$ small enough, we have
$$|g(x)| \leq (1-\eta_2)\|g\|_{L^\infty(A)}, \ \ \forall x \in A \setminus D,$$
since g is continuous on the compact set $A \setminus D$.
- We will now obtain a contradiction with the fact that p solves the best uniform approximation property (A.22).
For $i \in [\![1, k-1]\!]$ we set $w_i = \frac{x_i^+ + x_{i+1}^-}{2}$. By Proposition A.7.23, since $k \leq N$, there exists an element $\pi \in M(L)$ such that $\pi(w_i) = 0$ for any i, and such that $s(-1)^i \pi > 0$ on each $D_{i,\delta}$ and $\|\pi\|_{L^\infty(A)} \leq \|g\|_{L^\infty(A)}$.
We set $q = p + \eta\pi$ with $\eta > 0$ chosen such that $\eta < \min(\eta_1, \eta_2)$ and we will show that $\|f - q\|_{L^\infty(A)} < \|g\|_{L^\infty(A)}$.
Let $x \in A$.
  * If $x \in A \cap D_{i,\delta}$ for some i, then we write
  $$s(-1)^i(f-q)(x) = s(-1)^i(g(x) - \eta\pi(x)) = s(-1)^i g(x) - \eta s(-1)^i \pi(x),$$
  and by the sign property of π on $D_{i,\delta}$ we get
  $$-(1-\eta_1)\|g\|_{L^\infty(A)} - \eta\|g\|_{L^\infty(A)} \leq s(-1)^i(f-q)(x) < s(-1)^i g(x),$$
  so that we have the strict inequalities
  $$-\|g\|_{L^\infty(A)} < s(-1)^i(f-q)(x) < \|g\|_{L^\infty(A)},$$
  and consequently
  $$|(f-q)(x)| < \|g\|_{L^\infty(A)}.$$

* If $x \in A \setminus D$ we just write

$$|(f-q)(x)| = |g(x) - \eta\pi(x)| \leq |g(x)| + \eta|\pi(x)| \leq (1-\eta_2)\|g\|_{L^\infty(A)} + \eta\|g\|_{L^\infty(A)} < \|g\|_{L^\infty(A)}.$$

We have thus proved that $\|f-q\|_{L^\infty(A)} < \|f-p\|_{L^\infty(A)}$ which contradicts (A.22).

- We can now prove the uniqueness of the best uniform approximation in A.

  Let us define $d \stackrel{\text{def}}{=} \inf_{q\in M(L)} \|f-q\|_{L^\infty(A)}$ and we assume that $p_1, p_2 \in M(L)$ are such that $\|f-p_i\|_{L^\infty(A)} = d$. Then, by the triangle inequality, $p = \frac{p_1+p_2}{2}$ also satisfies $\|f-p\|_{L^\infty(A)} = d$. Thanks to the equi-oscillation property, there exists $N+1$ distinct points $x_1 < ... < x_{N+1}$ where

$$d = |f(x_i) - p(x_i)| = \frac{1}{2}|(f(x_i) - p_1(x_i)) + (f(x_i) - p_2(x_i))|,$$

  and since $|f(x_i) - p_1(x_i)|$, $|f(x_i) - p_2(x_i)|$ are both less than d, we obtain that necessarily $f(x_i) - p_1(x_i) = f(x_i) - p_2(x_i)$. We deduce that $p_1(x_i) = p_2(x_i)$ for any $i \in [\![1, N+1]\!]$. By the uniqueness property of the Tchebychev system, we conclude that $p_1 = p_2$.

- Finally we prove that any $p \in M(L)$ such that $f-p$ has the equi-oscillation property on A (we call $x_1 < ... < x_{N+1}$ the associated family of points) is indeed a best uniform approximation of f on A. To prove that claim, we assume that there exists $q \in M(L)$ such that

$$\|f-q\|_{L^\infty} < \|f-p\|_{L^\infty}.$$

  This implies in particular that

$$|(f(x_i) - p(x_i)) + (p(x_i) - q(x_i))| < \|f-p\|_{L^\infty(A)} = |f(x_i) - p(x_i)|,$$

  and since $f(x_i) - p(x_i)$ has the sign $s(-1)^i$, we deduce that the sign of $(p-q)(x_i)$ is $s(-1)^{i+1}$ (and of course this quantity cannot be zero). Hence, $p-q$ changes its sign at least $N+1$ times, and by the intermediate value theorem $p-q$ has at least N distinct zeros in $(0, +\infty)$. By point 1 of Proposition A.7.22, this implies $p = q$.

■

Proposition and Definition A.7.28 (Generalized Tchebychev polynomials)

Let A be a compact subset of $[0, +\infty)$ such that $\#A \geq N+1$. There exists a unique (up to a multiplicative factor) element in $M(L)$ that equi-oscillates in A at exactly N points.
We denote by $T_{L,A}$ the unique such function that, in addition, satisfies the normalisation properties

$$\|T_{L,A}\|_{L^\infty(A)} = 1,$$

$$T_{L,A}(\max A) > 0.$$

Moreover,

- *$T_{L,A}$ has exactly $N-1$ zeros in $[0, +\infty)$. They are all located in the open interval $(\inf A, \sup A)$.*
- *The map*

$$x \mapsto |T_{L,A}(x)|$$

  is decreasing on $[0, \inf A]$.

The function $T_{L,A}$ is called the generalized Tchebychev polynomial on the set A with respect to the family L.

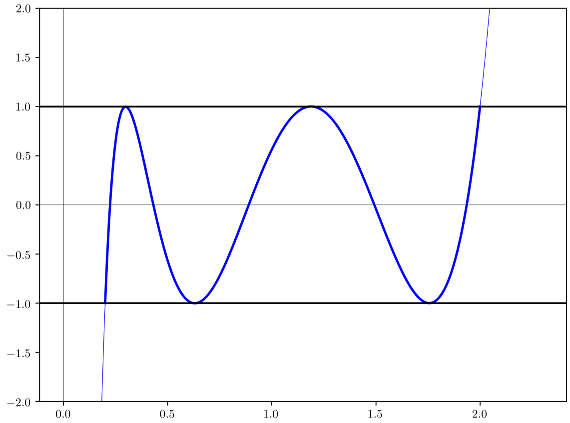

Figure A.3: The Tchebychev polynomial $T_{L,A}$ for $L = \{0, 1, 1.2, 1.5, 2, 2.5\}$ and $A = [0.2, 2]$.

We illustrate this definition in Figure A.3.

Proof :

If $L = \{0\}$, the result is straightforward (and $T_{L,A} = 1$).

Assume that $N > 1$ and let $\tilde{L} = L \setminus \mu_L$. We consider $\pi \in M(\tilde{L})$ the unique uniform best approximation of $x \mapsto x^{\mu_L}$ on A in $M(L')$ given by Theorem A.7.26. We know that the function $\tilde{T}(x) \stackrel{\text{def}}{=} x^{\mu_L} - \pi(x)$ belongs to $M(L)$ and equi-oscillates at least $\#L' + 1 = N$ times. Moreover, $\tilde{T}$ cannot equi-oscillate $N+1$ times because if it were the case $\tilde{T}$ would be the unique best uniform approximation of 0 on A in $M(L)$, and it will immediately imply that $\tilde{T} = 0$ on A which is not possible.

Note that the equi-oscillation property implies that $\tilde{T}$ has at least $N-1$ zeros in the open interval $I = (\inf A, \sup A)$. It is clear that $\tilde{T}$ cannot vanish on $[\sup A, +\infty)$ since in that case, the function would have N distinct zeros and thus will be identically equal to 0. Therefore, the normalisation conditions we consider are uniquely solvable.

Observe that, if $\inf A > 0$ we also have that $\tilde{T}$ cannot vanish on $[0, \inf A]$. Finally, if $\inf A = 0$, we also have $\tilde{T}(0) \neq 0$. Indeed, if we assume that $\tilde{T}(0) = 0$ and since we have $0 \in L$, we can easily see that $\tilde{T}$ actually belongs to $M(L \setminus \{0\})$. However, the only function in $M(L \setminus \{0\})$ that has at least N zeros in $(0, +\infty)$ is the function 0, which is a contradiction.

Finally, using Rolle's theorem, we know that $T'_{L,A}$ has at least $N-2$ zeros in $(\min A, \max A)$. Moreover, $(T_{L,A})' \in M(L \setminus \{0\})$ thus it cannot have another zero. In particular $(T_{L,A})'$ has a constant sign on $[0, \inf A)$ and $T_{L,A}$ does not vanish in this interval. The claim is proved. ■

Proposition A.7.29 (Maximality property of $T_{L,A}$)

Assume that $\inf A > 0$ *and let* $y \in [0, \inf A)$. *Then for any* $p \in M(L)$, *such that* $\|p\|_{L^\infty(A)} \leq 1$ *we have*

$$|p(y)| \leq |T_{L,A}(y)|.$$

Equivalently, we have

$$|p(y)| \leq |T_{L,A}(y)| \, \|p\|_{L^\infty(A)}, \quad \forall p \in M(L).$$

Proof :

The map $\Psi : p \in M(L) \mapsto |p(y)|$ is clearly continuous, thus it attains it maximum on the compact set $K = \{p \in M(L), \ \|p\|_{L^\infty(A)} \leq 1\}$.

It is clear that this maximum is achieved on a $p \in M(L)$ such that $\|p\|_{L^\infty(A)} = 1$.

Assume that p equi-oscillates exactly k times with $k < N$. As in the proof of A.7.26 we can build disjoint (ordered) open intervals $D_{i,\delta}$, $i = 1, ..., k$ such that

$$1 \geq s(-1)^i p(x) > -(1-\eta_1), \ \ \forall x \in A \cap D_{i,\delta},$$

for $D = \cup_i D_{i,\delta}$,

$$|p(x)| \leq 1 - \eta_2, \ \ \forall x \in A \setminus D.$$

For each $i \in [\![1, k-1]\!]$, we pick a set of point $w_{i+1/2}$ between $D_{i,\delta}$ and $D_{i+1,\delta}$ and we consider a $\pi \in M(L)$ such that

$$\begin{cases} \pi(w_{i+1/2}) = 0, & \forall i \in [\![1, k-1]\!], \\ \pi(y) = 0, \end{cases}$$

and

$$s(-1)^i \pi > 0, \ \ \text{on } D_{i,\delta}.$$

This is possible since $k < N$. We normalize π in such a way that

$$\|\pi\|_L^\infty(A) = 1.$$

For $\eta > 0$ small enough, we see that $\tilde{q} = p + \eta\pi \in M(L)$ satisfies

$$\tilde{q}(y) = p(y),$$

and

$$\|\tilde{q}\|_{L^\infty(A)} < 1.$$

Therefore the element $q = \tilde{q}/\|\tilde{q}\|_{L^\infty(A)}$ is in K and satisfies

$$\Psi(q) = |q(y)| > |p(y)| = \Psi(p),$$

which is a contradiction. ■

Proposition A.7.30 (Monotonicity of the generalised Tchebychev polynomial with respect to A)

Let A be any compact subset and I any compact interval of $(0, +\infty)$ such that

$$|A| \leq |I|, \ \text{and} \ \sup A \leq \sup I.$$

Then we have

$$\sup_{p \in M(L)} \frac{|p(0)|}{\|p\|_{L^\infty(A)}} \leq |T_{L,I}(0)|.$$

In particular, we have

$$|T_{L,A}(0)| \leq |T_{L,I}(0)|.$$

Proof :

- Let $\tilde{X} = \{\tilde{x}_1, ..., \tilde{x}_N\}$ be the equi-oscillations points in I of $T_{L,I}$. In particular we have

$$T_{L,I}(\tilde{x}_i) = (-1)^{N-i}, \ \ \forall i \in [\![1, N]\!]. \tag{A.24}$$

Introducing the elementary interpolants $\Phi_{L,\tilde{X},\bullet}$ we can write

$$T_{L,I} = \sum_{i=1}^{N} (-1)^{N-i} \Phi_{L,\tilde{X},i}.$$

- Let $\phi : s \in [0, +\infty[\mapsto |A \cap [s, +\infty)|$. This function is continuous, non-increasing, maps $[0, +\infty[$ onto $[0, |A|]$, and $\phi(s) = 0$ for $s \geq \sup A$. In particular, since $|I| \leq |A|$, there exists $0 \leq s_1 \leq ... \leq s_N < +\infty$ such that

$$\phi(s_i) = |I \cap [\tilde{x}_i, +\infty)|.$$

  We then define

$$x_i = \inf \Big(A \cap [s_i, +\infty) \Big).$$

  By compactness of A, we have that $x_i \in A$. From now on we set $X \overset{\text{def}}{=} \{x_1, ..., x_N\} \subset A$.

- Let us now compare $\tilde{X}$ and X.

  By definition of ϕ we have $\phi(x_i) = \phi(s_i)$ since $[s_i, x_i) \cap A = \emptyset$. This means that

$$|A \cap [x_i, +\infty)| = |I \cap [\tilde{x}_i, +\infty)|.$$

  Note that those quantities are positive and in particular we have $x_i < \sup A \leq \sup I$.

  Take now any $j, k \in [\![1, N]\!]$, $j \leq k$, we have

$$\begin{aligned}
|x_k - x_j| &= |[x_j, x_k)| \\
&\geq |A \cap [x_j, x_k)| \\
&= |A \cap [x_j, +\infty)| - |A \cap [x_k, +\infty)| \\
&= |I \cap [\tilde{x}_j, +\infty)| - |I \cap [\tilde{x}_j, +\infty)| \\
&= |I \cap [\tilde{x}_j, \tilde{x}_k)| \\
&= |\tilde{x}_k - \tilde{x}_j|,
\end{aligned}$$

  since I is an interval that contains $\tilde{x}_k$ and $\tilde{x}_j$.

  Similarly we have for any k

$$\begin{aligned}
|x_k - \max A| &\geq |A \cap [x_k, +\infty)| \\
&= |I \cap [\tilde{x}_k, +\infty)| \\
&= |\tilde{x}_k - \max I|,
\end{aligned}$$

  and since $\max I \geq \max A$, we deduce that $x_k \leq \tilde{x}_k$.

- Due to the previous properties, we can apply Proposition A.7.25 to X and $\tilde{X}$ and conclude that, for any $k \in [\![1, N]\!]$, we have

$$|\Phi_{L,X,k}(0)| \leq |\Phi_{L,\tilde{X},k}(0)|. \tag{A.25}$$

  Take now any $p \in M(L)$ and let us decompose it in the Lagrange basis $(\Phi_{L,X,k})_k$

$$p(x) = \sum_{k=1}^{N} p(x_k)\Phi_{L,X,k}(x), \quad \forall x \in [0, +\infty).$$

  We evaluate this formula at $x = 0$ and we apply the triangle inequality

$$|p(0)| \leq \left(\sum_{k=1}^{N} |\Phi_{L,X,k}(0)| \right) \|p\|_{L^\infty(A)},$$

  where we have used that all the $(x_k)_k$ belong to the set A, by construction.

Applying (A.25), we get

$$|p(0)| \leq \left(\sum_{k=1}^{N} |\Phi_{L,\tilde{X},k}(0)| \right) \|p\|_{L^\infty(A)},$$

but the sign of $\Phi_{L,\tilde{X},k}(0)$ is $(-1)^{k+1}$ and thus by (A.24),

$$\sum_{k=1}^{N} |\Phi_{L,\tilde{X},k}(0)| = \left| \sum_{k=1}^{N} (-1)^{k+1} \Phi_{L,\tilde{X},k}(0) \right| = \left| \sum_{k=1}^{N} T_{L,I}(\tilde{x}_k) \Phi_{L,\tilde{X},k}(0) \right| = |T_{L,I}(0)|.$$

The proof is complete.

It is clear that we can apply the above result to $p = T_{L,A}$ since, by definition, $\|T_{L,A}\|_{L^\infty(A)} = 1$. ■

Combining the previous results we finally obtain the following result that was actually the main aim of this appendix.

Theorem A.7.31 (A Remez inequality)

Let A be a compact subset of $(0, +\infty)$, I a compact interval of $(0, +\infty)$ such that

$$|A| \leq |I|, \text{ and } \sup A \leq \sup I.$$

Then for any $p \in M(L)$ we have

$$\|p\|_{L^\infty(0,\inf A)} \leq |T_{L,I}(0)| \, \|p\|_{L^\infty(A)}.$$

Proof :

We take any $p \in M(L)$ and any $y \in (0, \inf A)$ and we apply Proposition A.7.29 to get

$$|p(y)| \leq |T_{L,A}(y)| \|p\|_{L^\infty(A)}.$$

Then we use the monotonicity of $T_{L,A}$ on $[0, \inf A)$ and the fact that $y < \inf A$ to obtain

$$|p(y)| \leq |T_{L,A}(0)| \|p\|_{L^\infty(A)}.$$

The conclusion comes from the inequality $|T_{L,A}(0)| \leq |T_{L,I}(0)|$ that we established in Proposition A.7.30. ■

Bibliography

[ABM16] Damien Allonsius, Franck Boyer, and Morgan Morancey, *Spectral analysis of discrete elliptic operators and applications in control theory*, submitted (2016).

[AKBGBT11] Farid Ammar-Khodja, Assia Benabdallah, Manuel González-Burgos, and Luz Teresa, *Recent results on the controllability of linear coupled parabolic problems: A survey*, Mathematical Control and Related Fields **1** (2011), no. 3, 267–306.

[All07] Grégoire Allaire, *Numerical analysis and optimization*, Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford, 2007, An introduction to mathematical modelling and numerical simulation, Translated from the French by Alan Craig. MR 2326223

[BE95] Peter Borwein and Tamás Erdélyi, *Polynomials and polynomial inequalities*, Springer New York, 1995.

[BE97] Peter Borwein and Tamás Erdélyi, *Generalizations of Müntz's theorem via a Remez-type inequality for Müntz spaces*, Journal of the American Mathematical Society **10** (1997), no. 02, 327–350.

[BO14] Franck Boyer and Guillaume Olive, *Approximate controllability conditions for some linear 1D parabolic systems with space-dependent coefficients*, Mathematical Control and Related Fields **4** (2014), no. 3, 263–287.

[Boy13] Franck Boyer, *On the penalised HUM approach and its applications to the numerical approximation of null-controls for parabolic problems*, ESAIM: Proceedings **41** (2013), 15–58.

[Bre83] Haïm Brezis, *Analyse fonctionnelle*, Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree], Masson, Paris, 1983, Théorie et applications. [Theory and applications]. MR 697382 (85a:46001)

[Cor07] Jean-Michel Coron, *Control and nonlinearity*, Mathematical Surveys and Monographs, vol. 136, American Mathematical Society, Providence, RI, 2007. MR 2302744 (2008d:93001)

[EN00] Klaus-Jochen Engel and Rainer Nagel, *One-parameter semigroups for linear evolution equations*, Graduate Texts in Mathematics, vol. 194, Springer-Verlag, New York, 2000, With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt. MR 1721989

[Erv17] S. Ervedoza, *Carleman estimates for elliptic PDEs and applications*, http://www.lama.univ-savoie.fr/ControOpt2017/Slides/Slides-Ervedoza-Contropt2017.pdf, 2017, Exposé à l'école ContrOpt 2017 à Monastir.

[FCGBdT10] Enrique Fernández-Cara, Manuel González-Burgos, and Luz de Teresa, *Boundary controllability of parabolic coupled equations*, Journal of Functional Analysis **259** (2010), no. 7, 1720–1758.

[FI96] A. V. Fursikov and O. Yu. Imanuvilov, *Controllability of evolution equations*, Lecture Notes Series, vol. 34, Seoul National University, Research Institute of Mathematics, Global Analysis Research Center, Seoul, 1996. MR 1406566

[FR71] H. O. Fattorini and D. L. Russell, *Exact controllability theorems for linear parabolic equations in one space dimension*, Arch. Rational Mech. Anal. **43** (1971), 272–292. MR 0335014 (48 #13332)

[FR75] ______, *Uniform bounds on biorthogonal functions for real exponentials with an application to the control theory of parabolic equations*, Quart. Appl. Math. **32** (1974/75), 45–69. MR 0510972 (58 #23325)

[KBGBdT16] Farid Ammar Khodja, Assia Benabdallah, Manuel González-Burgos, and Luz de Teresa, *New phenomena for the null controllability of parabolic systems: Minimal time and geometrical dependence*, Journal of Mathematical Analysis and Applications **444** (2016), no. 2, 1071–1113.

[KdT09] Otared Kavian and Luz de Teresa, *Unique continuation principle for systems of parabolic equations*, ESAIM: COCV **16** (2009), no. 2, 247–274.

[LR95] G. Lebeau and L. Robbiano, *Contrôle exact de l'équation de la chaleur*, Communications in Partial Differential Equations **20** (1995), no. 1-2, 335–356.

[LRL11] Jérôme Le Rousseau and Gilles Lebeau, *On carleman estimates for elliptic and parabolic operators. applications to unique continuation and control of parabolic equations*, ESAIM: COCV **18** (2011), no. 3, 712–747.

[Oli14] Guillaume Olive, *Boundary approximate controllability of some linear parabolic systems*, Evolution Equations and Control Theory **3** (2014), no. 1, 167–189.

[Sei88] Thomas I. Seidman, *How violent are fast controls?*, Mathematics of Control, Signals, and Systems **1** (1988), no. 1, 89–95.

[TW09] Marius Tucsnak and George Weiss, *Observation and control for operator semigroups*, Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks], Birkhäuser Verlag, Basel, 2009. MR 2502023

[Zab07] Jerzy Zabczyk, *Mathematical control theory*, Birkhäuser, 2007.