UNIVERSITÀ DEGLI STUDI DI PADOVA

Dottorato di Ricerca in
Ingegneria dei Sistemi

XI Ciclo

Sede Amministrativa: Università di Bologna
Sedi Consorziate: Università di Firenze, Padova e Siena

Delay–Differential Systems in the Behavioral Approach

Paolo Vettori

Il Coordinatore
Prof. Giovanni Marro

Il Tutore
Prof. Ettore Fornasini

Il Correlatore
Prof. Sandro Zampieri

AA. AA. 1995–1996, 1996–1997, 1997–1998

Contents

Chapter 1

Introduction

1.1 Models for dynamical systems

Finite dimensional linear systems are a fundamental tool in many applications: their mathematical properties are rather simple and nowadays very well understood, while their capability to approximate a wide class of real dynamical systems is one of their major issues.

There is no need to say that using nonlinear or infinite dimensional systems leads to a much higher mathematical level, sometimes involving only particular subclasses of systems, that often implies a diminished capacity to achieve concrete results out of the theoretic framework.

Nevertheless, the choice of a finite–dimensional linear systems for modeling and control of dynamical systems may lead to poor results or even to unjustified complications (e.g. control of rotating rigid bodies is a typical problem much easier handled with nonlinear methods).

This thesis is concerned with delay differential systems, a particular class of linear infinite dimensional systems, that we want to introduce here without the mathematical details that will be necessary in the following chapters.

Delay–differential systems are dynamical systems that can be modeled with differential equations that depend also on past values of the variables. Typical examples ([FM95]; see also [BFL97] for an 'unusual' control application of delay–differential systems) are chemical reactors where actuators, varying fluxes of different fluids, modify concentrations with some delay due to the length of pipes that transport liquids. A controller of such a plant must take the time delay into account.

Recently interest has grown on imposing communication constraints on control systems (see e.g. [TSM98]): the possibility of controlling systems remotely implies not only a careful analysis of the information content of 'control messages' but also an even greater attention on time lags involved: robustness issues are fundamental since delays are not fixed (neither bounded *a priori* in the worst case).

Even partial differential equations may be reduced to delay–difference equation, as in the case of a flexible rod ([MRPF95]): if a torque or a force is applied on one end, the system is naturally described by a wave equation but algebraic manipulations permit to obtain a delay–differential model.

Infinite–dimensional linear systems theory (see e.g. [CZ95] or [BPDM93]) can certainly deal with delay–differential system but, actually, this class has been treated also with other more algebraic techniques such as those developed by the theory of linear systems over commutative rings ([BBV86]).

Our approach to delay–differential systems, as in the latter case, has a very algebraic nature.

1.2 The behavioral approach

In the eighties J. C. Willems proposed a new approach to dynamical systems based on the concept of *behavior*, which is the set of time trajectories that could be exhibited by a dynamical system. Unlike classical control theory or system theory, this approach does not classify variables *a priori* as inputs, outputs or state. For example a variable is an input if it is 'free', that is to say, if the set of admissible time trajectories of this variable is the whole function space.

It is important to decide in which function space the trajectories live in: it could be the set of all continuous functions, the set of smooth (infinitely derivable) functions, one of the L^p spaces and so on. Depending on this choice, properties of a system may change. For example a system like $\dot{y} = u$ can have only u as input when the class of permitted functions is the set of continuous functions: in fact y must be differentiable, a constraint that does not permit to choose y freely.

If the function space choosen were the set of smooth functions, y could be chosen arbitrarily, since the derivative of a smooth function is still a smooth function: therefore it could be an input.

This dependence on the trajectories space is not present in the algebraic framework proposed by M. Fliess: this is, for example, the reason why optimal control problem have never been considered in this context (while very interesting progresses on this subject were announced by J. C. Willems at MTNS98) and nonlinear systems are best arranged by M. Fliess' theory.

Actually researchers succeeded in finding a good function space for trajectories only for linear systems: discrete–time [Wil86], continuous–time [Wil89], multidimensional [Obe90], delay–differential [GL97a], time–varying [OF98]. No suitable function space has been discovered for generic nonlinear systems (see [MW95] and [ZWZ97]) and, without this object, behavioral concepts do not make sense.

On the converse M. Fliess' algebraic definitions has been extended from finite–dimensional, linear, time–invariant continuous–time systems [Fli90b] to a wide range of dynamical systems: discrete–time, time–varying, delay–differential, infinite–dimensional, nonlinear and others (see e.g. [Fli90a], [Fli92], [Fli93], [FG93], [FLMR97], [HF98]). Moreover this approach gave rise to a very important notion within nonlinear systems theory: *flatness* ([FLMR95], [FLMOR97]); this concept is also useful in infinite–dimensional linear systems, topic on which M. Fliess' research has been focusing lately [FM98].

Anyway, as we shall show, once a function space is chosen, the module–theoretic objects and concepts in M. Fliess' theory can be interpreted, inside the behavioral setup, as duals to behavioral objects.

1.3 A closer look at behaviors

Given a behavior (the set of trajectories of a dynamical system) our aim is at trying to obtain its representation, i.e. we search for a formal, mathematical way (*behavioral equations*) to decide which trajectories the behavior contains.

The representation with *latent variables* is the most general: it expresses the link, with some functional relation, between the manifest trajectories and other variables that are not directly observable in the system dynamics or not essential to modeling purposes. For examples in economics, sales are the *manifest* variables, constituting the behavior, while consumer demand could be considered an unknown latent variable. In electrical circuits, internal voltages and currents can be ignored if we are interested only in the external port behavior, but are necessary to build equations from basic physical principles.

Two particular cases of a latent variable representation are useful in different contexts: when we want to easily check whether a trajectory belongs to a behavior we should have a kernel representation; on the other hand if we want to generate the whole behavior, an image representation is the best choice.

Kernel representations, corresponding, roughly speaking, to AR systems, express the behavior as the kernel of an operator; therefore, as in fault detection problems, they permit to verify immediately if a given trajectory is an admissible one for a dynamical system.

Image representations, analogous to MA systems, characterize the behavior as the range of an operator: it is the most appropriate one for simulation purposes.

If we are given a behavior with some representation, a fundamental problem in the behavioral approach is the following: is it possible to find another representation (even of a different type) and, conversely, when does another representation define the same behavior?

If we consider linear, finite–dimensional, time–invariant, continuous–time behaviors and operators that are linear differential operators of any order, then the aforementioned questions receive quite simple answers: first of all, every behavior which has an image representation admits a kernel representation.

A behavior that has a representation with latent variables admits an equivalent kernel representation disregarding smoothness issues, in the sense that this equivalence always holds true only when the function space contains only smooth functions (see remark 3.8 for a clarifying example or [Pol97, sect. 2.5] for further details).

Finally, a behavior defined by a kernel representation admits an image representation if and only if it is controllable. This property, in the behavioral approach, is defined only in terms of trajectories: loosely speaking, a behavior is controllable when, given two admissible trajectories, there is another admissible one consisting of the *past* of the first one, of a suitable connection of finite duration and of the *future* of the second one.

1.4 Summary of the thesis

The main purpose of this thesis is to give an outline of results concerning delay–differential systems that may constitute a foundation for a behavioral theory of this class of linear systems. The first papers on this subject, namely [RW97] and [GL97a], appeared at the beginning of 1997; they provided some remarkable notions and structures on which a satisfactory theory has grown up, unfortunately, only for a particular type of delay–differential system: the so-called systems with commensurate delays.

This kind of systems, also called *one delay systems*, is characterized by the fact that every time delay that appears in the equations is an integral multiple of one single value.

A fundamental structure, often used in the behavioral framework, is the Smith form for matrices: the existence of the Smith form implies that every matrix is similar to a diagonal matrix; in this way systems of differential equation are reduced to the scalar case. Operators with commensurate delays admit the Smith form while operators with noncommensurate delays do not.

Another important tool, which permits to develop a satisfactory theory when the Smith form does not exist, is duality: here duality is intended between trajectories and operators acting on them. There are many situations that are simplified by considering dual objects. A corner-stone in the behavioral approach is [Obe90], where duality is the main ingredient to analyse a great variety of linear systems with constant coefficients. Unfortunately there is not a well-established duality theory for delay–differential systems.

So, since usual techniques fail with non commensurate delay–differential systems, we have to find other, often quite complicated, ways to deal with the problems the behavioral approach poses.

Chapter 2 introduces the class of trajectories we will be concerned with in this thesis, i.e. the set of real smooth functions; the concepts are presented from a functional analytical point of view, hence also distributions and continuous linear operators on smooth functions are defined, in order to have the basic elements necessary to investigate delay–differential systems. The operatorial notation we will use is mainly explained in this chapter.

The following chapter 3 makes clear what we mean by *dynamical systems* introducing two rather different approaches to systems theory: the behavioral approach of J. C. Willems and the module theoretic approach of M. Fliess. The fundamental concepts and problems related to these two approaches are explained, showing examples and important theorems for linear time–invariant differential systems. However also delay–differential systems are introduced and generalized to the wider class of *convolutional behaviors*, which seems to have no preceding in the systems theory literature.

Differential operators are renowned, but delay–differential equations introduce new operators whose properties were discovered quite recently. Chapter 4, besides defining and proving some linear algebra results about Smith forms, Bézout equations and generalized inverses of matrices, treats in an unified manner various kinds of operators showing that they are isomorphic to particular subrings of holomorphic functions; in this way many properties and relations between these classes of operators become more intuitive.

Other mathematical details are proved in chapter 5, where more is said about the strict relation that links behaviors and the corresponding modules in the M. Fliess' approach: they are indeed the algebraic and, under suitable hypotheses, topological dual one of each other, as is shown in theorem 5.17 which seems to be a new result.

Delay–differential behaviors with commensurate delay are the subject of chapter 6 that shows two important theorems, 6.2 and 6.5, on the representation problem of behaviors and the fundamental theorem 6.6 on controllability. As we pointed out, these theorems owe their simplicity to the existence of the Smith form for delay–differential operators with one delay.

Delay–differential behaviors with non commensurate delays, treated in chapter 7, as well as convolutional behaviors, present many subtle mathematical difficulties (even of number theoretic nature) that make results less elegant and proofs much harder.

This last chapter contains, with the exception of two results of L. Habets (theorem 7.1 on equivalence of behaviors) and of H. Mounier (theorem 7.4 on controllability in the module theoretic approach), new results on delay–differential and convolutional systems.

Both results cited above, only valid for delay–differential systems defined by full row rank matrices, are generalized to convolutional systems defined by matrices of generic rank: theorems 7.2 and 7.3 extend the result on behavior equivalence while theorems 7.5 and 7.7 are relative to the problem of controllability inside the M. Fliess' approach; also theorem 7.8 shows an important necessary and sufficient condition for a particular type of controllability that is defined in this framework.

As regards controllability of behaviors, section 7.2.2 proves the existence of a necessary condition (existence of an image representation) and of a sufficient condition (a rank condition over $\mathbb{C}$ of the defining matrix called *spectral controllability*) for behavioral controllability of convolutional systems; these conditions are both necessary and sufficient with commensurate delays, but example 7.14 shows that this is not the case even with two incommensurable delays.

Other results (as theorems 7.12 and 7.21) give conditions that make the above facts equivalent. Theorem 7.21, in particular, is an algebraic criterion that may be implemented by Gröbner bases.

There are two results on delay–differential behaviors that admit an image representation (therefore are controllable): theorem 7.26 states that it is always possible to find an full column rank delay–differential image representation while theorem 7.28 proves that if the system has a single input then the defining matrix admits a generalized inverse (if the matrix is full row rank, then it is invertible on the left); this condition is also equivalent to controllability for delay–differential systems with one delay, as is showed in theorem 6.6.

Last section investigates another condition that is equivalent to controllability for linear differential systems (see e.g. [Wil91, p. 266] for discrete-time systems): let us consider only the trajectories of a given behavior that are zero both in the 'past' and in the 'future'; the closure of this set (with respect to the topology of the function space trajectories belong to) is equal to the behavior if and only if it is controllable. The main result (theorem 7.34) states that for every convolutional behavior that admits a full row rank kernel representation, this condition is equivalent to spectral controllability.

To help the reader to face this puzzling situation, pages 109, 110 and 111 show a graphical representation of the relations between different controllability conditions in the behavioral and in the module theoretic approaches for delay–differential and convolutional systems.

Chapter 2

Smooth functions and convolutions

This chapter reviews some known results in functional analysis about smooth functions and their topological dual: distributions with compact support.

These concepts, in particular convolutional equations, constitute a very generic framework for the study of linear systems: the notation introduced in this chapter will be used throughout the whole thesis.

2.1 Topological preliminaries

Some basic topological definition are given in appendix B.

Definition 2.1. *A topological vector space is called **F-space** if it is metrizable and complete; a locally convex F-space is called **Fréchet space**.*

This kind of topological vector spaces is very important: continuous functions that are derivable infinitely many times or only up to order k, holomorphic functions, power series and smooth functions rapidly decreasing at infinity are Fréchet but not Banach spaces, i.e. are not normable [Tre67, ch. 10].

In particular

Definition 2.2. *The set of **smooth functions** $C^\infty(\mathbb{R}, \mathbb{R})$ equipped with the standard topology (uniform convergence of derivatives of any order on every compact) will be denoted by the symbol $\mathcal{E}$; the **holomorphic functions** on $\mathbb{C}$, that being smooth carry the same topology as $\mathcal{E}$, will be denoted by $\mathcal{O}$.*

Both spaces have the **Heine–Borel property**: every closed and bounded subset is compact. A consequence of this important fact is that these spaces are Montel spaces and therefore **reflexive** (see [Tre67, ch. 34 and 36]).

There are some very important results about topological vector spaces that will be necessary in the following sections. Their proofs can be found in most books on topology or functional analysis, e.g. [Rud73].

Definition 2.3. *A* ***functional*** *on a vector space $\mathcal{V}$ is a map $\alpha : \mathcal{V} \to \mathbb{F}$ where $\mathbb{F}$ is the field of scalars of $\mathcal{V}$.*

If $\mathcal{V}$ is equipped with a topology τ, then we will be mainly concerned with continuous functionals.

Proposition 2.4. *If Λ is a non zero linear functional on $\mathcal{V}$, then $\ker\Lambda = \Lambda^{-1}(0)$ is a proper subspace of $\mathcal{V}$.*

The following conditions are equivalent:

$$\Lambda \text{ continuous } \Leftrightarrow \ker\Lambda \text{ closed } \Leftrightarrow \ker\Lambda \text{ not dense in } \mathcal{V} \Leftrightarrow \Lambda \text{ bounded}$$

Proposition 2.5. *If $\mathcal{U}$ is a closed subspace of $\mathcal{V}$ with topology τ, then the quotient vector space $\mathcal{V}/\mathcal{U} = \{x + \mathcal{U} : x \in \mathcal{V}\}$ may be equipped with the* ***quotient topology*** *$\tau_{\mathcal{U}} = \{T + \mathcal{U} : T \in \tau\}$.*

Moreover, if $\mathcal{V}$ is a Fréchet space or an F-space, so is $\mathcal{V}/\mathcal{U}$.

Next two theorems, the 'open mapping theorem' and the 'Hahn–Banach theorem' may be stated in a much more general form and, especially for the second one, in many different ways. Here we state them in a form that is suited to our purposes (see [Rud73, cor. 2.12, thm. 3.6, 3.5]).

Theorem 2.6 (Open mapping theorem). *If Λ is a continuous linear mapping of an F-space onto an F-space and is one-to-one, then Λ^{-1} is continuous.*

Theorem 2.7 (Hahn–Banach theorem, analytical form). *If λ is a continuous linear functional on a subspace $\mathcal{U}$ of a locally convex topological vector space $\mathcal{V}$, then there exists a continuous linear functional Λ on $\mathcal{V}$ such that $\lambda = \Lambda$ on $\mathcal{U}$.*

Corollary 2.8. *If $\mathcal{U}$ is a subspace of a locally convex topological vector space $\mathcal{V}$, $x \in \mathcal{V}$ but $x \notin \overline{\mathcal{U}}$, the closure of $\mathcal{U}$, then there exists a continuous linear functional $\lambda : \mathcal{V} \to \mathbb{F}$ such that $\lambda(x) = 1$ and $\lambda(y) = 0$ for every $y \in \mathcal{U}$.*

2.2 Distributions

We recall here some definitions and important facts about distributions.

Definition 2.9. *The **dual** of the topological vector space $\mathcal{V}$ is the set of all continuous linear functionals on $\mathcal{V}$ and will be denoted by $\mathcal{V}'$.*

Distributions constitute the dual ($\mathcal{D}'$) of the test functions, i.e. smooth functions with compact support ($\mathcal{D} \subset \mathcal{E}$). Following the standard notation, if α is a distribution and $\phi(t)$ a test function, the value of α at ϕ is denoted by $\langle \alpha, \phi \rangle \in \mathbb{R}$.

Distributions extend functions in the following sense: if f is a locally integrable function, i.e. (Lebesgue) integrable on every compact, then it is a distribution

$$f : \mathcal{D} \to \mathbb{R},\ \phi \mapsto \langle f, \phi \rangle \triangleq \int_{\mathbb{R}} f(t)\phi(t)dt. \tag{2.1}$$

Obviously not every distribution may be associated to a locally integrable function: if δ, the evaluation at the origin $\langle \delta, \phi \rangle = \phi(0)$, were a function $\delta(t)$, it should be zero everywhere except at the origin, but in this case the integral of $\delta(t)\phi(t)$ would be zero.

Remark 2.10. *In what follows, the conditions that insure convergence of integrals will be always supposed satisfied, if not otherwise stated (due to their prevalently heuristic flavour).*

Note that if f is differentiable then, upon integrating by parts and remembering that ϕ is zero for large $|t|$, we get

$$< f', \phi >= \int_{\mathbb{R}} f'(t)\phi(t)dt = -\int_{\mathbb{R}} f(t)\phi'(t)dt = -\langle f, \phi' \rangle \tag{2.2}$$

that makes natural the definition of **derivative of a distribution** α up to any order $k \in \mathbb{N}$:

$$\left\langle \alpha^{(k)}, \phi \right\rangle = (-1)^k \left\langle \alpha, \phi^{(k)} \right\rangle. \tag{2.3}$$

The derivatives of the δ distributions map a function into its derivatives evaluated at the origin (regardless of the sign).

2.3 Convolutions

In order to define convolutions between functions and distributions, we have to introduce a fundamental operator: the **shift operator** σ_τ mapping a generic function f (defined at least on a group) into the function

$$(\sigma_\tau f)(t) \triangleq f(t-\tau). \tag{2.4}$$

Note that, given any distribution α and a test function $\phi(t)$, $\langle \alpha, \sigma_\tau \phi \rangle$ is a function of τ. Moreover

$$\phi(\tau) = (\sigma_{-\tau}\phi)(0) = \langle \delta, \sigma_{-\tau}\phi \rangle . \tag{2.5}$$

Since for functions f and g

$$\langle \sigma_\tau f, g \rangle = \int_{\mathbb{R}} f(t-\tau)g(t)\,dt = \int_{\mathbb{R}} f(x)g(x+\tau)\,dx = \langle f, \sigma_{-\tau} g \rangle \tag{2.6}$$

we may define the **shift** of a distribution as follows

$$\langle \sigma_\tau \alpha, \phi \rangle \triangleq \langle \alpha, \sigma_{-\tau}\phi \rangle . \tag{2.7}$$

Another standard notation in functional analysis is the symbol

$$\check{f}(t) \triangleq f(-t), \tag{2.8}$$

the **symmetric** of f with respect to the origin: we have that

$$\langle f, \check{g} \rangle = \int_{\mathbb{R}} f(t)g(-t)\,dt = \int_{\mathbb{R}} f(-x)g(x)\,dx = \left\langle \check{f}, g \right\rangle . \tag{2.9}$$

This property also extends to distributions by assuming:

$$\langle \check{\alpha}, \phi \rangle \triangleq \left\langle \alpha, \check{\phi} \right\rangle . \tag{2.10}$$

Symmetry and shift operators do not commute, but:

$$(\sigma_\tau f)^\vee(t) = (\sigma_\tau f)(-t) = f(-t-\tau) = \check{f}(t+\tau) = (\sigma_{-\tau}\check{f})(t). \tag{2.11}$$

Combining last property with (2.7) and (2.10) we have for a generic distribution

$$\left\langle(\sigma_\tau\alpha)^\vee,\phi\right\rangle=\left\langle\sigma_\tau\alpha,\check{\phi}\right\rangle=\left\langle\alpha,\sigma_{-\tau}\check{\phi}\right\rangle=\left\langle\alpha,(\sigma_\tau\phi)^\vee\right\rangle=\langle\check{\alpha},\sigma_\tau\phi\rangle=\langle\sigma_{-\tau}\check{\alpha},\phi\rangle\,. \tag{2.12}$$

The **convolution** of two functions is the function

$$(f\star g)(\tau)\triangleq\int_{\mathbb{R}}f(t)g(\tau-t)\,dt=\int_{\mathbb{R}}f(\tau-t)g(t)\,dt \tag{2.13}$$

when the (Lebesgue) integral exists. Using (2.9) we obtain

$$(f\star g)(\tau)=\int_{\mathbb{R}}f(t)\check{g}(t-\tau)\,dt=\int_{\mathbb{R}}f(t)\sigma_\tau\check{g}(t)\,dt=\langle f,\sigma_\tau\check{g}\rangle \tag{2.14}$$

This formula suggest the definition

$$(\alpha\star\phi)(\tau)\triangleq\left\langle\alpha,\sigma_\tau\check{\phi}\right\rangle=\langle\check{\alpha},\sigma_{-\tau}\phi\rangle\in\mathcal{E} \tag{2.15}$$

when $\alpha\in\mathcal{D}'$ is a distribution and $\phi(t)\in\mathcal{D}$ a test function [Tre67, p. 287]. In order to have a well defined convolution between distributions we need an additional hypothesis.

Actually, if f and g are functions acting as distributions on $\phi\in\mathcal{D}$, then it is easy to see that $\langle f\star g,\phi\rangle$ is well defined when the function $f(t)g(\tau)\phi(t+\tau)$ has compact support in $\mathbb{R}^2$; this happens if f, or g, has compact support,ll but even if the support of both functions is bounded either on the left or on the right.

Definition 2.11. *A distribution α is said to* ***vanish*** *in an open set U if $\langle\alpha,\phi\rangle=0$ for every test function ϕ having its support in U. The* ***support of the distribution*** *α is the complement of the largest open set in which α vanishes.*

Proposition 2.12. *The expression $\langle\alpha,\langle\beta,\sigma_{-\tau}\phi\rangle\rangle$ makes sense if either one of the distributions α or β has compact support or both have a bounded support on the same side. In this case that expression defines a functional which is called the* ***convolution*** *of α and β:*

$$\langle\alpha\star\beta,\phi\rangle\triangleq\langle\alpha,\langle\beta,\sigma_{-\tau}\phi\rangle\rangle. \tag{2.16}$$

If the support of α and β is bounded on one side, then the same fact holds for $\alpha\star\beta$.

If we only consider the set of distributions with compact support (which, as is shown in [Tre67, thm. 27.7], is a ring with respect to '+' and '$\star$' whose identity with respect to convolution is δ, fact trivial to verify from (2.5) and from the definition given above), then (2.16) always holds and defines a distribution $\gamma = \alpha \star \beta$ with compact support.

This hypothesis is indeed not restrictive: as [Tre67, thm. 24.2] states:

Theorem 2.13. *The ring of distributions with compact support is the topological dual of the topological vector space of infinitely differentiable functions. Therefore it is denoted by $\mathcal{E}'$.*

This theorem shows that, since we are going to employ infinitely differentiable functions, we need only distributions with compact support;

In order to treat linear systems we need not only functionals, but also a ring of operators mapping $\mathcal{E}$ into itself. For our purposes (linear time invariant systems), we should require that these operators are linear, continuous and commute with the shift operator. We show that this ring is isomorphic to $\mathcal{E}'$.

Lemma 2.14. *Convolution 'commutes' with shift, i.e. for every $\alpha \in \mathcal{E}'$, $f \in \mathcal{E}$ and $\tau \in \mathbb{R}$*

$$\sigma_\tau(\alpha \star f) = \alpha \star (\sigma_\tau f). \tag{2.17}$$

Proof. The proof follows at once from definitions:

$$(\sigma_\tau(\alpha \star f))(t) = (\alpha \star f)(t - \tau) = \left\langle \check{\alpha}, \sigma_{(\tau - t)} f \right\rangle = \langle \check{\alpha}, \sigma_{-t}\sigma_\tau f \rangle = (\alpha \star (\sigma_\tau f))(t).$$

□

Then we prove that

Lemma 2.15. *Every linear continuous operator $\Lambda : \mathcal{E} \to \mathcal{E}$ which commutes with the shift operator $\Lambda\sigma_\tau = \sigma_\tau\Lambda$ is a convolution: there exists a unique $\alpha \in \mathcal{E}'$ such that $\Lambda f = \alpha \star f$ for every $f \in \mathcal{E}$.*

Proof. Let λ be the functional that maps $\mathcal{E} \ni f \mapsto \Lambda f(0)$: it is a composition $\lambda = \delta \circ \Lambda$ of a continuous linear operator and a continuous linear functional therefore $\lambda \in \mathcal{E}'$ and $\langle \lambda, f \rangle = \Lambda f(0)$.

Since Λ commutes with the shift,

$$(\Lambda f)(\tau) = (\sigma_{-\tau}(\Lambda f))(0) = (\Lambda(\sigma_{-\tau} f))(0) = \langle \lambda, \sigma_{-\tau} f \rangle = \check{\lambda} \star f(\tau)$$

from definition (2.15); thus $\alpha = \check{\lambda}$.

The distribution λ is uniquely determined: if $\langle \lambda_1, f \rangle = \langle \lambda_2, f \rangle = \Lambda f(0)$ then $\langle \lambda_1 - \lambda_2, f \rangle = 0$ for every $f \in \mathcal{E}$ therefore $\lambda_1 = \lambda_2$. □

So we have the following:

Theorem 2.16. *Let L be the set of all continuous linear operators on $\mathcal{E}$ which commute with the shift operator. Then the map*

$$\tilde{}\colon \mathcal{E}' \to L, \ \alpha \mapsto \tilde{\alpha} \text{ such that } \tilde{\alpha}(f) = \alpha \star f$$

is a ring isomorphism, i.e. the composition of $\tilde{\alpha}$ and $\tilde{\beta}$ is $\tilde{\alpha}\tilde{\beta} = \widetilde{\alpha \star \beta}$.

Proof. The map $\tilde{}$ is well defined because for every $\alpha \in \mathcal{E}'$, $\tilde{\alpha}$ is linear, commutes with shift by lemma 2.14 and is continuous [Tre67, thm. 27.3]; lemma 2.15 states that the map is onto and one-to-one.

The map is a ring isomorphism if we have $\alpha \star (\beta \star f) = (\alpha \star \beta) \star f$. If we suppose, for notation purposes, that in the following formula $\beta \star f$ is a function of τ, then, from definition (2.15):

$$\begin{aligned}(\alpha \star (\beta \star f))(t) &= \langle \check{\alpha}, \sigma_{-t}(\beta \star f) \rangle = \langle \check{\alpha}, \beta \star (\sigma_{-t} f) \rangle = \left\langle \check{\alpha}, \left\langle \check{\beta}, \sigma_{-\tau}\sigma_{-t} f \right\rangle \right\rangle \\ &= \left\langle \check{\alpha} \star \check{\beta}, \sigma_{-t} f \right\rangle = \langle (\alpha \star \beta)^{\vee}, \sigma_{-t} f \rangle = ((\alpha \star \beta) \star f)(t).\end{aligned}$$

□

Through this isomorphism we may consider the set $\mathcal{E}'$ not only as the topological dual of $\mathcal{E}$, but also as a ring of operators acting on it.

2.4 The Paley-Wiener theorem

Consider the set of **rapidly decreasing smooth functions**

$$\mathcal{S} = \left\{ f(t) \in \mathcal{E} : \ \forall i, k \in \mathbb{N} \ \lim_{t \to \infty} |t^k| f^{(i)}(t) = 0 \right\}, \tag{2.18}$$

which is a Fréchet space (and $\mathcal{D} \subset \mathcal{S} \subset \mathcal{E}$).

This space is important since we can define the **Fourier operator** on its elements:

$$\mathcal{F} : \mathcal{S} \to \mathcal{S},\ f(t) \mapsto (\mathcal{F}f)(\omega) \triangleq \int_{\mathbb{R}} f(t)e^{-2\pi i\omega t}\,dt \tag{2.19}$$

which [Tre67, ch. 25] is a topological isomorphism.

It is easy to extend the Fourier operator on the dual $\mathcal{S}'$, the set of **tempered distributions**:

$$\mathcal{F} : \mathcal{S}' \to \mathcal{S}',\ \alpha \mapsto \mathcal{F}\alpha \text{ such that } \forall f \in \mathcal{S}\ \langle \mathcal{F}\alpha, f\rangle = \langle \alpha, \mathcal{F}f\rangle \tag{2.20}$$

Regarding the function $f(t) \in \mathcal{S}$ as a distribution, we can rewrite its Fourier transform (2.19) as $(\mathcal{F}f)(\omega) = \langle f(t), e^{-2\pi i\omega t}\rangle$; in an analogous way we could define the Fourier transform of a distribution as a function of ω, but we need an additional hypothesis: it must have compact support [Tre67, pr. 29.1].

Proposition 2.17. *The Fourier transform of a distribution with compact support $\alpha \in \mathcal{E}'$ is the function*

$$(\mathcal{F}\alpha)(\omega) = \left\langle \alpha, e^{2\pi i\omega t}\right\rangle \tag{2.21}$$

where α acts on $e^{-2\pi i\omega t} = f_\omega(t)$ that is function of t with parameter ω.

We can extend $\mathcal{F}\alpha(\omega)$ to a complex holomorphic function just substituting $\omega \in \mathbb{R}$ with $s \in \mathbb{C}$ in equation 2.21, so that $\mathcal{F}\alpha(s) \in \mathcal{O}$.

Most of the mathematical literature is based on the Fourier transform while engineers prefer Laplace transform, probably due to its simpler form and its capability to operate on more general functions than the other transform does.

On the other side the Fourier transform can operate on tempered distributions while Laplace transform cannot. In our context (distributions with compact support) both transforms exist:

Definition 2.18. *The **Laplace transform** is the operator so defined*

$$\mathcal{L} : \mathcal{E}' \to \mathcal{O}\ \alpha \mapsto \hat{\alpha}(s) = (\mathcal{L}\alpha)(s) \triangleq \left\langle \alpha, e^{-st}\right\rangle \tag{2.22}$$

where α acts on $e^{-st} = f_s(t)$ that is a function of t with parameter s.

The relation between the two transforms is trivial: $\hat{\alpha}(s) = (\mathcal{F}\alpha)\left(\frac{s}{2\pi i}\right)$.

Example 2.19. The holomorphic function $h(s) = s^k$ corresponds to the k-th order derivative operator: indeed a simple computation shows that

$$(\mathcal{L}\delta^{(k)})(s) = \left\langle \delta^{(k)}, e^{-st} \right\rangle = (-1)^k \left(\tfrac{d^k}{dt^k} e^{-st} \right) \Big|_{t=0} = s^k.$$

Then, by equation (2.3) and definition (2.8) we have

$$\left\langle \delta^{(k)}, \check{f} \right\rangle = \left\langle \delta, (-1)^k \tfrac{d^k}{dt^k} \check{f} \right\rangle = \left\langle \delta, (\tfrac{d^k}{dt^k} f)^\vee \right\rangle$$

therefore

$$(\delta^{(k)} \star f)(t) = \left\langle \delta^{(k)}, \sigma_t \check{f} \right\rangle = \left\langle \delta, \sigma_t (f^{(k)})^\vee \right\rangle = f^{(k)}(t).$$

♣

Example 2.20. The shift operator satisfies the hypotheses of theorem 2.16: therefore there are a distribution in $\mathcal{E}'$ and an holomorphic function that correspond to it.

From (2.5) and (2.7) we get immediately that if δ_τ is the functional that evaluates a function at τ then

$$\langle \delta_\tau, f \rangle = f(\tau) = \langle \delta, \sigma_{-\tau} f \rangle = \langle \sigma_\tau \delta, f \rangle \tag{2.23}$$

therefore $\delta_\tau = \sigma_\tau \delta$; now by (2.16) and (2.9)

$$(\tilde{\delta}_\tau f)(t) = \left\langle \delta_\tau, \sigma_t \check{f} \right\rangle = \left\langle \delta, \sigma_t \sigma_{-\tau} \check{f} \right\rangle = \langle \delta, \sigma_t (\sigma_\tau f)^\vee \rangle = (\sigma_\tau f)(t). \tag{2.24}$$

Using the Laplace transform

$$\hat{\delta}_\tau(s) = \left\langle \delta_\tau, e^{-st} \right\rangle = e^{-s\tau} \tag{2.25}$$

hence the holomorphic function corresponding to $\sigma_\tau = \tilde{\delta}_\tau$ is $e^{-s\tau}$. ♣

We can say a little bit more about the holomorphic functions that are Laplace transforms of distributions in $\mathcal{E}'$: let

$$p(s) \triangleq \log\left(1 + |s|^2\right) + |\operatorname{Re} s| \tag{2.26}$$

and define the algebra of holomorphic functions[1]

$$\mathcal{A} \triangleq \left\{ f \in \mathcal{O} : \exists A, B > 0,\ |f(s)| \leq A e^{Bp(s)}\ \forall s \in \mathbb{C} \right\} \tag{2.27}$$

which will be called ring of **Paley–Wiener functions**. Then

Theorem 2.21 (Paley–Wiener). *The space $\mathcal{A}$ defined in (2.27) is topologically isomorphic to $\hat{\mathcal{E}}'$, the set of Laplace transforms of distributions with compact support.*

This is a fundamental theorem because it permits to characterize those holomorphic functions that have an 'operatorial' meaning. Distributions that have a punctual support have rather simple Laplace transforms (see examples 2.19 and 2.20); the following example shows, employing definition (2.27), a non trivial holomorphic function $h(s) \in \mathcal{A}$ that is the Laplace transform of some distribution in $\mathcal{E}'$.

Example 2.22. Let $h(s)$ be as follows:

$$h(s) = \frac{e^s - e^{-s}}{s}.$$

The function is holomorphic; to verify that $h(s) \in \mathcal{A}$ let us consider what happens if s is in a neighborhood of $0 \in \mathbb{C}$ and then in its complement.

Being $h(s)$ continuous (even at $s = 0$), it is bounded on every compact set:

$$|s| \leq 1 \ \Rightarrow\ \exists A' \text{ such that } |h(s)| \leq A' \leq A' e^{Bp(s)}.$$

Then, we know that e^s and e^{-s} are Paley–Wiener functions, being Laplace transforms (see equation (2.25)) of the distributions δ_{-1} and δ_1. Therefore

$$|s| \geq 1 \ \Rightarrow\ |h(s)| = \frac{|e^s - e^{-s}|}{|s|} \leq |e^s - e^{-s}| \leq |e^s| + |e^{-s}| \leq 2\max\left\{ |e^s|, |e^{-s}| \right\}$$

so there are $A'', B > 0$ such that $|h(s)| \leq A'' e^{Bp(s)}$ for $|s| \geq 1$.

In conclusion, taking $A = \max\{A', A''\}$, $h(s)$ satisfies definition (2.27). ♣

The 'operatorial' meaning of a similar function is explained in example 6.1.

[1] See for more general definitions [Str83, ch. 3], where the function $p(s)$ differs from that given here since we are going to use Laplace instead of Fourier transforms; we omit to indicate the dependence of $\mathcal{A}$ on p, thus we will not use a notation like $\mathcal{A}_p$, because we will deal only with this particular class of holomorphic functions.

2.5 Systems of convolutional equations

The concepts developed so far are still valid if we no longer confine ourselves to scalar operators. Moreover

Remark 2.23. *The expression written in this chapter do not change at all if we follow this convention: the elements with functional or operatorial meaning, e.g. elements in $\mathcal{A}^k$, are row vectors while vectors of functions, e.g. elements in $\mathcal{E}^q$, are column vectors.*

Let us say something more precise about the non scalar case: first of all the evaluation of a distribution $\alpha \in \mathcal{E}'^p$ at a function $v(t) \in \mathcal{E}^p$ is

$$\langle \alpha, v \rangle \triangleq \sum_{j=1}^{p} \langle \alpha_j, v_j \rangle \tag{2.28}$$

and analogously a system of convolutional equation will be treated as follows: if we consider the distribution $\beta \in \mathcal{E}^{p\times q}$, $\tilde{\beta}$ maps $w(t) \in \mathcal{E}^q$ to $v = \tilde{\beta}w \in \mathcal{E}^p$ in the usual way:

$$v_i = \sum_{j=1}^{q} \tilde{\beta}_{ij} w_j,\ \forall i = 1, \dots , p. \tag{2.29}$$

Note that from (2.16) and theorem 2.16 we get

$$\langle \alpha \star \check{\beta}, w \rangle = \left\langle \alpha, \langle \check{\beta}, \sigma_{-\tau} w \rangle \right\rangle = \langle \alpha, \beta \star w \rangle = \left\langle \alpha, \tilde{\beta} w \right\rangle \tag{2.30}$$

i.e. the adjoint of the operator $\tilde{\beta}$, denoted by $\tilde{\beta}'$, maps α to $\tilde{\beta}'\alpha = \alpha \star \check{\beta}$. The above-mentioned convention ensures the correctness of these symbolic manipulations once the dimensions of matrices and vectors are well chosen.

Thus, as operators,

$$\tilde{\alpha} v = \tilde{\alpha}(\tilde{\beta} w) = (\tilde{\alpha}\tilde{\beta}) w = \widetilde{\alpha \star \beta} w \tag{2.31}$$

where the matrix $\tilde{\beta}$ is “multiplied” by the column vector w on the right and by the row vector $\tilde{\alpha}$ on the left.

The Paley–Wiener theorem and the isomorphic relation between $\alpha \in \mathcal{E}'$, $\hat{\alpha}(s)$ and the operator $\tilde{\alpha}$ permits us to identify $\mathcal{E}'$, $\hat{\mathcal{E}}'$, $\mathcal{A}$ and the ring of linear continuous operators on $\mathcal{E}$ commuting with the shift; last equation may be also written, with

$a(s) = \hat{\alpha}(s) \in \mathcal{A}^p$ and $B(s) = \hat{\beta}(s) \in \mathcal{A}^{p\times q}$, as

$$a(s)v = a(s)(B(s)w) = (a(s)B(s))w.$$

We note that $a(s)B(s)$ is effectively a multiplication in $\mathcal{O}$ of particular holomorphic functions while $B(s)w$ has the operatorial meaning of equation (2.29). Due to these different "multiplications" of elements in $\mathcal{A}$ (or in other operator rings), when we are concerned with kernels and images of matrices, we have to be particularly careful.

Remark 2.24. *If we are considering the kernel of a matrix $B(s) \in \mathcal{A}^{p\times q}$ then we can view it as a matrix of operators acting on the right on $\mathcal{E}$:*

$$\ker_{\mathcal{E}} B(s) \ \textit{or} \ \ker_{\mathcal{E}} B(s)\circ \ \textit{is the set} \ \{w(t) \in \mathcal{E}^q :\ B(s)w = 0\}\,;$$

or think of $B(s)$ as a matrix taking elements in $\mathcal{A}$, which may be multiplied on the left by (row) vectors in $\mathcal{A}^p$:

$$\ker_{\mathcal{A}} \circ B(s) \ \textit{is the set} \ \{a(s) \in \mathcal{A}^p :\ a(s)B(s) = 0\}\,.$$

Analogously, the image of $B(s)$ may be twofold: as an operator

$$\operatorname{im}_{\mathcal{E}} B(s) \ \textit{or} \ \operatorname{im}_{\mathcal{E}} B(s)\circ \ \textit{is the set} \ \{v = B(s)w :\ w(t) \in \mathcal{E}^q\} \subseteq \mathcal{E}^p;$$

while as a matrix on the ring $\mathcal{A}$ we have, using in this case also another intuitive notation,

$$\operatorname{im}_{\mathcal{A}} \circ B(s) \ \textit{or} \ \mathcal{A}^p B(s) \ \textit{is the set} \ \{b(s) = a(s)B(s) :\ a(s) \in \mathcal{A}^p\} \subseteq \mathcal{A}^q.$$

Chapter 3

Algebraic approaches

Dynamical systems may be defined and treated in many different ways; in this chapter we present two important approaches that recently gave rise to interesting results on delay–differential systems. A more specific approach to system over rings, that still needs further development, will be shortly presented in chapter 7.

3.1 The behavioral approach

This section is devoted to a basic, but necessary, explanation of the behavioral approach as it was developed by J. C. Willems. Definitions and theorems will be introduced for the well-known case of linear continuous time differential systems.

3.1.1 Dynamical systems and behaviors

The first step is the definition of the objects the behavioral approach deals with.

Definition 3.1. *A* ***dynamical system*** *is a triple* $\Sigma = (T, W, \mathcal{B})$*:* $\mathcal{B}$*, the* ***behavior*** *of the dynamical system, is a subset of* $W^T = \{w : T \to W\}$*, the set of functions from* T*, the* ***time set****, into* W*, the* ***signal alphabet****.*

The above definition assumes no particular structure on the sets T and W. We will require two more properties.

Definition 3.2. *A dynamical system* $\Sigma = (T, W, \mathcal{B})$ *is* ***linear*** *if* W *is a vector space (the same structure is induced on* W^T*) and* $\mathcal{B}$ *is a subspace of* W^T*.*

A dynamical system $\Sigma = (T, W, \mathcal{B})$ is ***time-invariant*** *if $(T, +)$ is a semigroup and*

$$\forall \tau \in T,\ w(\cdot) \in \mathcal{B} \Rightarrow w(\cdot + \tau) \in \mathcal{B}. \tag{3.1}$$

Remark 3.3. *When the time set is not only a semigroup (like $\mathbb{N}$ or $\mathbb{R}_+$), but a group (assuring, roughly speaking, that the '$-$' operation is well defined, like in $\mathbb{Z}$ and $\mathbb{R}$), we can use the shift operator σ_τ defined in (2.4) that maps a function $w \mapsto (\sigma_\tau w)(t) = w(t - \tau)$. In this case property (3.1) is equivalent to*

$$\forall \tau \in T,\ w \in \mathcal{B} \Rightarrow \sigma_\tau w \in \mathcal{B}. \tag{3.2}$$

We will only deal with time-invariant linear continuous-time systems, i.e. systems for which $T = \mathbb{R}$ and $W = \mathbb{R}^q$ or $W = \mathbb{C}^q$: a more general set-up [Wil89, Wil91] is not so useful in our context. Other classes of dynamical systems will be occasionally touched upon: for a deeper analysis of their properties we refer the interested reader to the bibliography.

When the dynamical system is $\Sigma = (\mathbb{R}, \mathbb{R}^q, \mathcal{B})$, its behavior $\mathcal{B}$ is a subset of the set of all **time trajectories** $(\mathbb{R}^q)^\mathbb{R} = \{w : \mathbb{R} \to \mathbb{R}^q\}$. In order to obtain a clear physical meaning, we will always implicitly assume that the behavior $\mathcal{B}$ is a subset of $C^\infty(\mathbb{R}, \mathbb{R}^q)$, the set of all smooth (i.e. infinitely differentiable) functions that will be denoted by $\mathcal{E}^q$ following definition 2.2; we will call these systems **linear smooth systems**.

3.1.2 Behavior specification

The simplest way to specify a behavior is to introduce the characteristic function of $\mathcal{B}$, i.e. $\chi : W^T \to \{0, 1\}$ so that $\chi(w) = 0$ if and only if $w \in \mathcal{B}$. Therefore $\mathcal{B} = \chi^{-1}(0)$, which is called **behavioral equation**.

That is too abstract for our purposes: so we will suppose that the behavior is specified by some set of differential equations. In other words the trajectories belonging to $\mathcal{B}$ satisfy a system of differential equations. The notation will be more compact and clear if we introduce the following class of differential operators.

Definition 3.4. *A polynomial differential operator*

$$r\left(\tfrac{d}{dt}\right) = \sum_0^n a_i \tfrac{d^i}{dt^i} = a_n \tfrac{d^n}{dt^n} + a_{n-1}\tfrac{d^{n-1}}{dt^{n-1}} + \cdots + a_1 \tfrac{d}{dt} + a_0 \in \mathbb{R}\left[\tfrac{d}{dt}\right] \tag{3.3}$$

maps $w(t) \in \mathcal{E}$ *into*

$$r\left(\tfrac{d}{dt}\right) w = \sum_0^n a_i w^{(i)}(t) = a_n w^{(n)}(t) + a_{n-1} w^{(n-1)}(t) + \cdots + a_1 w^{(1)}(t) + a_0 w(t).$$

where $w^{(k)}(t) \triangleq \frac{d^k}{dt^k} w(t)$.

A matrix operator $R\left(\frac{d}{dt}\right) \in \mathbb{R}\left[\frac{d}{dt}\right]^{p\times q}$ *maps in the usual way* $w(t) \in \mathcal{E}^q$ *into* $v(t) \in \mathcal{E}^p$, *such that* $v_i = \sum_j R_{ij}\left(\frac{d}{dt}\right) w_j$.

Remark 3.5. *We note that the operators just introduced in this chapter are consistent with the previously defined ones. If the distribution* α *is a linear combination of* $\delta^{(k)}$ *then example 2.19 shows that* $\hat{\alpha}(s)$ *is a linear combination of* s^k*:*

$$\alpha = \sum_0^n a_k \delta^{(k)} \;\Rightarrow\; \hat{\alpha}(s) = \sum_0^n a_k s^k \tag{3.4}$$

in other words $\hat{\alpha}(s) \in \mathbb{R}[s]$ *and since* s^k *(or, better,* $\widetilde{s^k}$*) is the* k*-th order derivative operator,*

$$(\tilde{\alpha} w)(t) = \sum_0^n a_k w^{(k)}(t)$$

i.e. $\tilde{\alpha}$ *and* $r\left(\frac{d}{dt}\right)$*, as defined in (3.3), are the same operator.*

Therefore, in the end, $\tilde{\alpha} = \hat{\alpha}\left(\frac{d}{dt}\right)$*. This allows us to treat the polynomial* $\hat{\alpha}(s)$ *as a differential polynomial operator identifying the* symbols s *and* $\frac{d}{dt}$.

Anyway, for all $a(s) \in \mathcal{A}$*, since there is a unique* $\alpha \in \mathcal{E}'$ *such that* $a(s) = \hat{\alpha}(s)$*, the expression* $a(s)w$ *will have the following intuitive meaning:*

$$(a(s)w)(t) = (\tilde{\alpha} w)(t) = (\alpha \star w)(t) = \langle \alpha, \sigma_t \check{w} \rangle\,. \tag{3.5}$$

In this section, however, we will deal only with polynomial operators: therefore we will prefer a more classical and simple notation.

There are many ways to write down a system of differential equations; the most important ones that will be used in this thesis, are called **kernel representation**, **image representation** and **latent variables representation**, which is the generalization of the previous ones. We are already in a position to define the first two representations:

Definition 3.6. *A linear smooth system has a* ***differential kernel representation*** *if the elements of its behavior satisfy a system of homogeneous differential equations: there is a matrix of polynomial differential operators* $R\left(\frac{d}{dt}\right) \in \mathbb{R}\left[\frac{d}{dt}\right]^{p\times q}$ *such that* $w(t) \in \mathcal{B} \subseteq \mathcal{E}^q$ *if and only if* $R\left(\frac{d}{dt}\right) w = 0$*. Following remark 2.24 we may write*

$$\mathcal{B} = \ker_{\mathcal{E}} R\left(\tfrac{d}{dt}\right).$$

A linear smooth system has a ***differential image representation*** *if its behavior is the image of a matrix of polynomial differential operators* $M\left(\frac{d}{dt}\right) \in \mathbb{R}\left[\frac{d}{dt}\right]^{q\times d}$*; i.e.* $w(t) \in \mathcal{B} \subseteq \mathcal{E}^q$ *if and only if there exists* $v(t) \in \mathcal{E}^d$ *such that* $w = M\left(\frac{d}{dt}\right) v$*. Again, more compactly, this definition is equivalent to*

$$\mathcal{B} = \operatorname{im}_{\mathcal{E}} M\left(\tfrac{d}{dt}\right).$$

Equations get often a simpler structure if we introduce some auxiliary variables called **latent variables**; indeed they are sometimes necessary, but not explicit in a certain sense (e.g. state variables in input/output system or internal energy and entropy in thermodynamics).

Following [Wil89], a **dynamical system with latent variables** is a quadruple $\Sigma_i = (T, W, V, \mathcal{B}_i)$ where V is the signal space of the latent variables and $\mathcal{B}_i \subseteq (W \times V)^T$. If we introduce the projection $P_w : W \times V \to W$, $(w, v) \mapsto P_w(w, v) = w$, then we get the **induced dynamical system** $\Sigma_e = (T, W, P_w\mathcal{B}_i)$. The behavior of this system is often called **external behavior**; $\mathcal{B}_i$ is called **internal behavior**.

If the external behavior $\mathcal{B}$ admits a kernel representation the procedure that constructs this representation starting from the given internal behavior is called **latent variables elimination** .

If we restrict ourselves to linear smooth systems, we have the following:

Definition 3.7. *A linear smooth system has a* ***differential latent variables representation*** *if there exist two matrices of polynomial differential operators:*

$R\left(\frac{d}{dt}\right) \in \mathbb{R}\left[\frac{d}{dt}\right]^{p\times q}$ and $M\left(\frac{d}{dt}\right) \in \mathbb{R}\left[\frac{d}{dt}\right]^{p\times d}$ such that

$$w(t) \in \mathcal{B} \subseteq \mathcal{E}^q \;\Leftrightarrow\; \exists v(t) \in \mathcal{E}^d \text{ such that } R\left(\tfrac{d}{dt}\right) w = M\left(\tfrac{d}{dt}\right) v.$$

In other words there exists a linear smooth system with latent variables whose external behavior coincides with $\mathcal{B}$.

A standard linear input/state/output system like

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned}$$

fits naturally into the behavioral framework. Actually, considering the state as latent variables, it can be rewritten as

$$\begin{bmatrix} B & 0 \\ -D & I \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix} = \begin{bmatrix} \frac{d}{dt}I - A \\ C \end{bmatrix} x.$$

Its external behavior corresponds to the input/output system.

Remark 3.8. *If we consider a linear dynamical system whose trajectories are continuous functions (therefore every differential equation must be intended in a distributional sense (see (2.1)), even a very elementary behavior with latent variable $v(t)$ does not admit latent variables elimination:*

$$\mathcal{B} = \left\{ \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix} \in C(\mathbb{R}, \mathbb{R}^2) : \exists v(t) \in C(\mathbb{R}, \mathbb{R}),\ \begin{bmatrix} 1 & -1 \\ 0 & \frac{d}{dt} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} v \right\} \tag{3.6}$$

contains trajectories that are continuous and differentiable; there is no kernel representation of $\mathcal{B}$ since this kind of representation cannot give rise to a 'differentiability' constraint.

On the contrary, linear smooth systems with latent variables always admit a differential kernel representation (see [Rap98, ch. 2.5] and references therein). If the same equations (3.6) were the equations defining a linear smooth system, i.e. if $w_1(t)$, $w_2(t)$ and $v(t)$ were smooth functions, then we would have

$$\mathcal{B} = \left\{ \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix} \in \mathcal{E}^2 : \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = 0 \right\}.$$

Linear smooth systems that admit kernel, image or latent variable differential representations will be called **differential systems** and **differential behaviors** their behaviors.

Latent variables elimination is just one of the problems related to the representation of the behavior that this approach poses: more generally we could ask: if we are given two behaviors with different representations then the inclusion $\mathcal{B}_1 \subseteq \mathcal{B}_2$ has an algebraic counterpart?

Theory developed on differential systems gives sufficiently exhaustive answers; sections 6.2 and 7.1 contains theorems that generalize the following ones:

Theorem 3.9. *If $\mathcal{B}_1$ and $\mathcal{B}_2$ are behaviors with differential kernel representations given by matrices $R_1\left(\frac{d}{dt}\right) \in R\left[\frac{d}{dt}\right]^{p_1 \times q}$ and $R_2\left(\frac{d}{dt}\right) \in R\left[\frac{d}{dt}\right]^{p_2 \times q}$, then $\mathcal{B}_1 \subseteq \mathcal{B}_2$ if and only if there is a matrix $X\left(\frac{d}{dt}\right) \in R\left[\frac{d}{dt}\right]^{p_2 \times p_1}$ such that $X\left(\frac{d}{dt}\right) R_1\left(\frac{d}{dt}\right) = R_2\left(\frac{d}{dt}\right)$ (see for similar statements [PW97, thm. 2.5.4, 3.6.2]).*

Theorem 3.10. *If $\mathcal{B}$ has a differential image representation, then it also admits a differential kernel representation.*

The converse of this theorem does not always hold: we need an additional property, controllability.

3.1.3 Controllability of behaviors

In this section we aim at introducing a structural property of dynamical systems, i.e. a property that does not depend on the particular representation of the behavior.

Definition 3.11. *A linear time invariant system $\Sigma = (\mathbb{R}, \mathbb{R}^q, \mathcal{B})$ is **controllable** if*

$$\forall w_1(t), w_2(t) \in \mathcal{B}\ \exists \bar{w}(t) \in \mathcal{B},\ \tau \geq 0 \text{ such that } \bar{w}(t) = \begin{cases} w_1(t) & t \leq 0 \\ w_2(t-\tau) & t \geq \tau \end{cases} \tag{3.7}$$

Loosely speaking, given two trajectories of a controllable behavior, there exist a trajectory that shares the 'past' with the first one and the 'future' with the second one.

The notion of classical controllability of linear system theory and behavioral controllability are formally different, but they coincide: indeed a fundamental theorem [PW97, thm. 5.2.10] states that

Theorem 3.12. *A differential behavior with kernel representation given by a matrix $R\left(\frac{d}{dt}\right)$ is controllable if and only if the rank of $R(\lambda)$ does not change for every $\lambda \in \mathbb{C}$.*

This reduces to the well-known Popov-Belevith-Hautus test for controllability when $R\left(\frac{d}{dt}\right) = \left[\frac{d}{dt}I - A \quad -B\right]$, which is the matrix operator used in the kernel representation of the system $\dot{x} = Ax + Bu$.

Another useful and elegant property of controllable differential behaviors is that they admit an image representation and vice versa [PW97, thm. 6.6]:

Theorem 3.13. *A differential behavior is controllable if and only if it has a differential image representation:*

$$\exists M\left(\tfrac{d}{dt}\right) \in \mathbb{R}\left[\tfrac{d}{dt}\right]^{q \times d} \text{ such that } \mathcal{B} = \operatorname{im}_{\mathcal{E}} M\left(\tfrac{d}{dt}\right).$$

A simple consequence of this theorem is that:

Proposition 3.14. *A differential system is controllable if and only if*

$$\forall \tau > 0,\, \forall w(t) \in \mathcal{B}\ \exists \bar{w}(t) \in \mathcal{B} \text{ such that } \bar{w}(t) = \begin{cases} w(t) & t < 0 \\ 0 & t \geq \tau \end{cases}$$

Before we prove this proposition we state a useful lemma:

Lemma 3.15. *There is a smooth function $\psi(t) \in \mathcal{E}$ such that $\psi(t) = 0$ for $t \leq 0$ and that $\psi(t) = 1$ for $t \geq 1$.*

Proof. Let us construct directly $\psi(t)$ from the well-known smooth function [PW97, rem. 2.4.5] $\tilde{\psi}(t)$ with compact support $[-1, 1]$:

$$\tilde{\psi}(t) = \begin{cases} e^{\frac{1}{t^2-1}} & |t| < 1 \\ 0 & |t| \geq 1 \end{cases} \qquad \Rightarrow \qquad \psi(t) = \frac{\int_{-\infty}^{2t-1} \tilde{\psi}(x)\, dx}{\int_{-\infty}^{+\infty} \tilde{\psi}(x)\, dx}.$$

□

Proof of proposition 3.14. If the system is controllable, then this proposition is trivially verified because for every $\tau > 0$, using $\psi(t)$ of lemma 3.15, we can define the smooth function

$$\tilde{v}(t) = \psi\left(\frac{\tau - t}{\tau}\right) \text{ such that } \tilde{v}(t) = 1 \text{ for } t \leq 0 \text{ and } \tilde{v}(t) = 0 \text{ for } t \geq \tau.$$

Since $w(t) \in \mathcal{B}$ implies $w = Mv$ for some smooth function $v(t)$ by theorem 3.13, $\bar{w} = M(v\tilde{v})$ is the desired trajectory.

Conversely if for any $\tau > 0$ and $w(t) \in \mathcal{B}$ we have a $\bar{w}(t)$ steering it to zero, then given $w_1(t)$, $w_2(t) \in \mathcal{B}$ and a fixed $\bar{\tau} > 0$ we can take $0 < \tau < \bar{\tau}$ and $w(t) = w_1(t) - w_2(t - \bar{\tau})$: in this way the function $\bar{w}(t) + w_2(t - \bar{\tau})$ is equal to $w_1(t)$ for $t \leq 0$ and to $w_2(t - \bar{\tau})$ for $t \geq \tau$, *a fortiori* for $t \geq \bar{\tau}$. □

Strictly related to controllability is the concept of autonomous system:

Definition 3.16. *A linear time invariant system $\Sigma = (\mathbb{R}, \mathbb{R}^q, \mathcal{B})$ is* ***autonomous*** *if given any $w_1(t), w_2(t) \in \mathcal{B}$, $w_1(t) = w_2(t)$ $\forall t \leq 0$ implies that $w_1(t) = w_2(t)$ $\forall t \in \mathbb{R}$.*

It is really trivial to verify that a controllable system is not autonomous and vice versa. It is more difficult to prove [PW97, thm. 5.2.14] that

Theorem 3.17. *Given a differential behavior $\mathcal{B}$, it is always possible to find two differential behaviors $\mathcal{B}_a$ autonomous and $\mathcal{B}_c$ controllable, such that $\mathcal{B} = \mathcal{B}_a \oplus \mathcal{B}_c$; the controllable subsystem is uniquely determined.*

3.1.4 Delay–differential equations

The main goal of this thesis is to show how delay–differential systems may be treated with behavioral techniques. First of all we must give an exact definition of this particular class of linear systems.

Delay–differential equations are functional equations involving both derivative and shift operators. Since derivative and shift operators commute, the composition of these operators is independent of the order; the operator derivating n times and shifting m times with time shift τ is therefore denoted by $\left(\frac{d}{dt}\right)^n \sigma_\tau^m$ or else $\left(\frac{d}{dt}\right)^n \sigma_{m\tau}$.

If every time shift is a multiple of one single delay then the system is said to have **commensurate delays** or **one delay**; in this case, without loosing generality, we will consider delays as belonging to $\mathbb{Z}$ and denote the shift operator as σ (assuming implicitly that the delay is unitary.)

If the time shifts do not have a common multiple, i.e. there are (necessarily more than one) **incommensurate delays**, it is customary (see proposition 4.27) to consider the set of delays as a subset of a (finite and direct) sum of modules like

$$\tau_i\mathbb{Z} = \{k\tau_i : k \in \mathbb{Z}\} = \{t \in \mathbb{R} : t/\tau_i \in \mathbb{Z}\}.$$

Since the sum of these modules is a free module, it has a well defined rank, i.e. the number of elements $\{\tau_i\}$ of its basis: that is why such systems are also said to have $\boldsymbol{m}$ **delays** if m is the rank mentioned above.

For the same reason a single time-delay is a (uniquely determined) $\mathbb{Z}$–linear combination of the basis $\tau_1, \dots, \tau_m \in \mathbb{R}$:

$$\tau = \sum_i k_i\tau_i \ (k_i \in \mathbb{Z}) \ \Rightarrow \ \sigma_\tau = \sigma_{k_1\tau_1} \cdots \sigma_{k_m\tau_m} = \sigma_{\tau_1}^{k_1} \cdots \sigma_{\tau_m}^{k_m}. \tag{3.8}$$

In this case an abbreviated notation like $\boldsymbol{\sigma} = (\sigma_{\tau_1} \cdots \sigma_{\tau_m})$ will be often used; we shall write the operator of equation (3.8) as $\boldsymbol{\sigma}^{\boldsymbol{k}}$ and even something like $\boldsymbol{k\tau} = \sum_i k_i\tau_i$. The $\boldsymbol{k} \in \mathbb{Z}^m$ are also known as multi-indices.

A suitable set of operators for dealing with delay–differential equations having m delays $\tau_1, \dots, \tau_m$ could be the ring of **delay–differential polynomials** in $m+1$ variables $\mathbb{R}\left[\frac{d}{dt}, \boldsymbol{\sigma}\right]$ such that

$$\left(r\left(\tfrac{d}{dt}, \boldsymbol{\sigma}\right) w\right)(t) = \left(\sum_{i,\boldsymbol{k}} a_{i\boldsymbol{k}} \tfrac{d^i}{dt^i} \boldsymbol{\sigma}^{\boldsymbol{k}} w\right)(t) = \sum_{i,\boldsymbol{k}} a_{i\boldsymbol{k}} w^{(i)}(t - \boldsymbol{k\tau}) \tag{3.9}$$

with indices i and $\boldsymbol{k}$ belonging to finite sets.

Remark 3.18. *As regards the equivalence problem, we note immediately that the operators just defined are not adequate: in fact the equations*

$$x(t-1) = u(t-1) \ \Leftrightarrow \ \sigma x(t) = \sigma u(t) \ \text{and} \ x(t) = u(t)$$

represent the same behavior but the kernel representation of the second one, $[1\ {-1}]$*, cannot be deduced algebraically from the the representation* $[\sigma\ -\sigma]$.

A simple way to obviate this difficulty consists in introducing a larger ring of operators, namely analogous to the ring commonly used with discrete–time, especially multi–dimensional, behaviors: the ring of **delay–differential Laurent polynomials** $\mathbb{R}\left[\frac{d}{dt}, \boldsymbol{\sigma}, \boldsymbol{\sigma}^{-1}\right]$, containing polynomial with both positive and negative powers of the delay operators.[1]

Remark 3.19. *Another simple example shows that neither Laurent polynomials suffices: if we consider the behaviors represented by the equations*

$$\frac{d}{dt}x = 0 \text{ and } (1-\sigma)x(t) = x(t) - x(t-1) = 0$$

the first one (constant functions) is obviously strictly included into the second one (periodic functions) but there is no polynomial $x(\frac{d}{dt}, \sigma)$ such that, as in theorem 3.9, $x(\frac{d}{dt}, \sigma)\frac{d}{dt} = 1 - \sigma$.

To overcome this much harder problem we have to extend our operator ring to include more general operators, as we will show in section 4.2.3 and in chapters 6 and 7.

3.1.5 Convolutional behaviors

The most general class of behaviors that will be discussed in this thesis are convolutional behaviors:

Definition 3.20. *A* ***convolutional behavior*** *in* ***kernel representation*** *is a linear smooth system that satisfies the following convolutional equation*

$$\mathcal{B} = \ker_{\mathcal{E}} R(s) \subseteq \mathcal{E}^q \tag{3.10}$$

where $R(s) \in \mathcal{A}^{p\times q}$ (see remark 2.24).

Analogously the ***image representation*** *of a convolutional behavior is*

$$\mathcal{B} = \operatorname{im}_{\mathcal{E}} M(s) \subseteq \mathcal{E}^q \tag{3.11}$$

where $M(s)$ belongs to $\mathcal{A}^{q\times d}$.

In general a ***convolutional behavior with latent variables*** *is a linear smooth system so defined:*

$$\mathcal{B} = \left\{w(t) \in \mathcal{E}^q : \ \exists x(t) \in \mathcal{E}^d, \ R(s)w = M(s)v\right\} \tag{3.12}$$

[1] We note that σ^{-1} is a well defined operator, being the shift bijective.

and $R(s) \in \mathcal{A}^{p\times q}$, $M(s) \in \mathcal{A}^{p\times d}$.

Nevertheless we are mainly interested in delay–differential systems, so we will call **delay–differential systems** those linear smooth systems that admit a **delay–differential kernel representation** a **delay–differential image representation** or a **delay–differential latent variables representation** i.e. they admit a representation of the form (3.10), (3.11) or (3.12) where $R(s)$ and $M(s)$, with suitable dimensions, take their values in subrings of $\mathcal{A}$ that correspond to delay–differential (Laurent) polynomials or to their extensions as will be explained in chapter 4.

Such systems, usually in state space form, have been investigated as systems over rings (see e.g. [Hab94]) or as infinite dimensional systems (see e.g. [Hal77]).

In both cases there has been proposed many different definitions that generalize the notion of controllability (see for instance [RW97]). Behavioral controllability, definition 3.11, depends only on trajectories, thus applies also to delay–differential systems.

The problem is the following: given a behavior with some representation, is there some algebraic criterion to check behavioral controllability (as e.g. theorem 3.12)? We will try to give an answer.

3.2 The module theoretic approach

Michel Fliess proposed in the eighties a new algebraic point of view in the analysis of dynamical systems. Starting from linear differential systems his set of definitions and theorems has grown to embrace a wide class of systems: continuous and discrete time, linear, delay differential and non linear systems; recently he and his coworkers are trying to give new insight in the world of multidimensional systems involving partial differential equations.

We will only deal with some results on delay differential systems. However, in order to understand the relation between these definitions and theorems and the ones that belong to the behavioral approach, it is necessary to follow these ideas from the beginning, i.e. from the simpler and well-known ordinary linear systems, clarifying their (sometimes too algebraic) meaning.

In this section we will follow an abstract notation: $\mathcal{R}$ is a generic domain that, as we show more precisely in chapter 5, is a ring of operators like $\mathcal{A}$ or $\mathbb{R}\left[\frac{d}{dt}\right]$. Even if we do not deal with trajectories in this section, we consider vectors in $\mathcal{R}^p$ as rows and we denote a matrix $R \in \mathcal{R}^{p\times q}$ (even more explicitly than in remark 2.24) as $\circ R$ when it is considered as a (multiplicative) operator on the left.

3.2.1 Systems are modules

Algebraic definitions and theorems that are used but not introduced in this section may be found in appendix A or in books like [Lan93], [AM69], [AF92].

Definition 3.21. *A linear system is a finitely generated $\mathcal{R}$–module $\mathcal{M}$.*

This definition is the starting point of Fliess' theory. Using standard algebraic language, we can say that a linear system is the **cokernel** of some matrix $R \in \mathcal{R}^{p\times q}$, called the **presentation matrix**, i.e.

$$\mathcal{M} = \operatorname{coker}_{\mathcal{R}} \circ R \triangleq \mathcal{R}^q / \mathcal{R}^p R. \tag{3.13}$$

Actually, a linear system is a module with finite free presentation; using an exact sequence of modules:

$$\mathcal{R}^p \xrightarrow{\circ R} \mathcal{R}^q \xrightarrow{\phi} \mathcal{M} \longrightarrow 0 \tag{3.14}$$

where ϕ is the canonical (surjective) projection

$$\phi : \mathcal{R}^q \to \mathcal{R}^q / \ker\phi = \mathcal{R}^q/\mathcal{R}^p R = \operatorname{coker}_{\mathcal{R}} \circ R = \mathcal{M}$$

When $\mathcal{R} = \mathbb{R}\left[\frac{d}{dt}\right]$, which is a Noetherian ring, then we can deduce R from the generators of $\mathcal{M}$; it is sufficient to define ϕ mapping a base of the free module $\mathcal{R}^q$ onto $\mathcal{M}$: then the rows of R are the generators of $\ker\phi$ (which is finitely generated by Noetherianity). When we cannot assure that $\ker\phi$ is finitely generated, e.g. when $\mathcal{R} = \mathcal{A}$ which is not Noetherian, we have to start directly from R.

The definition of input and output variables is rather simple: a set of elements $\mathsf{u}_1, \ldots, \mathsf{u}_m \in \mathcal{M}$, (independent) generators of the set U, is an input if the quotient module $\mathcal{M}/U$ is torsion (see A.2); every set $\{\mathsf{y}_i\} \subset \mathcal{M}$ is an output.

Similarly to a behavior, the module $\mathcal{M}$, contains input, output and latent variables without any *a priori* classification. This resemblance is not fortuitous, as we shall see in chapter 5.

Example 3.22. The well–known input/state system

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{3.15}$$

where $x(t) \in \mathcal{E}^n$ and $u(t) \in \mathcal{E}^m$ admits a (behavioral) kernel representation as $\mathcal{B} = \ker_{\mathcal{E}} R\left(\frac{d}{dt}\right) \circ$ with $R\left(\frac{d}{dt}\right) \triangleq [\frac{d}{dt}I - A \quad -B]$

The same matrix, this time operating on the left, is the presentation matrix of a module $\mathcal{M}$ that is formally equivalent to (3.15). Indeed, let $\mathcal{R} = \mathbb{R}\left[\frac{d}{dt}\right]$ and consider $\mathcal{R}^{n+m}$ generated by the basis $X_1, \ldots, X_n$, $U_1, \ldots, U_m$, (these are only symbols, elements of a basis: they have no other meaning. We may think e.g. that $\{X_i\} \cup \{U_j\}$ is simply the canonical basis of $\mathcal{R}^{n+m}$).

Every element $a\left(\frac{d}{dt}\right) \in \mathcal{R}^{m+n}$ is a linear combination of $\{X_i\}$ and $\{U_i\}$: if $X^\top = [X_1, \ldots, X_n]$ and $U^\top = [U_1, \ldots, U_m]$ then every $a\left(\frac{d}{dt}\right)$ is equal to $b\left(\frac{d}{dt}\right)\begin{bmatrix} X \\ U \end{bmatrix}$ for some $b\left(\frac{d}{dt}\right) \in \mathcal{R}^{n+m}$ (trivially $a\left(\frac{d}{dt}\right) = b\left(\frac{d}{dt}\right)$ in the case of the canonical basis).

If we suppose that $R\left(\frac{d}{dt}\right)$ is expressed with respect to this basis we may write the operator as $\circ R\left(\frac{d}{dt}\right)\begin{bmatrix} X \\ U \end{bmatrix}$. If ϕ, as in (3.14), is the canonical projection onto the quotient and $\mathsf{x}_i \triangleq \phi(X_i)$ and $\mathsf{u}_i \triangleq \phi(U_i)$ then we know that $\operatorname{im}_{\mathcal{R}} \circ R\left(\frac{d}{dt}\right)\begin{bmatrix} X \\ U \end{bmatrix} = \ker_{\mathcal{R}} \phi$, thus

$$0 = \phi\left(R\left(\tfrac{d}{dt}\right)\begin{bmatrix} X \\ U \end{bmatrix}\right) = R\left(\tfrac{d}{dt}\right)\phi\left(\begin{bmatrix} X \\ U \end{bmatrix}\right) = R\left(\tfrac{d}{dt}\right)\begin{bmatrix} \mathsf{x} \\ \mathsf{u} \end{bmatrix}$$

that is to say, by definition of $R\left(\frac{d}{dt}\right)$, that

$$\tfrac{d}{dt}\mathsf{x} = A\mathsf{x} + B\mathsf{u}.$$

Note, once again, that last two expressions, albeit similar to a kernel representation and to (3.15), do not have the same *concrete* sense: $\mathsf{x} = \phi(X)$ and $\mathsf{u} = \phi(U)$ are not trajectories, but purely algebraic symbols!

They simply generate $\mathcal{M}$ while $R\left(\frac{d}{dt}\right)$ defines the (algebraic) relations between them.

♣

3.2.2 Abstract controllabilities

Controllability does not have in this context a well defined definition, it is rather a family of definitions.

Definition 3.23. *For any $\mathcal{R}$-algebra $\mathcal{G}$, a system $\mathcal{M}$ is $\mathcal{G}$–torsion free (or $\mathcal{G}$–projective or $\mathcal{G}$–free) controllable if the $\mathcal{G}$–module $\mathcal{G} \otimes_{\mathcal{R}} \mathcal{M}$ is torsion free (or projective or free).*

It is necessary to explain many things: first of all the meaning of $\mathcal{G} \otimes_{\mathcal{R}} \mathcal{M}$.

The very definition of the tensor product of two modules over the same ring is in appendix A.2; however it is not so helpful: in this case tensor product is a standard way to do a *change of base ring*, i.e. it is a sort of algebraic immersion of the scalars of the module $\mathcal{M}$ into a greater ring, namely $\mathcal{G}$.

The particular structure of $\mathcal{M} = \operatorname{coker}_{\mathcal{R}} \circ R$ is essential in finding a simpler form for $\mathcal{G} \otimes_{\mathcal{R}} \mathcal{M}$: the functor $\mathcal{G}\otimes_{\mathcal{R}} : \mathcal{R}\text{-modules} \to \mathcal{G}\text{-modules}$ preserves cokernels [Coh95, p. 146]: $\mathcal{G} \otimes_{\mathcal{R}} \operatorname{coker}_{\mathcal{R}} \circ R = \operatorname{coker}_{\mathcal{G}\otimes_{\mathcal{R}}\mathcal{R}} \circ R$. In other words, since $\mathcal{G} \otimes_{\mathcal{R}} \mathcal{R}^n \cong \mathcal{G}^n$ [Coh95, p. 159], $\mathcal{G} \otimes_{\mathcal{R}} \operatorname{coker}_{\mathcal{R}} \circ R = \operatorname{coker}_{\mathcal{G}} \circ R$.

Indeed, using exact sequences, since $\mathcal{G}\otimes_{\mathcal{R}}$ is a right exact (covariant) functor [Bro92, p. 144], we can apply it to sequence (3.14) and get

$$\begin{array}{ccccccccc} \mathcal{R}^p & \xrightarrow{\circ R} & \mathcal{R}^q & \xrightarrow{\phi} & \mathcal{M} & \longrightarrow & 0 & & \\ & & & & & & & \Big\downarrow \mathcal{G}\otimes_{\mathcal{R}} \cdot & \\ \mathcal{G}^p & \xrightarrow{\circ R} & \mathcal{G}^q & \xrightarrow{\phi} & \mathcal{G}\otimes_{\mathcal{R}}\mathcal{M} & \longrightarrow & 0 & & \end{array}$$

hence we obtain the very simple result:

$$\mathcal{G} \otimes_{\mathcal{R}} \mathcal{M} \cong \mathcal{G}^q / \mathcal{G}^p R.$$

Afterwards, the three controllability properties are not independent: a free module is projective and a projective one is torsion free. Moreover torsion free modules over principal ideal domains are free: polynomials in one variable, for instance, are principal ideal domains. Therefore ordinary linear time-invariant systems, both continuous and discrete time, which may be represented by modules over $\mathcal{R} \cong \mathbb{R}[s]$, are simply said controllable if they satisfy one of the definitions of controllability (with $\mathcal{G} = \mathcal{R}$).

Freeness, the strongest property, implies that the module $\mathcal{G}^q/\mathcal{G}^p R$ has a basis and is isomorphic to $\mathcal{G}^{q-r}$ where r is the rank of R: the module $\mathcal{G} \otimes_{\mathcal{R}} \mathcal{M}$ behaves like a vector space.

Example 3.24. As we saw in example 3.22, in the input/state case we have $p = n$ (the dimension of the state) and $q = n + m$. If the system is controllable, since $R\left(\frac{d}{dt}\right)$ is clearly a full row rank matrix $p = r$, we obtain that the dimension of $\mathcal{M}$ is $q - p = m$, that is exactly the dimension of the inputs.

Being the inputs independent, they are a basis of the module; in other words every other element of the system (state variables and outputs) is an $\mathbb{R}\left[\frac{d}{dt}\right]$–linear combination of the inputs, i.e. a linear function of the inputs and their derivatives. ♣

A more down to eath meaning of projective controllability will be shown in section 7.2.1; we only recall here a useful basic property of projective modules [AF92, ex. 11, p. 203].

Lemma 3.25 (Dual basis). *An $\mathcal{R}$–module P finitely generated is projective if and only if there exist $\{x_i\} \subset P$ and $f_i \in \mathrm{Hom}_{\mathcal{R}}(P, \mathcal{R})$, $i = 1, \ldots, n$ such that*

$$\text{for every } x \in P,\ x = \sum_{i=1}^{n} f_i(x) x_i.$$

The weakest of the three properties, torsion free-controllability, will be further investigated in section 7.2.1. Here we simply show what torsion freeness of a module implies in this specific context: $\mathcal{G} \otimes_{\mathcal{R}} \mathcal{M}$ is torsion free if any non zero element cannot be annihilated by non zero elements in $\mathcal{G}$. In other words, if $\mathcal{M} = \mathrm{coker}_{\mathcal{R}} \circ R$ with $R \in \mathcal{R}^{p \times q}$,

$$\text{if } g \in \mathcal{G} \text{ and } gx \in \mathcal{G}^p R \text{ then } x \in \mathcal{G}^p R.$$

The module $\mathcal{G}^p R$ is the zero of $\mathcal{G} \otimes_{\mathcal{R}} \mathcal{M}$ and $x + \mathcal{G}^p R$ is zero in $\mathcal{G} \otimes_{\mathcal{R}} \mathcal{M}$ if and only if $x \in \mathcal{G}^p R$; thus previous equation states that $gx + \mathcal{G}^p R$ is zero only if $x + \mathcal{G}^p R$ itself is zero.

We omitted a great amount of definitions and interesting results of this algebraic theory of dynamical systems: we refer the reader to the papers written by Fliess and his coworkers (for example [Mou98a], [FM95], [FMRR95], [Mou98b]).

3.2.3 Controllability for delay–differential systems

We have seen that the operator ring corresponding to delay–differential equations with m non commensurate delays is $\mathcal{R} = \mathbb{R}\left[\frac{d}{dt}, \boldsymbol{\sigma}\right]$ which is isomorphic (proposition 4.27) to a polynomial ring $\mathbb{R}[s, \boldsymbol{z}]$ in $m+1$ variables. In this case every projective module is free thanks to the well-known Quillen–Suslin theorem [Lan93, p. 850] and we have only $\mathcal{R}$-free or torsion free controllable systems.

The literature about delay–differential systems is a subset of the wider one about system over rings which, beyond having embraced various branches of mathematics, gave rise to many notions of controllability.

For example, $\mathbb{R}\left[\frac{d}{dt}, \boldsymbol{\sigma}\right]$-free controllability is equivalent to *reachability* whereas $\mathbb{R}(\boldsymbol{\sigma})\left[\frac{d}{dt}\right]$-free controllability is also known as *weak controllability* (see [Mou98a]).

Because of theorems 3.12 and 3.13 we are mainly interested in a stronger form of controllability that was defined for the first time for a particular class of systems:

Definition 3.26. *The system*

$$\dot{x} = A(\sigma)x + Bu$$

where $A(s) \in \mathbb{R}[s]^{n\times n}$, *is called* ***spectrally controllable*** *if*

$$\operatorname{rank}[\lambda I - A(e^{-\lambda}) \quad -B] = n, \ \forall \lambda \in \mathbb{C}.$$

This definition has been extended to other similar cases (see [RW97]). We will use *spectral controllability* to indicate a 'constant rank condition on $\mathbb{C}$': in [Mou98a] the exact definition is the following: if $R \in \mathbb{R}\left[\frac{d}{dt}, \boldsymbol{\sigma}\right]^{p\times q}$ is the full row rank presentation matrix of the linear delay system $\mathcal{M}$, it is **spectrally controllable** if

$$\operatorname{rank}_{\mathbb{C}} R\left(\lambda, e^{\boldsymbol{\tau}\lambda}\right) = p, \ \forall \lambda \in \mathbb{C}.$$

We may generalize further the above definition to generic operator rings:

Definition 3.27. *If we are given a matrix* $R(s) \in \mathcal{A}^{p\times q}$ *with rank* r, *the linear system* $\mathcal{M} = \operatorname{coker}_{\mathcal{A}} \circ R(s)$ *is spectrally controllable if*

$$\operatorname{rank}_{\mathbb{C}} R(\lambda) = r, \ \forall \lambda \in \mathbb{C}. \tag{3.16}$$

In this case, by duality reasons that will be explained in chapter 5, we shall *say that also the behavior in kernel representation* $\mathcal{B} = \ker_{\mathcal{E}} R(s)$ *is spectrally controllable.*

Chapter 4

Operator rings

We have already introduced many classes of operators, simple differential operators, delay–differential operators, distributions with compact support. A key point is that there is a link between these operators and particular sets of holomorphic functions in $\mathbb{C}$.

The aim of this chapter is at giving a more unitary treatment of these and other new operator rings along with their most important properties.

4.1 Algebraic preliminaries

Throughout this section we will use the symbol $\mathcal{R}$ to denote a domain (of operators). Following the notation of remark 2.24, the symbol $\circ R$ states that the matrix $R \in \mathcal{R}^{p\times q}$ operates on the left and $R\circ$ on the right.

4.1.1 Smith form

The Smith form is a very advantageous way to factorize matrices over rings. Unfortunately not every matrix admits such a factorization but there are important rings (e.g. polynomials in one variable) such that every matrix taking values in them has a Smith form.

Definition 4.1. *The matrix $R \in \mathcal{R}^{p\times q}$ admits a* ***Smith form*** *if $R = P\bar{R}Q$ where the matrices $P \in \mathcal{R}^{p\times p}$ and $Q \in \mathcal{R}^{q\times q}$ are invertible in $\mathcal{R}$ and $\bar{R} \in \mathcal{R}^{p\times q}$ has $r \leq \min\{p, q\}$ non zero entries only on the main diagonal; we will denote with $\breve{R}$ the maximal square full rank submatrix of $\bar{R}$ that contains on its diagonal every non zero element of $\bar{R}$.*

Moreover, the (non-zero) elements on the diagonal of $\breve{R}$ are unique up to a multiplication by a non-zero constant and are ordered in such a way that each element divides the following one.

Remark 4.2. *It must be noted that over any field (e.g. the field of fractions of every domain) the Smith form always exists with $\breve{R} = I_r$.*

Definition 4.3. *A domain in which every matrix admits a Smith form is called an* ***elementary divisor domain****.*

Once the existence of the Smith form of a matrix has been established, there are several ways to write it and many simple consequences. They will be used quite often throughout the thesis and are collected here for future references.

The following equation provides a more detailed description of the Smith form. We remind that R and $\bar{R}$ are $p \times q$ matrices, while P, Q and $\breve{R}$ are square with dimension p, q and r, rank of R.

$$R = P\bar{R}Q = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} \breve{R} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} = \begin{bmatrix} P_{11} \\ P_{21} \end{bmatrix} \breve{R} \begin{bmatrix} Q_{11} & Q_{12} \end{bmatrix}. \tag{4.1}$$

We can express $\bar{R}$ and $\breve{R}$ using the inverses of P and Q, respectively U and V:

$$\bar{R} = \begin{bmatrix} \breve{R} & 0 \\ 0 & 0 \end{bmatrix} = URV = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} R \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \Rightarrow \breve{R} = \begin{bmatrix} U_{11} & U_{12} \end{bmatrix} R \begin{bmatrix} V_{11} \\ V_{21} \end{bmatrix}. \tag{4.2}$$

Using a shorthand notation (appending a star denotes a whole 'block-row' or 'block-column'), since $UP = PU = I_p$ and $QV = VQ = I_q$:

$$R = P_{*1}\breve{R}Q_{1*},\ \breve{R} = U_{1*}RV_{*1},\ U_{1*}P_{*1} = I_r = Q_{1*}V_{*1}, \tag{4.3}$$

$$P_{*1}U_{1*} + P_{*2}U_{2*} = I_p,\ V_{*1}Q_{1*} + V_{*2}Q_{2*} = I_q. \tag{4.4}$$

Lemma 4.4. *If the elements on the diagonal of $\breve{R}$ are left-injective*[1] *or right-*

[1] *As we already pointed out, $R\circ$ and $\circ R$ may have different operatorial meaning; therefore the side must be explicitly indicated here. For example we have seen that we can deal with linear differential systems employing the operator ring $\mathcal{R} = \mathbb{R}[\frac{d}{dt}]$: if we take non zero elements, the left operation (multiplication of polynomials) is clearly injective, since the ring is a domain, and not surjective; the right operation (polynomial differential operator on $\mathcal{E}$) is not injective but, as we shall see in section 4.2.2, surjective.*

injective elements of $\mathcal{R}$ (if and only if $\circ\breve{R}$ or $\breve{R}\circ$ is injective), then

$$\ker\circ R = \ker\circ P_{*1} = \operatorname{im}\circ U_{2*} \text{ or, respectively, } \ker R\circ = \ker Q_{1*}\circ = \operatorname{im} V_{*2}\circ. \quad (4.5)$$

If the elements on the diagonal of $\breve{R}$ are left-surjective or right-surjective elements of $\mathcal{R}$ (if and only if $\circ\breve{R}$ or $\breve{R}\circ$ is surjective):

$$\operatorname{im}\circ R = \operatorname{im}\circ Q_{1*} = \ker\circ V_{*2} \text{ or, respectively, } \operatorname{im} R\circ = \operatorname{im} P_{*1}\circ = \ker U_{2*}\circ. \quad (4.6)$$

If one relation in (4.5) [or in (4.6)] is satisfied, the kernel [or image] of R is also kernel of a surjective matrix and image of an injective matrix.

Proof. Let us prove the first equality of (4.5): from (4.2) we see that

$$U_{2*}R = 0 \text{ therefore } \operatorname{im}\circ U_{2*} \subseteq \ker\circ R.$$

From (4.3) we have

$$vR = 0 \;\Rightarrow\; vRV_{*1} = 0 \;\Rightarrow\; vP_{*1}\breve{R} = 0$$

but, since $\breve{R}$ is a diagonal matrix with elements d_i, then the i-th component of $vP_{*1}\breve{R}$ is the product of the i-th component of vP_{*1} and of the i-th element of the diagonal of $\breve{R}$, $d_i \neq 0$. Hence, if every d_i is left-injective, the product is zero if and only if $vP_{*1} = 0$, therefore $\ker\circ R \subseteq \ker\circ P_{*1}$. Then, using (4.4) we obtain

$$v \in \ker\circ P_{*1} \;\Rightarrow\; v = vI_p = vP_{*1}U_{1*} + vP_{*2}U_{2*} = vP_{*2}U_{2*} \;\Rightarrow v \in \operatorname{im}\circ U_{2*}.$$

The second equation in (4.5) is proven in a similar way.

Relations (4.6) are ‘dual’: from $RV_{*2} = 0$ we obtain $\operatorname{im}\circ R \subseteq \ker\circ V_{*2}$; from (4.4)

$$x \in \ker\circ V_{*2} \;\Rightarrow\; x = xV_{*1}Q_{1*} + xV_{*2}Q_{2*} = xV_{*1}Q_{1*} \;\Rightarrow\; \ker\circ V_{*2} \subseteq \operatorname{im}\circ Q_{1*}.$$

Finally the operator $\circ P_{*1}$ is surjective since from (4.3) we have

$$U_{1*}P_{*1} = I_r \;\Rightarrow\; \forall a\ \exists b = aU_{1*} \text{ such that } bP_{*1} = a. \quad (4.7)$$

Then as soon as every single element on the diagonal of $\breve{R}$ is left-surjective, so does $\circ P_{*1}\breve{R}$ and $\operatorname{im}\circ Q_{1*} \subseteq \operatorname{im}\circ P_{*1}\breve{R}Q_{1*} = \operatorname{im}\circ R$.

Surjectivity of $\circ V_{*2}$, $Q_{1*}\circ$ and $U_{2*}\circ$ may be shown as we did for $\circ P_{*1}$ using relations similar to (4.7). The same equations prove that those operators are injective 'on the other side': for instance P_{*1° is injective since if $P_{*1}a = 0$ then $0 = U_{1*}P_{*1}a = I_r a = a$. □

Remark 4.5. *Invertible elements are both injective and surjective, hence if the elements of $\breve{R}$ are such, e.g. when $\mathcal{R}$ is a field or $\breve{R}$ is constant, equations (4.5) and (4.6) are true.*

4.1.2 Bézout equation

The so called **Bézout equation** is a fundamental tool in algebraic system theory: n elements in a ring $x_i \in \mathcal{R}$ satisfy a Bézout equation if there are n element $a_i \in \mathcal{R}$ such that $\sum_i a_i x_i = 1$. In this case the set $\{x_i\}$ generates the whole ring.

Definition 4.6. *A domain in which every finitely generated ideal is principal is called* ***Bézout domain****.*

If $\mathcal{R}$ is a Bézout domain, any n elements $x_i \in \mathcal{R}$ generate a principal ideal $(x_1, \ldots, x_n)_\mathcal{R} = (f)_\mathcal{R}$: f is the greatest common factor of x_i. In fact $f = \sum a_i x_i$ and also $x_i = y_i f$ with $a_i, y_i \in \mathcal{R}$. The factors y_i of x_i are coprime, that is to say they have only 1 as common factor, thus satisfy a Bézout equation:

$$f = \sum a_i x_i = \sum a_i y_i f \;\Rightarrow\; 1 = \sum a_i y_i.$$

The relation between elementary divisor domains and Bézout domains is still an open question. We know that

Proposition 4.7. *If $\mathcal{R}$ is an elementary divisor domain, then it is also a Bézout domain.*

Proof. Let X be a row vector with n components $x_i \in \mathcal{R}$. The Smith form of X is given by $P \in \mathcal{R}$, invertible, $0 \neq \breve{X} \in \mathcal{R}$ and a square $n \times n$ invertible matrix Q.

Without loss of generality we can say that $P = 1$ and the determinant of Q is also 1. In this case, if $f = \breve{X}$, y_i are the elements of the first row of Q and a_i are the n minors of order $n-1$ of the last rows of Q, we have exactly the relations we found for generic elements in a Bézout domain: $x_i = y_i f$ (from $X = P\bar{X}Q$) and $\sum a_i y_i = 1$ (the determinant of Q) thus $f = \sum a_i x_i$ and so $(f)_\mathcal{R} = (x_1, \ldots, x_n)_\mathcal{R}$. □

On the other hand every known Bézout domain is an elementary divisor domain but there is no proof of this equivalence.

Scalar Bézout equations are also useful to analyse properties of matrices.

Definition 4.8. *Given any matrix $R \in \mathcal{R}^{p\times q}$ and $r \leq \min\{p,q\}$, the* ***compound matrix*** *of order r of R, $C_r(R)$, is a matrix with $\binom{p}{r}$ rows and $\binom{q}{r}$ columns whose elements are the minors of order r of R.*

Once a way to order the minors of R into rows and columns of $C_r(R)$ has been fixed, a fundamental property of the compound matrix is the following (see e.g. [RM71])

$$XY = Z \;\Rightarrow\; C_r(X)C_r(Y) = C_r(Z).$$

Proposition 4.9. *A full row (column) rank matrix is right (left) invertible if and only if its maximal minors satisfy a Bézout equation.*

Proof. If a matrix $R \in \mathcal{R}^{p\times q}$ is right invertible, i.e. there is a matrix $X \in \mathcal{R}^{q\times p}$ such that $RX = I_p$ then $C_p(R)C_p(X) = C_p(I_p) = 1$; since $C_p(R)$ and $C_p(X)$ are a row and a column vector containing the $p \times p$ minors of R and X the sufficiency has been proved.

Conversely, if m_i are the minors of R satisfying a Bézout equation, the we can write $m_i = \det RS_i$ where S_i is a matrix built only with 0 and 1 that 'chooses' the desired columns of R. So, using the well-known formula $A \operatorname{adj}(A) = I \det(A)$,

$$\begin{aligned} I &= \sum x_i I m_i = \sum x_i I \det(RS_i) = \sum x_i RS_i \operatorname{adj}(RS_i) \\ &= R \sum x_i S_i \operatorname{adj}(RS_i) = RX. \end{aligned}$$

If R is full column rank, the proof symmetric to this one. □

4.1.3 Generalized inverses

If R does not have full rank, then it is still possible to find a nice result using generalized inverses [BIG74].

Definition 4.10. *A matrix $G \in \mathcal{R}^{q\times p}$ is the $\{1\}$-inverse[2] of the matrix $R \in \mathcal{R}^{p\times q}$ if $RGR = R$; it is a $\{1,2\}$-inverse if it is a $\{1\}$-inverse and $GRG = G$.*

Lemma 4.11. *A matrix has a $\{1\}$-inverse if and only if it has a $\{1,2\}$-inverse.*

Proof. Let $RGR = R$. If we let $\bar{G} = GRG$ then

$$R\bar{G}R = RGRGR = RGR = R$$

hence $\bar{G}$ is a $\{1\}$-inverse. Moreover

$$\bar{G}R\bar{G} = GRGRGRG = GRGRG = GRG = \bar{G}$$

showing that $\bar{G}$ is a $\{1,2\}$-inverse. □

This proposition permits us to use always $\{1,2\}$-inverses which will be simply called **generalized inverses**.

Theorem 4.12. *A matrix has a generalized inverse if and only if its minors satisfy a Bézout equation.*

Theorem 4.12 is in [BR83]; [Son80] showed a very similar, more abstract, theorem. Other related results are in [BBRMP90] and in [MPBR96].

The following results generalize further theorem 4.12:

Theorem 4.13. *Suppose that R has rank r. If a linear combination of its $r \times r$ minors is equal to a, then there is a matrix G such that $RGR = aR$; if $RGR = aR$, there is a linear combination of the $r \times r$ minors of R equal to a^r.*

Proof. The first part follows from the proof, quite complicated, of [BR83, thm. 8], simply replacing the '1' of the Bézout equation with a. The second part is simpler: if we use compound matrices we have that

$$RGR = aR \;\Rightarrow\; C_r(R)C_r(G)C_r(R) = C_r(aR) = a^r C_r(R);$$

$C_r(R)$ has rank one by [BR83, lem. 9]: if m_{ij} are the elements of $C_r(R)$, i.e. the minors of R, then $m_{ij}m_{hk} = m_{ik}m_{hj}$ for every i, j, h, k. So

$$a^r m_{ij} = \sum_{hk} m_{ik} g_{kh} m_{hj} = \sum_{hk} m_{ij} g_{kh} m_{hk} \;\Rightarrow\; a^r = \sum_{hk} g_{kh} m_{hk}$$

[2]This rather strange name is due to the existence of a list of four properties a true inverse satisfies; generalized inverses only satisfy some of these items: the ones here defined satisfy the first one or the first two.

that is exactly what we were searching for. □

Theorem 4.14. *For every matrix $R \in \mathcal{R}^{p\times q}$ there are $a \in \mathcal{R}$ and $G \in \mathcal{R}^{q\times p}$ such that $RGR = aR$.*

Proof. We have to use the field of fractions of $\mathcal{R}$: in this field R, of rank r, has the Smith form (4.3) $R = P_{*1}Q_{1*}$, P_{*1} full column rank, equal to r and Q_{1*} full row rank, still equal to r. Since we are on the field of fractions, then there is an element $a_d \in \mathcal{R}$, the common denominator, such that $a_d R = \tilde{P}\tilde{Q}$, where matrices $\tilde{P}$ and $\tilde{Q}$ are now in $\mathcal{R}$.

$\tilde{P}$ and $\tilde{Q}$ have still full column and row rank r, thus there are two elements a_u and a_v and matrices $U \in \mathcal{R}^{r\times p}$ and $V \in \mathcal{R}^{q\times r}$ such that $a_u I = U\tilde{P}$ and $a_v I = \tilde{Q}V$. From the first relation we have $a_d a_u R = a_u\tilde{P}\tilde{Q} = \tilde{P}U\tilde{P}\tilde{Q} = a_d\tilde{P}UR$ hence $a_u R = \tilde{P}UR$; now if we define $G = a_d VU$ and $a = a_u a_v$ then

$$RGR = a_d RVUR = \tilde{P}\tilde{Q}VUR = a_v\tilde{P}UR = a_u a_v R = aR.$$

□

4.2 Holomorphic functions

The ring $\mathcal{O}$ of complex holomorphic functions is itself an operator ring acting on the ring of complex holomorphic functions of exponential type,[3] that is to say, the complex holomorphic functions $h(s) \in \mathcal{O}$ that grow as exponentials as $|s| \to \infty$.

Since every other operator ring we will encounter is a subring of holomorphic functions, the way $\mathcal{O}$ operates on functions of exponential type is inherited by all operator rings in use throughout this thesis.

In particular it is important to show how holomorphic functions operate on **polynomial–exponential** functions so defined in the generic matrix case:

$$\mathcal{P}_e^{p\times q} \triangleq \left\{P(t)e^{\lambda t} :\ P(t) \in \mathbb{R}[t]^{p\times q},\ \lambda \in \mathbb{C}\right\}. \tag{4.8}$$

[3] More precisely: analytic functionals are isomorphic to holomorphic functions of exponential type [Tre67, ch. 22]; by reflexivity [Tre67, prop. 36.10] $\mathcal{O}$ is the dual of holomorphic function of exponential type. We can extend $\mathcal{O}$ to a ring of operators.

Proposition 4.15. *If $h(s) \in \mathcal{O}$ is a generic holomorphic function then there exists a unique operator*

$$\tilde{h} : t^n e^{\lambda t} \mapsto \tilde{h}\left(t^n e^{\lambda t}\right) \triangleq \sum_{i=0}^{n} \binom{n}{i} h^{(n-i)}(\lambda) t^i e^{\lambda t} \tag{4.9}$$

where $h^{(k)}(\lambda) = \frac{d^k}{ds^k} h(s)\Big|_{s=\lambda}$, that can be extended to a linear operator on $\mathcal{P}_e$. Moreover for any $h(s), k(s) \in \mathcal{O}$, $\tilde{h}\tilde{k} = \widetilde{hk}$.

Proof. The operator $\tilde{h}$ can be extended easily to linear combinations of elements $t^n e^{\lambda t}$ and then to non scalar cases, so to map the set of exponential–polynomial functions (4.8) into itself.

It remains to show that given $h(s), k(s) \in \mathcal{O}$ then $\tilde{h}(\tilde{k}p) = \widetilde{hk}(p)$ for every $p(t) \in \mathcal{P}_e$ or, without loss of generality, for every $p(t) = t^n e^{\lambda t}$:

$$\begin{aligned} \tilde{h}(\tilde{k}(t^n e^{\lambda t})) &= \sum_{i=0}^{n} \binom{n}{i} k^{(n-i)}(\lambda) \tilde{h}(t^i e^{\lambda t}) \\ &= \sum_{i=0}^{n} \binom{n}{i} k^{(n-i)}(\lambda) \sum_{j=0}^{i} \binom{i}{j} h^{(i-j)}(\lambda) t^j e^{\lambda t} \\ &= e^{\lambda t} \sum_{j=0}^{n} \sum_{i=j}^{n} \binom{n}{i} \binom{i}{j} k^{(n-i)}(\lambda) h^{(i-j)}(\lambda) t^j \end{aligned}$$

and, since it is relatively simple to show that $\binom{n}{i}\binom{i}{j} = \binom{n}{j}\binom{n-j}{i-j}$,

$$\begin{aligned} &= e^{\lambda t} \sum_{j=0}^{n} \binom{n}{j} t^j \sum_{i=j}^{n} \binom{n-j}{i-j} k^{(n-i)}(\lambda) h^{(i-j)}(\lambda) \\ &= e^{\lambda t} \sum_{j=0}^{n} \binom{n}{j} t^j \sum_{l=0}^{n-j} \binom{n-j}{l} k^{(n-j-l)}(\lambda) h^{(l)}(\lambda) p43 \\ &= e^{\lambda t} \sum_{j=0}^{n} \binom{n}{j} t^j \frac{d^{n-j}}{ds^{n-j}} \left(k(s)h(s)\right)\Big|_{s=\lambda} = \widetilde{hk}\left(t^n e^{\lambda t}\right). \end{aligned}$$

□

Remark 4.16. *Polynomial–exponential functions are useful as a class of* test functions *since they can detect zeros, and multiplicities, of holomorphic functions:*

$\lambda \in \mathbb{C}$ is a zero of multiplicity m of the holomorphic function $h(s) \in \mathcal{O}$ if and only if

$$\forall n \leq m,\ \tilde{h}\left(t^n e^{\lambda t}\right) = 0.$$

The ring $\mathcal{O}$ of holomorphic functions has two very important properties: it admits the Smith form [Hel43] hence it is a Bézout domain (see also [Rud87, p. 306]).

4.2.1 Paley–Wiener functions

We recall here the definition of the ring $\mathcal{A}$ of Paley–Wiener functions given in equation (2.26) and (2.27):

$$\mathcal{A} = \left\{f(s) \in \mathcal{O} :\ \exists A, B > 0,\ |f(s)| \leq Ae^{Bp(s)}\ \forall s \in \mathbb{C}\right\} \tag{4.10}$$

where

$$p(s) = \log\left(1 + |s|^2\right) + |\operatorname{Re} s|. \tag{4.11}$$

This definition can also be written in the following way:

$$\mathcal{A} = \left\{f \in \mathcal{O} :\ \exists A, B, C > 0,\ |f(s)| \leq A(1 + |s|^B)e^{C|\operatorname{Re}(s)|}\ \forall s \in \mathbb{C}\right\}. \tag{4.12}$$

This ring plays a fundamental role as we already noticed, but it lacks the nice properties (the ones formerly listed and others that will be soon introduced) that holomorphic functions possess.

First of all, Paley–Wiener functions without common zeros do not generate the whole ring, hence $\mathcal{A}$ is not a Bézout domain: as following example shows there exist two elements of $\mathcal{A}$ that do not have common zeros, i.e. they do not have common factors in $\mathcal{O}$, but do not satisfy a Bézout equation.

Example 4.17. Let us consider the following two functions in $\mathcal{A}$:

$$h_1(s) = \frac{e^s - e^{-s}}{2s} = \frac{i\sin(-is)}{s} \qquad h_2(s) = \frac{e^{as} - e^{-as}}{2} = i\sin(-ias).$$

We have shown that $h_1(s) \in \mathcal{A}$ in example 2.22; as a consequence of the same proof, even $h_2(s) \in \mathcal{A}$.

The zeros of $h_1(s)$ are exactly the points of the set $\mathcal{Z}_1 \triangleq \{ik\pi : \ 0 \neq k \in \mathbb{Z}\}$ and the zeros of $h_2(s)$ are $\mathcal{Z}_2 \triangleq \{in\pi/a : \ n \in \mathbb{Z}\}$.

When $a = p/q \in \mathbb{Q}$, $h_1(s)$ and $h_2(s)$ have common zeros:

$$\mathcal{Z} \triangleq \mathcal{Z}_1 \cap \mathcal{Z}_2 = \{ikp\pi : 0 \neq k \in \mathbb{Z}\} \neq \emptyset$$

and the factors $s - z$ with $z \in \mathcal{Z}$ are common to $h_1(s)$ and $h_2(s)$.

If $a \notin \mathbb{Q}$ then $\mathcal{Z} = \emptyset$: $h_1(s)$ and $h_2(s)$ have no common zeros. In this case we want to see if they satisfy a Bézout equation, i.e.

$$\exists x_1(s), x_2(s) \in \mathcal{A} \text{ such that } x_1(s)h_1(s) + x_2(s)h_2(s) = 1. \tag{4.13}$$

Let us recall here the definition of Liouville numbers [Niv56]: $a \notin \mathbb{Q}$ is a **Liouville number** if for every $C \in \mathbb{N}$ there are infinitely many fractions p/q such that

$$\left| a - \frac{p}{q} \right| \leq q^{-1-C}. \tag{4.14}$$

If $C = 0$ then this is true for every $a \in \mathbb{R}$ ($C = 1$ at least for irrational numbers) and is the basis of the standard approximation of real numbers by continued fractions; Liouville numbers, in this sense, are trascendental numbers that are *well approximated* by rational numbers up to any 'order' $C + 1$.

If we suppose that a is a Liouville number and equation (4.13) is true, then using a basic trigonometry result and (4.14):

$$|\sin aq\pi| = |\sin(aq\pi - p\pi)| \leq |aq\pi - p\pi| \leq q^{-C}\pi;$$

when $s = iq\pi \in \mathcal{Z}_1$, zero of $h_1(s)$, substituting it in (4.13) we get

$$ix_2(-iq\pi) \sin aq\pi = 1 \ \Rightarrow \ 1 = |x_2(-iq\pi)| \cdot |\sin aq\pi| \leq |x_2(-iq\pi)|\pi q^{-C} \tag{4.15}$$

But $x_2(s) \in \mathcal{A}$: by definition (4.12) there exist constants $A, B > 0$ such that

$$|x_2(-iq\pi)| \leq A(1 + q\pi)^B$$

and from (4.15)

$$1 \leq |x_2(-iq\pi)|\pi q^{-C} \leq A\pi(1+q\pi)^B q^{-C} \tag{4.16}$$

which is clearly impossible since it should be valid for infinitely many $q \in \mathbb{N}$ while B is fixed and C may be arbitrarily chosen. In other words, if we take $C > B$ there are infinitely many fractions p/q satisfying relation (4.14): we can let q go to infinity, so the right part of equation (4.16) tends to zero. ♣

If the elements $h_i(s) \in \mathcal{A}$ satisfy a Bézout equation in $\mathcal{A}$, then

$$1 = \sum h_i(s)k_i(s) \leq \sum |h_i(s)||k_i(s)| \leq \sum |h_i(s)|Ae^{Bp(s)}$$
$$\Rightarrow \sum |h_i(s)| \geq \frac{1}{A}e^{-Bp(s)}$$

This condition is not only necessary, as Hörmander proved in [Hör67]:

Theorem 4.18. *The set of functions $\{h_1(s), \dots, h_n(s)\} \subset \mathcal{A}$ is a set of generators of $\mathcal{A}$ if and only if there are constants $\epsilon, B, C > 0$ such that*

$$\sum_{i=1}^{n} |f_i(s)| \geq \epsilon(1+|s|^{-C})e^{-B|\operatorname{Re}(s)|} \; \forall s \in \mathbb{C}. \tag{4.17}$$

Condition (4.17) is apparently simple, but it is really hard to verify in general. There are some other properties of holomorphic functions strictly related to the solution of a Bézout equation, that will be useful in the following sections: in [Ehr60, p. 523] we find

Definition 4.19. *A function $h(s) \in \mathcal{A}$ is **slowly decreasing** if there is a real constant $\epsilon > 0$ such that*

$$\forall x \in \mathbb{R}, \; \exists y \in \mathbb{R} : \; |x-y| \leq \epsilon \log(1+|x|) \text{ and } |h(y)| \geq (\epsilon + |y|)^{-\epsilon}.$$

The same paper of Ehrenpreis contains the following

Theorem 4.20. *Let $\alpha \in \mathcal{E}'$ be a distribution with compact support. The following conditions are equivalent:*

- *$\hat{\alpha}(s)$ is slowly decreasing*
- *$\tilde{\alpha} : \mathcal{E} \to \mathcal{E}$ is a surjective operator*

- *the ideal $\mathcal{A}\hat{\alpha}(s)$ is closed.*

These definitions have been extended in various ways (see for example [Ehr70, ch. XI.1] and [Str83]); in [BT79] there is an important definition:

Definition 4.21. *The **local ideal** generated by $\{h_1(s), \dots, h_n(s)\} \subset \mathcal{A}$ is the set*

$$(h_1, \dots, h_n)^l_{\mathcal{A}} = \{f \in \mathcal{A} : \forall s \in \mathbb{C}\, \exists \text{open } U \ni s, k_i(s) \in \mathcal{O}(U),\ f = \textstyle\sum h_i k_i \text{ in } U\} \tag{4.18}$$

where $\mathcal{O}(U)$ is the set of holomorphic functions defined on the open set $U \subseteq \mathbb{C}$.

*The functions $\{h_i(s)\}$ are **jointly slowly decreasing** if*

$$(h_1, \dots, h_n)^l_{\mathcal{A}} = (h_1, \dots, h_n)_{\mathcal{A}}$$

i.e. the ideal they generate on $\mathcal{A}$ coincides with the local ideal.

Proposition 4.22. *Suppose that the functions $\{h_1, \dots, h_n\} \subset \mathcal{A}$ have no common zeros; they satisfy the Bézout equation if and only if they are jointly slowly decreasing.*

Proof. Since $(h_1, \dots, h_n)^l_{\mathcal{A}} \supseteq (h_1, \dots, h_n)_{\mathcal{A}}$, if the functions satisfy a Bézout equation in $\mathcal{A}$, then the ideal they generate coincides with $\mathcal{A}$ and so must do the local ideal.

If the functions have no common zeros, then they satisfy a Bézout equation on $\mathcal{O}$; this implies that $1 \in (h_1, \dots, h_n)^l_{\mathcal{A}}$ by its definition. If additionally the functions are jointly slowly decreasing, then also $1 \in (h_1, \dots, h_n)_{\mathcal{A}}$. □

Remark 4.23. *There exists a nice characterization of the local ideal generated by $\{h_i(s)\}$ in $\mathcal{A}$: it is the set of Paley–Wiener functions whose zeros are the set of common zeros of $\{h_i\}$ with greater or equal multiplicities.*

There are three important theorems, proved by Malgrange and by Schwartz, that regard matrices in $\mathcal{A}$.

The first one is similar to what is now called *Fundamental Principle* that Ehrenpreis and Palamodov discovered independently in [Ehr70] and [Pal70] (for generic linear partial differential equations) and can be stated for differential systems in a behavioral framework as follows: if $R\left(\frac{d}{dt}\right) \in \mathbb{R}\left[\frac{d}{dt}\right]^{p\times q}$ is a matrix of polynomial differential operators, then

$$\mathcal{B} = \ker_{\mathcal{E}} R\left(\tfrac{d}{dt}\right) = \overline{\ker_{\mathcal{P}_e} R\left(\tfrac{d}{dt}\right)} \tag{4.19}$$

where the closure of last term is with respect to the topology of $\mathcal{E}$.

The theorem proved earlier by Malgrange in [Mal56, p.318] needs a stronger hypothesis:

Theorem 4.24. *If $R(s) \in \mathcal{A}^{p\times q}$ is a full row rank matrix of Paley–Wiener functions, then*

$$\mathcal{B} = \ker_{\mathcal{E}} R(s) = \overline{\ker_{\mathcal{P}_e} R(s)}.$$

The second result generalizes the fact that for any two holomorphic functions $a(s), b(s) \in \mathcal{O}$, $a(s)/b(s) \in \mathcal{O}$ if and only if the zeros of $b(s)$ are also zeros of $a(s)$ with greater multiplicity.

Remembering that there is a relation between zeros of holomorphic functions and polynomial–exponential functions, as underlined in remark 4.16, the following theorem [Mal56, p. 282] should be quite intuitive:

Theorem 4.25. *Given two matrices of Paley–Wiener functions $R_1(s) \in \mathcal{A}^{p_1\times q}$ and $R_2(s) \in \mathcal{A}^{p_2\times q}$ then*

$$\exists X(s) \in \mathcal{O}^{p_2\times p_1} \text{ such that } X(s)R_1(s) = R_2(s) \;\Leftrightarrow\; \ker_{\mathcal{P}_e} R_1(s) \subseteq \ker_{\mathcal{P}_e} R_2(s).$$

Finally we recall the following theorem, a classical result of Schwartz [Sch47] (see also [BS93, p. 38]):

Theorem 4.26 (Spectral analysis theorem). *Let $R(s) \in \mathcal{A}^{p\times 1}$ be a column vector; if its elements have no common zeros, then* $\ker_{\mathcal{E}} R(s) = \{0\}$.

4.2.2 Exponential polynomials

Polynomials are undoubtedly the most known subring of holomorphic functions also from an *operatorial* point of view, because of their immediate meaning as polynomial differential operators (see definition 3.4).

The consistency we pointed out in remark 3.5 extends the definition of holomorphic operators since also on polynomial–exponential functions $\mathcal{P}_e$ the function s is a differential operator, as we can clearly see from (4.9).

The ring of polynomials has many well-known algebraic properties: basically it is a principal ideal domain, hence it is a Bézout domain; furthermore it is an elementary divisor domain. Besides these one there are, perhaps less known, *operatorial* properties that are surely important in this context.

The most important one is that polynomials, as Ehrenpreis proved in [Ehr54, p. 898], are surjective operators on $\mathcal{E}$; this fact still holds for polynomials in more than one variable (i.e. partial differential operators) and extends to systems of (partial) differential equations.

As regards delay–differential equations, in section 3.1.4 delay–differential polynomial operators in m delays $\tau_1, \dots, \tau_m$ were defined as polynomials in $m+1$ variables, the first one corresponding, roughly speaking, to derivation or to the function s and the other ones to the m delay operators $\sigma_{\tau_1}, \dots, \sigma_{\tau_m}$ that, as shown in example 2.20, correspond to the holomorphic functions $e^{-s\tau_1}, \dots, e^{-s\tau_m}$.

The question is: if we have m shift operators, the following rings (the first one is the so called ring of **exponential polynomials** not to be confused with polynomial–exponential functions $\mathcal{P}_e$)

$$\mathbb{R}\left[s, e^{-s\tau_1}, \dots, e^{-s\tau_m}\right],\ \mathbb{R}\left[\tfrac{d}{dt}, \sigma_{\tau_1}, \dots, \sigma_{\tau_m}\right] \text{ and } \mathbb{R}\left[s, z_1, \dots, z_m\right] \tag{4.20}$$

are still isomorphic?

First we note that the first two rings in (4.20) are naturally isomorphic since their generators (as a monoid algebra over $\mathbb{R}$) are the Laplace transform ($\hat{\alpha}(s) \in \mathcal{A}$) and, respectively, the *operatorial form* ($\tilde{\alpha}$) of the same set of distributions

$$\delta^{(1)},\ \delta_{\tau_1},\ \dots,\ \delta_{\tau_m}$$

viewed as a convolutional subalgebra of $\mathcal{E}'$: the monomials $\left(\frac{d}{dt}\right)^{n_0} \sigma_{\tau_1}^{n_1} \cdots \sigma_{\tau_m}^{n_m}$ and $s^{n_0} e^{-n_1\tau_1 s} \cdots e^{-n_m\tau_m s}$, which if we set $\tau = \sum n_i\tau_i$ are respectively $\left(\frac{d}{dt}\right)^{n_0} \sigma_\tau$ and $s^{n_0} e^{-\tau s}$, are in bijective relation with the distribution

$$\delta^{(n_0)} \star \delta_{n_1\tau_1} \star \cdots \star \delta_{n_m\tau_m} = \delta^{(n_0)} \star \delta_\tau = \delta_\tau^{(n_0)}.$$

Last passage is true since by (2.30), (2.23) and (2.24)

$$\forall f,\ \left\langle \delta^{(n_0)} \star \delta_\tau, f \right\rangle = \left\langle \delta^{(n_0)} \star \check{\delta}_{-\tau}, f \right\rangle = \left\langle \delta^{(n_0)}, \tilde{\delta}_{-\tau} f \right\rangle = \left\langle \delta^{(n_0)}, \sigma_{-\tau} f \right\rangle$$
$$= (-1)^{n_0} \left\langle \delta, \sigma_{-\tau} f^{(n_0)} \right\rangle = (-1)^{n_0} \left\langle \delta_\tau, f^{(n_0)} \right\rangle = \left\langle \delta_\tau^{(n_0)}, f \right\rangle.$$

Next proposition shows the relation between the ring of polynomials and the other ones; it justifies *a posteriori* the definitions of section 3.1.4 in case of incommensurate delays.

Proposition 4.27. *The rings listed in (4.20) are isomorphic if and only if the delays $\tau_1, \dots, \tau_m$ are independent over $\mathbb{Q}$, i.e there is no linear combination $\sum q_i \tau_i = 0$ with $q_i \in \mathbb{Q}$, at least one $q_i \neq 0$.*

Proof. The last ring in (4.20) may be projected onto the first one substituting $z_i \to e^{-\tau_i s}$, but this operation is not necessarily injective. Let us consider the projection of a generic monomial onto the convolution algebra formerly described:

$$\Theta(s^{n_0} z_1^{n_1} \cdots z_m^{n_m}) = \delta_\tau^{(n_0)},\ \tau = \sum_{i=1}^{m} n_i \tau_i$$

so that it is clear that another monomial $s^{n_0'} z_1^{n_1'} \cdots z_m^{n_m'}$ will be projected onto the same element if and only if $n_0 = n_0'$ and $\tau = \tau' = \sum n_i' \tau_i$; last relation may be written as

$$0 = \tau - \tau' = \sum_{i=1}^{m} (n_i - n_i')\tau_i.$$

Now the result follows since a $\mathbb{Z}$-linear independency is equivalent to $\mathbb{Q}$-linear independency. □

We note that the same result holds for Laurent exponential polynomials, i.e. also the following rings are isomorphic when the hypothesis of proposition 4.27 is satisfied:

$$\mathbb{R}\left[s, e^{-s\boldsymbol{\tau}}, e^{s\boldsymbol{\tau}}\right],\ \mathbb{R}\left[\tfrac{d}{dt}, \sigma_{\boldsymbol{\tau}}, \sigma_{-\boldsymbol{\tau}}\right] \text{ and } \mathbb{R}\left[s, \boldsymbol{z}, \boldsymbol{z}^{-1}\right]$$

written with compact notation (i.e. $e^{-s\boldsymbol{\tau}} = e^{-s\tau_1}, \dots, e^{-s\tau_m}$ and so on).

Remark 4.28. *Because of this isomorphism, we can use indifferently one of the three rings in (4.20), but only from a formal point of view.*

It is trivial that e.g. zeros of a polynomial in $m+1$ variables and of the corresponding exponential polynomial are generally different: the simplest example with one delay is $p(s,z)=z$ which is zero along the whole s axis whereas $p(s,e^{-s})=e^{-s}$ has no zeros. Therefore also factorization properties differ (see [EP97]).

Like polynomials, also exponential polynomials are surjective operators on $\mathcal{E}$ [Ehr55]:

Theorem 4.29. *Let $f(s)=\sum f_i(s)e^{\alpha_i s}$ where $f_i(s)\in\mathbb{C}[s]$ and $\alpha_i\in\mathbb{C}$; if $f(s)=\hat{\phi}(s)$, Laplace transform of $\phi\in\mathcal{E}'$, then $\tilde{\phi}:\mathcal{E}\to\mathcal{E}$ is injective.*

A fundamental property of polynomials is that $p(s)/q(s)$ is an holomorphic function if and only if the fraction is a polynomial.

This equivalence does not hold for exponential polynomials: in general holomorphic fractions of exponential polynomials are not exponential polynomials. Indeed there is a weaker[4] but fundamental property [BD74]:

Theorem 4.30. *Given the set of* complex *exponential polynomials*

$$E=\left\{f(s)=\sum_1^n f_i(s)e^{\alpha_i s}:\ f_i(s)\in\mathbb{C}[s],\ \alpha_i\in\mathbb{C}\right\}$$

then for any two $f(s),g(s)\in E$ we have that

$$\frac{f(s)}{g(s)}\in\mathcal{O}\ \Rightarrow\ \frac{f(s)}{g(s)}=\frac{h(s)}{p(s)},\ h(s)\in E,\ p(s)\in\mathbb{C}[s].$$

Moreover let $d_h(s)$ be the greatest common divisor of the polynomials $h_i(s)$, coefficients of the exponentials in

$$h(s)=\sum_1^k h_i(s)e^{\gamma_i s}.$$

Then, if $h(s)/p(s)$ is reduced such that $d_h(s)$ and $p(s)$ have no common factors, $p(s)$ divides $d_g(s)$.

This theorem suggests the definition of another important ring of operators.

[4] The original theorem applies to n–variables exponential polynomials; the theorem that is stated here is a simpler but rich enough 'corollary'.

4.2.3 The ring $\mathcal{H}_m$

Theorem 4.30 becomes in our context (delays and coefficients in $\mathbb{R}$):

Proposition 4.31. *If $f(s), g(s) \in \mathbb{R}\left[s, e^{-\boldsymbol{\tau} s}\right]$ then*

$$\frac{f(s)}{g(s)} \in \mathcal{O} \;\Rightarrow\; \frac{f(s)}{g(s)} = \frac{h(s)}{p(s)e^{-\tau s}}$$

where $h(s) \in \mathbb{R}\left[s, e^{-\boldsymbol{\tau} s}\right]$, $p(s) \in \mathbb{R}[s]$, and τ is an $\mathbb{N}$-linear combination of the delays: $\tau = \sum n_i \tau_i$, $n_i \in \mathbb{N}$.

Proof. (Sketch) Theorem 4.30 states that, following its notation, if

$$f(s) = \sum f_i(s)e^{\alpha_i s}, \; g(s) = \sum g_j(s)e^{\beta_j s} \text{ and } h(s) = \sum h_k(s)e^{\gamma_k s},$$

with α_i and β_j $\mathbb{N}$-linear combination of the delays, then $p(s)f(s) = g(s)h(s)$. Since $p(s)$ is a polynomial $p(s)f(s)$ must have the exponentials of $f(s)$ while $g(s)h(s)$ has exponentials like $e^{-(\beta_j+\gamma_k)s}$.

The exponential polynomials are equal if every $\beta_j + \gamma_k$ is equal to some α_i (and the polynomial coefficients coincide); therefore $\gamma_k = \alpha_i - \beta_j$ is a $\mathbb{Z}$-linear combination of the delays.

Finally we can let the negative combinations vanish dividing by a suitable monomial $e^{-\tau s}$. □

Definition 4.32. *The set $\mathcal{H}_m$ contains the holomorphic fractions of exponential polynomials with m delays:*

$$\mathcal{H}_m \triangleq \mathcal{O} \cap \left\{ \frac{f(s)}{g(s)} : \; f(s), g(s) \in \mathbb{R}\left[s, e^{-\boldsymbol{\tau} s}\right] \right\}.$$

Proposition 4.33. *The set $\mathcal{H}_m$ is a ring and whenever*

$$h(s), k(s) \in \mathcal{H}_m, \; \frac{h(s)}{k(s)} \in \mathcal{O} \;\Rightarrow\; \frac{h(s)}{k(s)} \in \mathcal{H}_m.$$

Proof. Trivially $\mathcal{H}_m$ is a ring; then by proposition 4.31 every fraction of elements in $\mathcal{H}_m$ is again a fraction of exponential polynomials, hence if it is holomorphic it belongs to $\mathcal{H}_m$. □

Remark 4.34. *The ring $\mathcal{H}_m$ may also be defined in this way:*

$$\mathcal{H}_m = \mathcal{O} \cap \mathbb{R}(s) \left[e^{-\boldsymbol{\tau} s}, e^{\boldsymbol{\tau} s}\right]$$

i.e. as the holomorphic elements of the ring of Laurent polynomials in $e^{-\boldsymbol{\tau} s}$ whose coefficients are rational functions.

The ring $\mathcal{H}_m$ extends the ring of exponential polynomials without becoming too large for our purposes:

Theorem 4.35. *The ring $\mathcal{H}_m$ is a subring of $\mathcal{A}$, and its elements operate on $\mathcal{E}$ in the following way: let $h(s) \in \mathcal{H}_m$ and let $f(s), g(s)$ be two exponential polynomials such that $h(s) = f(s)/g(s)$; so*

$$\forall w(t) \in \mathcal{E},\ h(s)w = v \in \mathcal{E} \text{ if and only if } f(s)w = g(s)v. \tag{4.21}$$

Proof. First we show that the way elements of $\mathcal{H}_m$ operate on $\mathcal{E}$ is well defined.

From theorem 4.29 we know that $g(s)$ is a surjective operator:

$$\forall w(t) \in \mathcal{E}\ \exists u(t) \in \mathcal{E}:\ g(s)u = w;$$

$u(t)$ is defined up to sum with elements in $\ker_{\mathcal{E}} g(s)$, i.e.

$$g(s)u = g(s)\tilde{u} = w \ \Rightarrow\ u(t) - \tilde{u}(t) = n(t) \in \ker_{\mathcal{E}} g(s). \tag{4.22}$$

Combining theorems 4.24 and 4.25 we see that

$$\ker_{\mathcal{E}} g(s) = \overline{\ker_{\mathcal{P}_e} g(s)} \subseteq \overline{\ker_{\mathcal{P}_e} f(s)} = \ker_{\mathcal{E}} f(s)$$

thus for every $n(t) \in \ker_{\mathcal{E}} g(s)$ also $f(s)n = 0$. Therefore for any two $u(t)$ and $\tilde{u}(t)$ as in (4.22) $f(s)u = f(s)(\tilde{u} + n) = f(s)\tilde{u} = v(t)$ so $v(t)$ is uniquely determined for every $w(t)$.

The fact that $\mathcal{H}_m \subseteq \mathcal{A}$ may be proved as in example 2.22: we give only a sketch of the proof.

Suppose that $h(s) = f(s)/g(s)$ with $f(s)$ a Laurent exponential polynomial, thus $f(s) \in \mathcal{A}$, and $g(s)$ a polynomial whose zeros are zeros of $f(s)$. If $\mathcal{Z}_i$ are non intersecating neighborhoods of the zeros of $g(s)$, then outside $\cup \mathcal{Z}_i$, $|g(s)| \geq 1/G$

for some constant $G > 0$, therefore

$$\forall s \neq \cup \mathcal{Z}_i,\ |h(s)| = \frac{|f(s)|}{|g(s)|} \leq G|f(s)|$$

thus $h(s)$ also satisfies the definition of $\mathcal{A}$.

When $s \in \mathcal{Z}_i$, a compact set, $h(s)$ is a continuous function, hence bounded; *a fortiori*

$$\forall s \in \mathcal{Z}_i,\ |h(s)| \leq A_i e^{B_i p(s)} \text{ for some constants } A_i, B_i > 0$$

as in (4.10) proving in conclusion that $h(s) \in \mathcal{A}$. □

Eventually, the most important property of $\mathcal{H}_1$ (unfortunately it is false when there are more incommensurable delays) is the existence of the Smith form. In [GL97a], a fundamental paper for the behavioral approach to delay–differential systems, there are some noteworhty properties of $\mathcal{H}_1$:

Theorem 4.36. *The ring $\mathcal{H}_1$ is not a unique factorization domain and not a Noetherian ring. But it is a greatest common divisor domain, a Bézout domain and an elementary divisor domain.*

We will come back to the fruitful consequences of this theorem in chapter 6, showing how delay–differential systems with commensurable delays may be treated very satisfactorily with behavioral techniques.

Chapter 5

Duality

Steenrod called 'abstract nonsense' the *arrow-theoretic* constructions he and other mathematicians like Cartan and MacLane developed mainly in the forties and fifties; these ideas had been used in the study of topological spaces, partial differential equations, algebraic geometry.

This chapter will only touch on these topics in order to prove that there is a deep relation between the two different approaches to system theory that were introduced in chapter 3: we try to link the 'trajectory' side of behaviors and the 'algebraic' side of Fliess' approach showing a kind of duality between them.

5.1 Behaviors are homomorphisms

Behaviors have their own place inside the algebraic constructions suggested by Fliess; however a neat treatment of this subject should involve homological algebraic tools, that would need more than one chapter only to be introduced.

Therefore we will only show some important results, without getting too deep into mathematic details (appendix A contains more precise definitions), mainly to be able to link the module theory and the behavioral world. A good reference, for linear time–invariant systems, is [Obe90]; time–varying linear system has also been studied as is shown in [OF98].

There are also some works on delay–differential systems or in general on convolutional systems (see e.g. [Str83]) but they are not so satisfactory from an engineering point of view since they do not deal with typical problems of systems theory, but with different functional analytical questions.

Remark 5.1. *If $\mathcal{R}$ is a ring and M and N are left modules over R then we denote the set of $\mathcal{R}$–homomorphisms (i.e. the $\mathcal{R}$-linear maps) of M into N by $\mathrm{Hom}_{\mathcal{R}}(M, N)$.*

It can be proved easily that if $\mathcal{S}$ is another ring, then if M [or N] is also a right $\mathcal{S}$–module, the set $\mathrm{Hom}_{\mathcal{R}}(M, N)$ is a left [or right] $\mathcal{S}$–module, the multiplication by an element $s \in \mathcal{S}$ being defined as

$$(sf)(x) = f(xs) \text{ [or } (fs)(x) = f(x)s\text{]}. \tag{5.1}$$

We will suppose that the rings we consider are rings of operators on $\mathcal{E}$ ($\mathcal{R}$, for instance, may be equal to $\mathcal{A}$, $\mathcal{H}_m$ and so on) or, in other words, that $\mathcal{E}$ is a left $\mathcal{R}$–module.

Proposition 5.2. *Given a matrix $R \in \mathcal{R}^{p\times q}$, it can be both a presentation matrix of the linear system $\mathcal{M} \triangleq \mathrm{coker}_{\mathcal{R}} \circ R$ and a kernel representation of the behavior $\mathcal{B} = \ker_{\mathcal{E}} R$. We have $\mathcal{B} \cong \mathrm{Hom}_{\mathcal{R}}(\mathcal{M}, \mathcal{E})$.*

Proof. This is more an explanation rather than a proof; to be more concrete we suppose that $\mathcal{R} = \mathcal{A}$. Since $\mathcal{M}$ is not only a left but also a right $\mathcal{A}$–module, $\mathcal{B}$ gets the structure of left $\mathcal{A}$–module by remark 5.1.

First let us show that every trajectory in $\mathcal{B}$ is an homomorphism: let $w(t) \in \mathcal{B}$; the map

$$\Phi_w : \mathcal{M} \to \mathcal{E},\ \mathsf{m} = m(s) + \mathcal{A}^p R(s) \mapsto \Phi_w(\mathsf{m}) \triangleq m(s)w \in \mathcal{E} \tag{5.2}$$

is a well defined linear map: indeed any other $n(s)$ in the same equivalence class of $m(s)$ differs from it up to elements in the image of $R(s)$, thus

$$n(s)w = (m(s) + a(s)R(s))w = m(s)w + a(s)R(s)w = m(s)w$$

since $w(t)$ is in the kernel of $R(s)$.

Conversely every linear map Φ is given by (5.2): let $\{e_i\}$ be the standard basis of $\mathcal{A}^q$, the set of vectors with i-th component 1 and zero elsewhere and $\{\mathsf{e}_i\}$ the corresponding generators of $\mathcal{M}$. In this case

$$w_i(t) \triangleq \Phi(\mathsf{e}_i) \ \Rightarrow\ \Phi = \Phi_w \text{ with } w(t) \triangleq [w_1(t) \ \cdots \ w_q(t)]^\top$$

by linearity.

We note that the left $\mathcal{A}$–module structure of $\mathcal{B}$ is exactly the one defined by (5.1):

$$\forall a(s) \in \mathcal{A},\ a(s)\Phi_w(\mathsf{m}) = \Phi_w(\mathsf{m}a(s)) = m(s)a(s)w = \Phi_{a(s)w}(\mathsf{m}).$$

□

Next theorem shows that, at least in the particular case of linear differential systems, this duality between $\mathcal{B}$ and $\mathcal{M}$ has a deep meaning.

Theorem 5.3. *Given a matrix* $R\left(\frac{d}{dt}\right) \in \mathbb{R}\left[\frac{d}{dt}\right]^{p\times q}$*, the linear system* $\mathcal{M} = \operatorname{coker}_{\mathbb{R}[\frac{d}{dt}]} \circ R\left(\frac{d}{dt}\right)$ *is controllable if and only if the corresponding behavior* $\mathcal{B} = \ker_{\mathcal{E}} R\left(\frac{d}{dt}\right)$ *is controllable.*

Proof. About the notation: an element $\mathsf{m} \in \mathcal{M}$ is an equivalence class $\mathsf{m} = m\left(\frac{d}{dt}\right) + \mathbb{R}\left[\frac{d}{dt}\right]^p R\left(\frac{d}{dt}\right)$ and Φ operates as in (5.2): $\Phi_w(\mathsf{m}) = m\left(\frac{d}{dt}\right) w$ for every $w(t) \in \mathcal{B} \subseteq \mathcal{E}^q$.

Let us point out that the trajectory $f \triangleq \Phi_w(\mathsf{m}) = m\left(\frac{d}{dt}\right) w \in \mathcal{E}$ depends only locally on $w(t)$, i.e. if $w_1(t)$ and $w_2(t)$ coincide on an interval I then also $\Phi_{w_1}(\mathsf{m})$ and $\Phi_{w_2}(\mathsf{m})$ coincide on I; conversely if $\Phi_{w_1}(\mathsf{m})$ coincides with $\Phi_{w_2}(\mathsf{m})$ on an interval I for every $\mathsf{m} \in \mathcal{M}$, then also $w_1(t)$ and $w_2(t)$ coincide on I. Finally, by (5.2), $\Phi_{\sigma_\tau w}(\mathsf{m}) = \sigma_\tau \Phi_w(\mathsf{m})$.

Let us suppose that $\mathcal{M}$ is controllable. Then, by definition given in section 3.2.2, $\mathcal{M}$ is free hence it has a basis, say $\{\mathsf{m}_1, \ldots, \mathsf{m}_k\}$. Every homomorphism Φ_w with $w(t) \in \mathcal{B}$ is defined by linearity on the whole $\mathcal{M}$ once we know its values $\Phi_w(\mathsf{m}_i)$.

Given any two $w_1(t), w_2(t) \in \mathcal{B}$ then there is always a smooth way to connect the past $t \leq 0$ of $\Phi_{w_1}(\mathsf{m}_i)$ with the future $t \geq \tau \geq 0$ of $\Phi_{\sigma_\tau w_2}(\mathsf{m}_i) = \sigma_\tau \Phi_{w_2}(\mathsf{m}_i)$ (see e.g. proposition 3.14). Let us call $f_i(t)$ this trajectory. Then, if we let $\Phi(\mathsf{m}_i) \triangleq f_i$ we have defined a homomorphism of $\mathcal{M}$ into $\mathcal{E}$, hence a trajectory $w(t) \in \mathcal{B}$ that, by definition and following what we said before, coincides with $w_1(t)$ as $t \leq 0$ and with $\sigma_\tau w_2(t)$ as $t \geq \tau$.

Suppose now that the behavior $\mathcal{B}$ is controllable and that $\mathcal{M}$ is not free. Then since $\mathcal{M} = \mathcal{M}_f \oplus \mathcal{M}_t$ is a direct sum of a free and a torsion submodule [Lan93, p. 147], $\mathcal{M}_t \neq 0$.

So, if $0 \neq \mathsf{m} \in \mathcal{M}_t$ there is a $p\left(\frac{d}{dt}\right) \in \mathbb{R}\left[\frac{d}{dt}\right]$ such that $p\left(\frac{d}{dt}\right)\mathsf{m} = 0$, therefore $\Phi_w(p\left(\frac{d}{dt}\right)\mathsf{m}) = 0$ for every $w(t) \in \mathcal{B}$.

This means that if $f \triangleq m\left(\frac{d}{dt}\right) w$, $p\left(\frac{d}{dt}\right) f = p\left(\frac{d}{dt}\right) m\left(\frac{d}{dt}\right) w = 0$, i.e. $f(t)$ satisfies an homogeneous scalar differential equation therefore it is the zero trajectory if and only if it vanishes for some $t \in \mathbb{R}$.

Let $w(t) \in \mathcal{B}$ be such that $f = m\left(\frac{d}{dt}\right) w$ is not zero in some interval $I = [t_1, t_2]$ of the positive time-axis and $\bar{w}(t) \in \mathcal{B}$ a trajectory equal to $w(t)$ for $t \leq 0$ and zero for $t \geq \tau \geq 0$ for some $\tau \leq t_1$ that exists since $\mathcal{B}$ is controllable.

Then $\bar{f} = m\left(\frac{d}{dt}\right) \bar{w}$ is the zero trajectory since it vanishes for $t \geq \tau$; $f(t) = \bar{f}(t)$ for $t \leq 0$ hence also $f(t)$ is zero everywhere, in contradiction with the definition of $w(t)$. $\square$

We will investigate further the relation between controllability of behaviors and various types of algebraic controllabilities in chapter 7.

5.2 Algebraic duality

The relation between the module of relations $\mathcal{M}$ and the behavior $\mathcal{B}$, kernel of the representation matrix, can be proposed using exact sequences of modules:

$$
\begin{array}{ccc}
\mathcal{R}^p \xrightarrow{\circ R} \mathcal{R}^q \xrightarrow{\phi} \mathcal{M} \longrightarrow 0 & & \\
 & & \Big\downarrow \mathrm{Hom}_{\mathcal{R}}(\cdot, \mathcal{E}) \\
\mathcal{E}^p \xleftarrow{R\circ} \mathcal{E}^q \xleftarrow{i} \mathcal{B} \longleftarrow 0 & &
\end{array}
$$

Actually $\mathrm{Hom}_{\mathcal{R}}(\cdot, \mathcal{E})$ is a contravariant functor (i.e. it 'reverses arrows') that maps the category of left and right $\mathcal{R}$–modules, where the operation is the ring product of $\mathcal{R}$, into the category of left modules over $\mathcal{R}$: this time the operation consists in applying operators (elements of $\mathcal{R}$) to functions (in $\mathcal{E}$). So, since $\mathrm{Hom}_{\mathcal{R}}(\mathcal{R}^k, \mathcal{E}) \cong \mathcal{E}^k$, this simple diagram yealds, in a more abstract way, the result of proposition 5.2, i.e. $\mathcal{B} = \ker_{\mathcal{E}} R \cong \mathrm{Hom}_{\mathcal{R}}(\mathcal{M}, \mathcal{E})$.

Moreover the functor $\mathrm{Hom}_{\mathcal{R}}(\cdot, \mathcal{E})$ maps morphisms of one category into morphisms of the other one: e.g. the projection ϕ becomes an injection i for functions while matrix operators simply operate 'on the other side'.

When $\mathcal{R} = \mathbb{R}\left[\frac{d}{dt}\right]$ (and many other rings indeed, see [Obe90]) the functor becomes a so called *categorical duality*, in other words there is a perfect duality between the algebraic modules and the functions modules: for example $\ker_{\mathcal{E}} R\left(\frac{d}{dt}\right) = \mathrm{im}_{\mathcal{E}} M\left(\frac{d}{dt}\right)$ if and only if $\mathrm{im}_{\mathbb{R}[\frac{d}{dt}]} \circ R\left(\frac{d}{dt}\right) = \ker_{\mathbb{R}[\frac{d}{dt}]} \circ M\left(\frac{d}{dt}\right)$.

This duality implies the Fundamental Principle for non homogeneous equations (see theorem 3.10): actually, given a behavior in image representation $\mathcal{B} = \operatorname{im}_{\mathcal{E}} M\left(\frac{d}{dt}\right)$, we can find its kernel representation simply constructing (generators of) the kernel $\ker_{\mathbb{R}[\frac{d}{dt}]} \circ M\left(\frac{d}{dt}\right)$ over the polynomial ring, fact that is allowed by Noetherianity of $\mathbb{R}\left[\frac{d}{dt}\right]$.

Dealing with more general linear systems, with other operator rings, makes it necessary to take topology into consideration, in order to deal with continuous operators.

5.3 Topological duality

The first fundamental fact involving duality is theorem 4.20, (see [BS93, p. 35] for a sketch of the proof). The following proposition is an immediate consequence that generalizes theorem 4.35:

Proposition 5.4. *Let $a(s), b(s) \in \mathcal{A}$ such that the fraction $c(s) = a(s)/b(s)$ is holomorphic. Then $c(s)$ is a Paley–Wiener function if and only if $\mathcal{A}b(s)$ is closed.*

Proof. The proposition may be restated more concisely: $\mathcal{O}b(s) \cap \mathcal{A} = \mathcal{A}b(s) \Leftrightarrow \mathcal{A}b(s)$ is closed. By Cartan's Theorem B (see [Hör73, p. 182]), $\mathcal{O}b(s)$ is closed, so implication '$\Rightarrow$' follows.

Conversely we must prove that if $\mathcal{A}b(s)$ is closed then $\mathcal{O}b(s) \cap \mathcal{A} \subseteq \mathcal{A}b(s)$ (the opposite inclusion being trivial).

From the definition 4.18 of local ideal generated by $b(s)$, we see that

$$\mathcal{A}b(s) = (b(s))_{\mathcal{A}} \subseteq (b(s))_{\mathcal{O}} \cap \mathcal{A} = \mathcal{O}b(s) \cap \mathcal{A} \subseteq (b(s))^{l}_{\mathcal{A}};$$

now, being $\mathcal{A}b(s)$ closed, $b(s)$ is slowly decreasing by theorem 4.20 therefore $(b(s))_{\mathcal{A}} = (b(s))^{l}_{\mathcal{A}}$ proving the proposition. □

Corollary 5.5. *Let $a(s) \in \mathcal{A}$ and $b(s) \in \mathcal{H}_m$ such that $a(s)/b(s) \in \mathcal{O}$. Then $a(s)/b(s) \in \mathcal{A}$.*

Proof. By definition 4.32 of $\mathcal{H}_m$, $b(s) = c(s)/d(s)$ where $c(s)$ and $d(s)$ are exponential polynomials; therefore $a(s)/b(s) = a(s)d(s)/c(s)$ that is a Paley–Wiener function by theorems 4.29 and 4.20. □

Remark 5.6. *We note that when proposition 5.4 holds, theorem 4.20 permits us to give $c(s) = a(s)/b(s) \in \mathcal{A}$ an operatorial meaning like the one described in theorem 4.35: for every $w(t) \in \mathcal{E}$ there is a $v(t) \in \mathcal{E}$ such that $b(s)v(t) = w(t)$; so, since $\ker_{\mathcal{E}} b(s) \subseteq \ker_{\mathcal{E}} a(s)$ (employing both theorems 4.24 and 4.25), the map*

$$c(s) = \frac{a(s)}{b(s)} : w(t) \mapsto v(t) + \ker_{\mathcal{E}} b(s) \mapsto a(s)v(t)$$

is well defined.

We have employed often the fact that the ideal in $\mathcal{A}$ generated by surjective operators is closed (theorem 4.20); this result may be extended also to matrix operators [Tre67, thm. 37.2]:

Theorem 5.7. *A matrix $R(s) \in \mathcal{A}^{p\times q}$ is surjective as operator on $\mathcal{E}^q$ if and only if $\ker_{\mathcal{A}} \circ R(s) = \{0\}$ and $\operatorname{im}_{\mathcal{A}} \circ R(s)$ is closed.*

We can obtain a result that is dual to previous theorem; we begin with a very useful lemma:

Lemma 5.8. *Let $\mathcal{R}$ be a generic ring of operators on $\mathcal{E}$; if $R(s) \in \mathcal{R}^{p\times q}$ admits a generalized inverse $G(s) \in \mathcal{R}^{q\times p}$ then*

$$\operatorname{im}_{\mathcal{E}} R = \ker_{\mathcal{E}}(I - RG) \text{ and } \ker_{\mathcal{E}} R = \operatorname{im}_{\mathcal{E}}(I - GR);$$

dually we have that

$$\ker_{\mathcal{R}} \circ R = \operatorname{im}_{\mathcal{R}} \circ (I - RG) \text{ and } \operatorname{im}_{\mathcal{R}} \circ R = \ker_{\mathcal{R}} \circ (I - GR).$$

Proof. Since $R = RGR$ we have $(I - RG)R = 0$ and $R(I - GR) = 0$; so

$$\operatorname{im}_{\mathcal{E}} R \subseteq \ker_{\mathcal{E}}(I - RG) \text{ and } \operatorname{im}_{\mathcal{E}}(I - GR) \subseteq \ker_{\mathcal{E}} R.$$

Conversely if $w \in \ker_{\mathcal{E}}(I - RG)$ then

$$0 = (I - RG)w = w - RGw \;\Rightarrow\; w = RGw \in \operatorname{im}_{\mathcal{E}} R$$

thus $\ker_{\mathcal{E}}(I - RG) \subseteq \operatorname{im}_{\mathcal{E}} R$; even more trivially we obtain that

$$w \in \ker_{\mathcal{E}} R \;\Rightarrow\; w = (I - GR)w \;\Rightarrow\; w \in \operatorname{im}_{\mathcal{E}}(I - GR).$$

The dual inclusions follows similarly. □

Theorem 5.9. *Given $R(s) \in \mathcal{A}^{p\times q}$, there exists a matrix $X(s) \in \mathcal{A}^{q\times p}$ such that $X(s)R(s) = I$ if and only if $\ker_{\mathcal{E}} R(s) = \{0\}$ and $\operatorname{im}_{\mathcal{E}} R(s)$ is closed.*

Proof. If $R(s)$ has a left inverse $X(s)$, by lemma 5.8 $\operatorname{im}_{\mathcal{E}} R(s) = \ker_{\mathcal{E}}(I - R(s)X(s))$ which is always closed. The operator $R(s)\circ$ is injective since $R(s)v = 0$ implies that $v = Iv = X(s)R(s)v = 0$.

On the other hand, since $\operatorname{im}_{\mathcal{E}} R(s)$ is closed, it is complete in the topology of $\mathcal{E}^p$ thus is itself an F-space; we can apply the open mapping theorem 2.6 to $\bar{R}(s) : \mathcal{E}^q \to \operatorname{im}_{\mathcal{E}} R(s)$ since it is clearly an injective and onto linear mapping between F-spaces: $\bar{X} = \bar{R}(s)^{-1}$ exists, is linear and continuous.

The definition of $\bar{X}$ implies that $R(s)f = v \Leftrightarrow \bar{X}v = f$. Since $R(s)$ is shift invariant

$$\sigma_\tau v = \sigma_\tau R(s)f = R(s)\sigma_\tau f \;\Rightarrow\; \bar{X}\sigma_\tau v = \bar{X}R(s)\sigma_\tau f = \sigma_\tau f = \sigma_\tau \bar{X} v$$

also $\bar{X}$ is shift invariant. Modifying slightly the proof of lemma 2.15 we have that $\bar{\chi} : v \mapsto \langle \bar{\chi}, v\rangle = (\bar{X}\check{v})(0)$ ($\check{v}(t) = v(-t)$ as in (2.8)) is a linear functional on $\operatorname{im}_{\mathcal{E}} R(s)$; by the Hahn–Banach theorem 2.7 it can be extended to a linear functional χ on $\mathcal{E}^p$.

Now, if $X(s) = \hat{\chi}(s)$, matrix Laplace transform of χ, then

$$\begin{aligned}\forall v(t) \in \operatorname{im}_{\mathcal{E}} R(s),\; (X(s)v)(\tau) &= \langle \chi, \sigma_\tau \check{v}\rangle = \langle \bar{\chi}, \sigma_\tau \check{v}\rangle = \langle \bar{\chi}, (\sigma_{-\tau} v)^\vee\rangle \\ &= (\bar{X}(\sigma_{-\tau} v))(0) = (\sigma_{-\tau}(\bar{X}v))(0) = (\bar{X}v)(\tau),\end{aligned}$$

so that $X(s)$ extends in the same way $\bar{X}$ on $\mathcal{E}^p$. By definition of $\bar{X}$, $X(s)R(s) = \bar{X}R(s) = I$. □

Corollary 5.10. *A matrix $R(s) \in \mathcal{A}^{p\times q}$, is left–surjective on $\mathcal{A}$, i.e. $\mathcal{A}^p R(s) = \mathcal{A}^q$, if and only if $\ker_{\mathcal{E}} R(s) = \{0\}$ and $\operatorname{im}_{\mathcal{E}} R(s)$ is closed.*

Proof. We only need to prove that $R(s)$ has a left inverse if and only if $\circ R(s)$ is surjective.

If $X(s)R(s) = I$, every $a(s) \in \mathcal{A}^p$ then $a(s) = b(s)R(s)$ with $b(s) = a(s)X(s)$. Conversely when $\circ R(s)$ is surjective, then we can construct a left inverse $X(s) \in \mathcal{A}^{q\times p}$ having as rows $x_i(s) \in \mathcal{A}^p$ such that $x_i(s)R(s) = e_i$ where e_i is the standard basis of the free module $\mathcal{A}^q$. □

Remark 5.11. *This corollary is the dual of theorem 5.7, but there is no proper dual of theorem 5.9. Indeed there is no algebraic condition for surjectivity over $\mathcal{E}^q$ analogous to the existence of a left inverse.*

If $R(s)\circ$ is not surjective but its image $\mathrm{im}_{\mathcal{A}} \circ R(s)$ is closed, we still have an interesting result, namely a topological interpretation of the module $\mathcal{M} = \mathrm{coker}_{\mathcal{A}} \circ R(s)$ with respect to the behavior $\mathcal{B} = \ker_{\mathcal{E}} R(s)$.

Definition 5.12. *Given a subset of smooth functions $E \subseteq \mathcal{E}^q$, the **orthogonal** of E is the submodule of $\mathcal{A}^q$*

$$E^{\perp} = \{a(s) \in \mathcal{A}^q : a(s)v = 0 \, \forall v(t) \in E\} .$$

Dually the orthogonal of a subset $A \subseteq \mathcal{A}^q$ of Paley–Wiener functions is the $\mathcal{A}$-submodule of $\mathcal{E}^q$

$$A^{\perp} = \{v(t) \in \mathcal{E}^q : a(s)v = 0 \, \forall a(s) \in A\} .$$

We state in the following proposition some very basic but useful facts about orthogonals.

Proposition 5.13. *If A and B are subsets of $\mathcal{E}^q$ or of $\mathcal{A}^q$, the following statements are always true:*

- *$A^{\perp}$ is closed;*
- *$A \subseteq B \Rightarrow A^{\perp} \supseteq B^{\perp}$;*
- *$A \subseteq A^{\perp\perp}$;*
- *$A^{\perp} = A^{\perp\perp\perp}$.*

Proof. Let us suppose without loss of generality that $A, B \subseteq \mathcal{A}^q$.

Let $v_n(t) \in A^{\perp}$; for every $a(s) \in A$ we have $f_n = a(s)v_n = 0 \in \mathcal{E}$. Since $f_n(t) \to 0$, if $v_n(t) \to v(t)$, then $a(s)v_n \to a(s)v = 0$, therefore $v(t) \in A^{\perp}$.

Let $v(t) \in B^{\perp}$. Then, by definition, $b(s)v = 0$ for every $b(s) \in B$ and, *a fortiori*, $a(s)v = 0$ for every $a(s) \in A \subseteq B$. Therefore $v(t) \in A^{\perp}$.

Now, let $a(s) \in A$: for every $v(t) \in A^{\perp}$, $a(s)v = 0$, therefore $a(s) \in A^{\perp\perp}$.

In conclusion, since $A \subseteq A^{\perp\perp}$ we have that $A^{\perp} \supseteq (A^{\perp\perp})^{\perp} = A^{\perp\perp\perp}$; conversely we have also $A^{\perp} \subseteq (A^{\perp})^{\perp\perp} = A^{\perp\perp\perp}$. □

Next lemma shows which fundamental duality relates kernels and images.

Lemma 5.14. *For every* $R(s) \in \mathcal{A}^{p\times q}$

$$\ker_{\mathcal{E}} R(s) = (\mathrm{im}_{\mathcal{A}} \circ R(s))^{\perp}, \qquad (\ker_{\mathcal{E}} R(s))^{\perp} = \overline{\mathrm{im}_{\mathcal{A}} \circ R(s)},$$
$$\ker_{\mathcal{A}} \circ R(s) = (\mathrm{im}_{\mathcal{E}} R(s))^{\perp}, \qquad (\ker_{\mathcal{A}} \circ R(s))^{\perp} = \overline{\mathrm{im}_{\mathcal{E}} R(s)}.$$

Proof. We prove only first two equations; the proof of the other ones is similar.

Let for simplicity $I = \mathrm{im}_{\mathcal{A}} \circ R(s)$. The first equality is quite trivial: indeed $w(t) \in \ker_{\mathcal{E}} R(s) \Leftrightarrow R(s)w(t) = 0 \Leftrightarrow a(s)R(s)w(t) = 0\, \forall a(s) \in \mathcal{A}^p \Leftrightarrow w(t) \in I^{\perp}$.

Applying '$^{\perp}$' to both members of the equation we get $(\ker_{\mathcal{E}} R(s))^{\perp} = I^{\perp\perp}$, so we prove the second equality if $\overline{I} = I^{\perp\perp}$.

Let us now consider $\mathcal{I} \subseteq \mathcal{E}'$, topologically *isomorphic* to I (via Laplace transform). It is a subspace of $\mathcal{E}'$ and, by proposition 5.13, $\mathcal{I} \subseteq \mathcal{I}^{\perp\perp}$ and the latter set is closed; therefore also $\overline{\mathcal{I}} \subseteq \mathcal{I}^{\perp\perp}$. Let us suppose that

$$\alpha \in \mathcal{I}^{\perp\perp} \text{ but } \alpha \notin \overline{\mathcal{I}}; \tag{5.3}$$

by corollary 2.8 there is a linear functional $\lambda \in (\mathcal{E}')'$ such that $\lambda(\alpha) = 1$ and λ is zero for any other element in $\mathcal{I}$.

The reflexivity of $\mathcal{E}$ implies that there is a $w(t) \in \mathcal{E}$ such that $\lambda(\gamma) = \langle \gamma, w\rangle$ for every $\gamma \in \mathcal{E}'$. Therefore

$$\langle \alpha, w\rangle = 1 \text{ and } \langle \beta, w\rangle = 0\, \forall \beta \in \mathcal{I}. \tag{5.4}$$

Second equation in (5.4) implies that $w(t) \in \mathcal{I}^{\perp}$ so, by (5.3), we have that $\langle \alpha, w\rangle = 0$ that contradicts the first equation in (5.4). □

Corollary 5.15. *For every* $R(s) \in \mathcal{A}^{p\times q}$*,*

$$(\mathrm{im}_{\mathcal{A}} \circ R(s))^{\perp} = \left(\overline{\mathrm{im}_{\mathcal{A}} \circ R(s)}\right)^{\perp} \text{ and } (\mathrm{im}_{\mathcal{E}} R(s))^{\perp} = \left(\overline{\mathrm{im}_{\mathcal{E}} R(s)}\right)^{\perp}.$$

Proof. Let us consider without loss of generality only images over $\mathcal{A}$. We have just showed that

$$\left(\overline{\mathrm{im}_{\mathcal{A}} \circ R(s)}\right)^{\perp} = (\ker_{\mathcal{E}} R(s))^{\perp\perp} = (\mathrm{im}_{\mathcal{A}} \circ R(s))^{\perp\perp\perp};$$

proposition 5.13, stating that $(\mathrm{im}_{\mathcal{A}} \circ R(s))^{\perp} = (\mathrm{im}_{\mathcal{A}} \circ R(s))^{\perp\perp\perp}$, ends the proof. □

Lemma 5.16. *For any $R(s) \in \mathcal{A}^{p\times q}$, if $\operatorname{im}_{\mathcal{E}} R(s)$ is closed then $\operatorname{im}_{\mathcal{A}} \circ R(s)$ is closed too.*

Proof. Let $\bar{R}(s) : \mathcal{E}^q / \ker_{\mathcal{E}} R(s) \to \operatorname{im}_{\mathcal{E}} R(s) \subseteq \mathcal{E}^p$ be the canonical algebraic isomorphism induced by $R(s)$. Since $\operatorname{im}_{\mathcal{E}} R(s)$ is a closed subspace and $\mathcal{E}^q / \ker_{\mathcal{E}} R(s)$ is a Fréchet space by proposition 2.5, then $\bar{R}(s)^{-1}$ is a continuous linear map (open mapping theorem 2.6).

For every $a(s) \in (\ker_{\mathcal{E}} R(s))^\perp$ define $\bar{\rho}_a : \operatorname{im}_{\mathcal{E}} R(s) \to \mathbb{R}$ such that

$$R(s)w = v \;\Rightarrow\; \langle \bar{\rho}_a, \check{v} \rangle = (a(s)\check{w})(0);$$

where $\check{v}(t) = v(-t)$ as defined in (2.8); $\bar{\rho}_a$ is a continuous linear functional on $\operatorname{im}_{\mathcal{E}} R(s)$ since it is the composition of continuous linear mappings: $\bar{\rho}_a = \delta \circ a(s) \circ \bar{R}(s)^{-1} \circ^\vee$.

Applying the Hahn–Banach theorem 2.7 we extend $\bar{\rho}_a$ to a functional $\rho_a \in \mathcal{E}'^p$. Thus, remembering definition (2.15) and (2.11), we have

$$\begin{aligned}(\hat{\rho}_a(s)R(s)w)(t) &= (\hat{\rho}_a(s)v)(t) = \langle \rho_a, \sigma_t \check{v} \rangle = \langle \bar{\rho}_a, (\sigma_{-t} v)^\vee \rangle \\ &= (a(s)\sigma_{-t} w)(0) = (a(s)w)(t)\end{aligned}$$

i.e. every $a(s) \in (\ker_{\mathcal{E}} R(s))^\perp$ is in $\operatorname{im}_{\mathcal{A}} \circ R(s)$: by lemma 5.14 we obtain that

$$\overline{\operatorname{im}_{\mathcal{A}} \circ R(s)} = (\ker_{\mathcal{E}} R(s))^\perp \subseteq \operatorname{im}_{\mathcal{A}} \circ R(s) \;\Rightarrow\; \operatorname{im}_{\mathcal{A}} \circ R(s) \text{ is closed.}$$

□

Theorem 5.17. *Suppose that $R(s) \in \mathcal{A}^{p\times q}$ and let $\mathcal{B} = \ker_{\mathcal{E}} R(s)$ and $\mathcal{M} = \operatorname{coker}_{\mathcal{A}} \circ R(s)$. If $\operatorname{im}_{\mathcal{E}} R(s)$ is closed then $\mathcal{M}$ is the set of linear continuous shift invariant operators of $\mathcal{B}$ into $\mathcal{E}$ (i.e. $\mathcal{M} \cong \mathcal{B}'$, the topological dual of $\mathcal{B}$).*

Proof. It is easy to see that $\mathcal{A}^q / \mathcal{B}^\perp$ is the set of all linear shift invariant operators on $\mathcal{B}$: actually $a(s)w(t) = (a(s) + b(s))w(t)$, for every $w(t) \in \mathcal{B}$ and $b(s) \in \mathcal{B}^\perp$.

If $\operatorname{im}_{\mathcal{E}} R(s)$ is closed, lemma 5.16 shows that also $\operatorname{im}_{\mathcal{A}} \circ R(s)$ is closed. So, by lemma 5.14, $\mathcal{B}^\perp = \operatorname{im}_{\mathcal{A}} \circ R(s) = \mathcal{A}^p R(s)$, that is to say

$$\mathcal{M} = \mathcal{A}^q / \mathcal{A}^p R(s) = \mathcal{A}^q / \mathcal{B}^\perp.$$

Moreover if $R(s) = \hat{\rho}(s)$ with $\rho \in \mathcal{E}'^{p\times q}$, the following diagram ($\mathcal{L}$ is the topological Laplace isomorphism) commutes:

$$\begin{array}{ccccccc}
\mathcal{A}^p & \xrightarrow{\circ R} & \mathcal{A}^q & \xrightarrow{\phi} & \mathcal{M} & \longrightarrow & 0 \\
\downarrow{\scriptstyle \mathcal{L}} & & \downarrow{\scriptstyle \mathcal{L}} & & \downarrow & & \\
\mathcal{E}'^p & \xrightarrow{\star\rho} & \mathcal{E}'^q & \xrightarrow{\phi} & \mathcal{E}'^q/\mathcal{E}'^p \star \rho & \longrightarrow & 0
\end{array}$$

therefore $\mathcal{E}'^p \star \rho \cong \mathcal{A}^p R(s) \cong \mathcal{B}^{\perp}$ and $\mathcal{E}'^q/\mathcal{E}'^p \star \rho \cong \mathcal{M}$. If we let

$$\mathcal{B}^{(\perp)} = \{\alpha \in \mathcal{E}'^q : \forall w(t) \in \mathcal{B} \ \langle \alpha, w \rangle = 0\}$$

then it is easy to see that both $\mathcal{B}' \cong \mathcal{E}'^q/\mathcal{B}^{(\perp)}$. and $\mathcal{B}^{(\perp)} \cong \mathcal{B}^{\perp}$. So we obtain that

$$\mathcal{B}' \cong \mathcal{E}'^q/\mathcal{E}'^p \star \rho \cong \mathcal{M}.$$

□

Chapter 6

Systems with one delay

This chapter presents a brief survey on delay–differential behaviors with commensurate time delays. The shift operator is simply σ such that $(\sigma f)(t) = f(t-1)$; it corresponds to the function $e^{-s} \in \mathcal{A}$, also denoted with z.

6.1 The ring $\mathcal{H}_1$

As shown in section 4.2.3 the ring $\mathcal{H}_1$ is an elementary divisor domain: standard techniques involving the Smith form of matrices may be applied to delay–differential systems once we extend the ring of operators from delay–differential polynomials to $\mathcal{H}_1$.

This extension, which is the price we have to pay for having nice and elegant results, introduces so called *distributed* delays, a particular class of operators with compact support. These kind of operators are not new to researchers in the area of systems over rings: in [KKT86] there is a construction that leads to a ring (strictly included in $\mathcal{H}_1$) containing such operators.

A simple example will better explain how elements of $\mathcal{H}_1$ operate on functions.

Example 6.1. Let $h(s) \in \mathcal{H}_1$ be the function $h(s) = (1-e^{-s})/s$. Given a function $w(t) \in \mathcal{E}$ we have

$$h(s)w = \frac{1-e^{-s}}{s}w = v \;\Leftrightarrow\; (1-\sigma)w = sv \;\Leftrightarrow\; w(t) - w(t-1) = \tfrac{d}{dt}v(t).$$

In order to understand how we can express $v(t)$ in terms of $w(t)$ we have only to remember the way elements of $h(s)$ operate (see theorem 4.35):

$$h(s)w = \frac{1-e^{-s}}{s}w = v \;\Leftrightarrow\; \exists x(t) \in \mathcal{E}:\; sx = w,\; (1-e^{-s})x = v,$$

therefore we can take any function $x(t)$ such that $\frac{d}{dt}x = w$, for example

$$x(t) = \int_0^t w(\tau)\, d\tau$$

and then, since $v = (1-\sigma)x$

$$v(t) = x(t) - x(t-1) = \int_0^t w(\tau)\, d\tau - \int_0^{t-1} w(\tau)\, d\tau = \int_{t-1}^t w(\tau)\, d\tau.$$

This definition leads to correct results because $\ker_{\mathcal{E}} s$ contains constant functions while $\ker_{\mathcal{E}}(1-e^{-s})$ contains every periodic function with period one and this implies that $\ker_{\mathcal{E}} s \subseteq \ker_{\mathcal{E}} 1-e^{-s}$. We could not, for example, take $x = (1-e^{-s})w$ and then find a $v(t)$ such that $sv = x$: we would obtain $v(t) = C + \int_{t-1}^t w(\tau)\, d\tau$ with some constant $C \in \mathbb{R}$. ♣

6.2 Representations in $\mathcal{H}_1$

In this section we show how the fact that $\mathcal{H}_1$ is an elementary divisor domain always permits to find nice representations of any behavior. However, we will not be exhaustive: the following facts are the most important, but there are other results and also other important way to represent a behavior. For example, an input/output representation will be introduced for $\mathcal{H}_m$ in section 7.2.5; for the definition and properties of a *first order* representation see [GL97b].

First of all, given a behavior in latent variable representation (this case includes image representations), it is always possible to find a kernel representation.

Theorem 6.2. *Latent variable elimination is always possible for delay–differential systems with commensurate delays; therefore the projection of every delay–differential latent variable representation admits kernel representation.*

Proof. Suppose that the behavior is defined as

$$\mathcal{B} = \left\{w(t):\; \exists x(t) \in \mathcal{E}^d,\; R(s)w = M(s)x\right\},\; R(s) \in \mathcal{H}_1^{p\times q},\; M(s) \in \mathcal{H}_1^{p\times d}. \quad (6.1)$$

We can suppose without loss of generality that the Smith form (definition 4.1) of $M(s)$ is $\bar{M}(s)Q(s)$, i.e. that $P(s) = I$ (otherwise, premultiplying both $M(s)$ and $R(s)$ by the inverse of $P(s)$, we obtain such an equivalent behavior). So we have

$$M(s) = \begin{bmatrix} \breve{M}(s) & 0 \\ 0 & 0 \end{bmatrix} Q(s) = \begin{bmatrix} M_1(s) \\ 0 \end{bmatrix}$$

where $M_1(s) = [\breve{M}(s)\ 0]Q(s) \in \mathcal{H}_1^{r\times d}$ is a surjective matrix: actually $\breve{M}(s)$ is diagonal with surjective entries (theorem 4.29) and $Q(s)$ is invertible, i.e. there is a $V(s) \in \mathcal{H}_1^{d\times d}$ such that $V(s)Q(s) = Q(s)V(s) = I_d$; in conclusion

$$\forall y(t) \in \mathcal{E}^r\ \exists \bar{x}(t) \in \mathcal{E}^r : \ \breve{M}(s)\bar{x} = y;$$

so if the first r components of $\tilde{x}(t) \in \mathcal{E}^d$ coincide with $\bar{x}(t)$ and we pose $x = V(s)\tilde{x}$, we get

$$M_1(s)x = [\breve{M}(s)\ 0]Q(s)V(s)\tilde{x} = [\breve{M}(s)\ 0]\tilde{x} = \breve{M}(s)\bar{x} = y.$$

If we partition $R(s)$ accordingly to $M(s)$ we obtain

$$R(s)w = M(s)x \ \Leftrightarrow\ R_1(s)w = M_1(s)x \text{ and } R_2(s)w = 0.$$

We see that $\mathcal{B} = \ker_{\mathcal{E}} R_2(s)$: obviously $\mathcal{B} \subseteq \ker_{\mathcal{E}} R_2(s)$ and, vice versa, if $R_2(s)w = 0$ then there is an $x(t) \in \mathcal{E}^d$ such that $M_1(s)x = R_1(s)w$ by surjectivity of $M_1(s)$ and so $R(s)w = M(s)x$. Therefore $w(t) \in \mathcal{B}$. □

Corollary 6.3. *If there is only one delay, every delay–differential image representation has an injective delay–differential image representation and a surjective delay–differential kernel representation.*

Proof. An image representation corresponds to the latent variable representation (6.1) with $R(s) = I$. Using (4.6) we see immediately that $\operatorname{im}_{\mathcal{E}} M(s) = \operatorname{im}_{\mathcal{E}} P_{*1}(s) = \ker_{\mathcal{E}} U_{2*}(s)$, where $P_{*1}(s)$ is injective and $U_{2*}(s)$ surjective; latter relation follows also from the previous proof: with the notation therein used, if $U(s)$ is the inverse of $P(s)$ as in (4.6), then the behavior is defined by

$$U(s)w = \bar{M}(s)Q(s)x \ \Leftrightarrow\ U_{1*}(s)w = M_1(s)x \text{ and } U_{2*}(s)w = 0$$

and $\mathcal{B} = \ker_{\mathcal{E}} U_{2*}(s)$. □

A behavior that admits a kernel representation does not admit in general an image representation: this problem concerns controllability that will be treated in the following section.

Nevertheless, part of the above corollary still holds for kernel representations:

Proposition 6.4. *Every delay–differential kernel representation with commensurate delays always admits a surjective delay–differential kernel representation.*

Proof. This fact is very simple to prove: if $\mathcal{B} = \ker_{\mathcal{E}} R(s)$ with $R(s) \in \mathcal{H}_1^{p\times q}$ and the Smith form (4.1) of the matrix is $P(s)\breve{R}(s)Q(s)$, then

$$\mathcal{B} = \ker_{\mathcal{E}} \breve{R}(s)Q(s)$$

since $P(s)$ admits a left inverse; obviously $\breve{R}(s)Q(s)$ is surjective being the product of surjective matrices: the first by theorem 4.29, the second being right invertible. □

We show now which relation exists between two kernel representations of the same behavior or, more generally, of two behaviors when one is a subset of the other.

Actually, the following theorem provides a test which allows us to verify whether the kernel of an operator is a subset of the kernel of another operator. The proof may be found in [GL97a, prop. 4.4 (2)].

Theorem 6.5. *If $\mathcal{B}_1 = \ker_{\mathcal{E}} R_1(s)$, $R_1(s) \in \mathcal{H}_1^{p_1\times q}$ and $\mathcal{B}_2 = \ker_{\mathcal{E}} R_2(s)$, $R_2(s) \in \mathcal{H}_1^{p_2\times q}$, then $\mathcal{B}_1 \subseteq \mathcal{B}_2$ if and only if there is a matrix $X(s) \in \mathcal{H}_1^{p_2\times p_1}$ such that $X(s)R_1(s) = R_2(s)$.*

6.3 Controllability in $\mathcal{H}_1$

Validity of theorems 3.12 and 3.13, that give necessary and sufficient conditions for controllability, have been verified independently, and with rather different algebraic tools, by H. Glüsing Lüerßen [GL97a] and P. Rocha together with J. C. Willems [RW97] for delay–differential behaviors. The latter proof relies heavily on 2-dimensional algebraic techniques (applied to the operator ring $\mathbb{R}[s, z]$) that become false for n-dimensional systems as soon as $n > 2$ (see [YP84]) while the proof given in [GL97a] takes advantage of the existence of the Smith form of matrices therein established: both are not useful in the case of non commensurate delays.

The following theorem shows how differential behaviors are similar, with respect to controllability, to delay–differential behaviors with commensurate delays:

Theorem 6.6. *Given the delay–differential behavior $\mathcal{B} = \ker_{\mathcal{E}} R(s)$, $R(s) \in \mathcal{H}_1^{p\times q}$, the following conditions are equivalent:*

- *$R(s)$ admits a generalized inverse;*
- *$\mathcal{B}$ admits an image representation;*
- *$\mathcal{B}$ is controllable;*
- *the rank of the complex matrix $R(\lambda)$ does not depend on $\lambda \in \mathbb{C}$;*
- *the module $\mathcal{M} = \operatorname{coker}_{\mathcal{H}_1} \circ R(s) = \mathcal{H}_1^q/\mathcal{H}_1^p R(s)$ is (torsion free or projective or free) controllable.*

Proof. The equivalence of the first four conditions is a simple consequence of theorem 7.12 and of the fact that $\mathcal{H}_1$ is a Bézout domain. As regards last condition, we have to prove that the torsion free module $\mathcal{M}$ is free.

Since $\mathcal{H}_1$ is an elementary divisor domain, $R(s)$ admits a Smith form

$$R(s) = P(s)\bar{R}(s)Q(s) = P_1(s)\breve{R}(s)Q_1(s);$$

we will denote by $q_i(s)$ the rows of $Q(s)$; being rows of an invertible matrix, they are linearly independent. If r is the rank of $R(s)$, then the rows of $Q_1(s)$ are the first r rows $q_i(s)$. Moreover, we remark that, by surjectivity of $P_1(s)$,

$$\mathcal{H}_1^p R(s) = \mathcal{H}_1^r \breve{R}(s)Q_1(s). \tag{6.2}$$

First we want to show that $d_i(s)$, the r diagonal elements of $\breve{R}(s)$, are constants.

Every element in $\mathcal{H}_1^r \breve{R}(s)Q_1(s)$ is a linear combination with coefficients $h_i(s) \in \mathcal{H}_1(s)$ of $d_i(s)q_i(s)$. Obviously $d_j(s)q_j(s) \in \mathcal{H}_1^p R(s)$ therefore, under the hypothesis that $\mathcal{M}$ is torsion free, we have that $q_j(s) \in \mathcal{H}_1^p R(s) = \mathcal{H}_1^r \breve{R}(s)Q_1(s)$. So

$$q_j(s) = \textstyle\sum_{i=1}^r h_i(s)d_i(s)q_i(s) \quad \text{therefore}$$
$$h_1(s)d_1(s)q_1(s) + \cdots + (h_j(s)d_j(s) - 1)q_j(s) + \cdots + h_r(s)d_r(s)q_r(s) = 0;$$

being linearly independent, the coefficients of $q_i(s)$ must be zero and, since $\mathcal{H}_1$ is a domain, $h_i(s) = 0$ for $i \neq j$ and $h_j(s)d_j(s) = 1$. This equation implies that $h_j(s)$ and $d_j(s)$ have no zeros, thus are constants; hence, by 6.2,

$$\mathcal{H}_1^p R(s) = \mathcal{H}_1^r Q_1(s).$$

Now we show that the elements

$$\mathsf{q}_i = q_i(s) + \mathcal{H}_1^p R(s),\ i = r+1, \dots, q$$

are a basis of $\mathcal{M}$.

Actually they are independent: $\sum_{r+1}^{q} h_j(s)\mathsf{q}_j = \mathsf{0}$ if and only if

$$\sum_{i=r+1}^{q} h_i(s)q_i(s) \in \mathcal{H}_1^r Q_1(s) \ \Leftrightarrow \ \sum_{i=r+1}^{q} h_i(s)q_i(s) = \sum_{j=1}^{r} k_j(s)q_j(s); \tag{6.3}$$

if we set $h_j(s) = -k_j(s)$ for $j = 1, \dots, r$, then equation (6.3) is equivalent to $h(s)Q(s) = 0$, true if and only if $h(s) = 0$ by invertibility of $Q(s)$.

They are also a generating set: indeed let $\mathsf{a} \in \mathcal{M}$, i.e. $\mathsf{a} = a(s) + \mathcal{H}_1^r Q_1(s)$; if we let $b(s) = a(s)Q(s)^{-1}$, then partitioning $b(s) = [h(s)\ k(s)]$ with $k(s) \in \mathcal{H}_1^{q-r}$, we have also $a(s) = h(s)Q_1(s) + k(s)Q_2(s)$ therefore

$$\mathsf{a} = k(s)Q_2(s) + \mathcal{H}_1^r Q_1(s) = \sum_{i=r+1}^{q} k_i(s)\mathsf{q}_i.$$

□

We note that a delay–differential behavior is autonomous (definition 3.16) if and only if every admissible trajectory that is zero in the past (i.e. $\forall t \leq \tau$ for some τ) is zero everywhere (trivially by linearity and shift–invariance). For delay–differential behaviors in $\mathcal{H}_1$ we show that (see also [Val98])

Proposition 6.7. *A delay–differential behavior $\mathcal{B} = \ker_{\mathcal{E}} R(s)$ is autonomous if and only if $R(s) \in \mathcal{H}_1^{p\times q}$ has full column rank q.*

Proof. If $\mathcal{B}$ is autonomous, and $R(s)$ does not have full column rank, then, using the Smith form (4.1) of $R(s)$ we can write

$$R(s)w = 0 \ \Leftrightarrow \ P(s)\bar{R}(s)Q(s)w = 0 \ \Leftrightarrow \ \bar{R}(s)Q(s)w = 0 \ \Leftrightarrow \ [\breve{R}(s)\ 0]Q(s)w = 0$$

exploiting the invertibility of $P(s)$; if we take $\bar{w}(t) \in \mathcal{E}^q$ containing as last component a function that is zero in the past and non zero in the future, then $[\breve{R}(s)\ 0]\bar{w} = 0$; if $V(s)$ is the inverse of $Q(s)$ we can find a trajectory $0 \neq w(t) \in \mathcal{H}_1^q$ such that $w = V(s)\bar{w}$ hence $Q(s)w = \bar{w}$:

$$[\breve{R}(s)\ 0]Q(s)w = [\breve{R}(s)\ 0]\bar{w} = 0 \ \Rightarrow \ w(t) \in \mathcal{B}.$$

Being $V(s)$ the Laplace transform of a compact support distribution, even $w(t)$ is zero in the past, contradicting the hypothesis.

On the converse, if $R(s)$ has full column rank, then, using the same notation as before, $\breve{R}(s) \in \mathcal{H}_1^{d\times d}$; therefore

$$w(t) \in \mathcal{B} \ \Leftrightarrow \ \breve{R}(s)Q(s)w = 0 \ \Leftrightarrow \ \breve{R}(s)\bar{w} = 0, \ \bar{w} = Q(s)w.$$

The equation $\breve{R}(s)\bar{w} = 0$ is a system of d scalar equations $d_i(s)\bar{w}_i = 0$, i.e. $\bar{w}_i(t)$ are finite linear combinations of polynomial–exponential functions (see remark 4.16); they are real holomorphic functions that cannot vanish only on bounded open intervals. Therefore $\bar{w}(t)$, and consequently $w(t)$, is zero in the past if and only if it is the zero trajectory. □

Finally we get a decomposition of delay–differential behaviors that extends theorem 3.17:

Theorem 6.8. *Given a delay–differential behavior $\mathcal{B} = \ker_{\mathcal{E}} R(s)$, $R(s) \in \mathcal{H}_1^{p\times q}$, it is always possible to find two delay–differential behaviors $\mathcal{B}_a$, autonomous, and $\mathcal{B}_c$, controllable, such that $\mathcal{B} = \mathcal{B}_a \oplus \mathcal{B}_c$; the controllable subsystem is uniquely determined.*

See, for the proof, [Val98].

Chapter 7

Systems with m delays

Delay–differential behaviors with noncommensurate time delays have intrinsic properties that make it difficult to extend results and sometimes even definitions given in the commensurate case.

This chapter may seem to be incomplete, but this is effectively the state of the art of behavioral theory: there are some partial results and many open problems.

7.1 Representations in $\mathcal{H}_m$

L. Habets developed a different approach to dynamical systems recently proposed in [Hab98] that, roughly speaking, starts from a module $\mathcal{M}$ over a ring $\mathcal{R}$; unlike the module theory of Fliess described in section 3.2, $\mathcal{M}$ is a function space considered as a module over a ring of operators acting on it.

In our context $\mathcal{E}$ is such a module whereas the ring could be the polynomial ring, the ring $\mathcal{H}_m$, or the algebra of Paley–Wiener functions $\mathcal{A}$.

His approach substantially aims at giving conditions which allow us to verify whether two different kernel representations are equivalent, i.e. define the same behavior.

Noting that some operator rings cannot describe properly the algebraic transformations from a representation to an equivalent one, he extends the operator ring in a way that resembles the definition of $\mathcal{H}_m$. He considers the set fractions of elements in the ring such that, as operators, the denominator is surjective and its kernel is included in the numerator's kernel; in formulas:

$$\mathcal{R}_{\mathcal{M}} = \left\{ \frac{p}{q} : \, p, q \in \mathcal{R}, \, \mathrm{im}_{\mathcal{M}}\, q = \mathcal{M} \text{ and } \ker_{\mathcal{M}} q \subseteq \ker_{\mathcal{M}} p \right\}.$$

This definition permits to describe the way $r = p/q \in \mathcal{R}_{\mathcal{M}}$ operates on $\mathcal{M}$:

$$\frac{p}{q} : \mathcal{M} \to \mathcal{M}, \, m \mapsto \frac{p}{q} m = w \;\Leftrightarrow\; \exists x \in \mathcal{M} \text{ such that } qx = m \text{ and } px = w;$$

which is well defined as we showed in 4.35 for $\mathcal{H}_m$.

This analogy with $\mathcal{H}_m$ is concrete: actually Habets proves [Hab98, thm. 5.4] that if $\mathcal{R} = \mathbb{R}\left[\frac{d}{dt}, \boldsymbol{\sigma}\right]$ is the delay–differential polynomial ring with m time delays and $\mathcal{M} = \mathcal{E}$, then $\mathcal{R}_{\mathcal{M}} = \mathcal{H}_m$. Therefore his results about $\mathcal{R}_{\mathcal{M}}$ hold true in our context too. The most important one is the following [Hab98, p. 10-11]:

Theorem 7.1. *Let $\mathcal{B}_1 = \ker_{\mathcal{E}} R_1(s)$ and $\mathcal{B}_2 = \ker_{\mathcal{E}} R_2(s)$ with $R_1(s) \in \mathcal{H}_m^{p_1 \times q}$ and $R_2(s) \in \mathcal{H}_m^{p_2 \times q}$; suppose that $R_1(s)$ has full row rank. Then*

$$\mathcal{B}_1 \subseteq \mathcal{B}_2 \;\Leftrightarrow\; \exists X(s) \in \mathcal{H}_m^{p_2 \times p_1} \text{ such that } X(s) R_1(s) = R_2(s).$$

The proof of this theorem is based on this idea: if $R_1(s)$ is a full row rank matrix, then it is surjective as operator on $\mathcal{E}^q$, because there is a matrix $Y(s) \in \mathcal{H}_m^{q \times p_1}$ and a scalar $q(s) \in \mathcal{H}_m$ (always surjective by theorem 4.29) such that $R_1(s) Y(s) = q(s) I$. Therefore

$$\forall v(t) \in \mathcal{E}^p \, \exists \bar{v}(t) \in \mathcal{E}^p : \, q(s)\bar{v} = v \;\Rightarrow\; w = Y(s)\bar{v} \in \mathcal{E}^q \text{ and } R(s) w = q(s)\bar{v} = v.$$

So we can define $X(s) = Y(s) R_2(s) q^{-1}(s)$ where the 'fraction' is still in $\mathcal{H}_m$ due to the kernel inclusion $\mathcal{B}_1 \subseteq \mathcal{B}_2$.

The above result can be extended to generic convolutional behaviors:

Theorem 7.2. *Let $\mathcal{B}_1 = \ker_{\mathcal{E}} R_1(s)$ and $\mathcal{B}_2 = \ker_{\mathcal{E}} R_2(s)$ with $R_1(s) \in \mathcal{A}^{p_1 \times q}$ and $R_2(s) \in \mathcal{A}^{p_2 \times q}$; suppose that $R_1(s)$ has full row rank. Then*

$$\mathcal{B}_1 \subseteq \mathcal{B}_2 \;\Leftrightarrow\; \exists X(s) \in \mathcal{O}^{p_2 \times p_1} \text{ such that } X(s) R_1(s) = R_2(s).$$

Proof. We know that $\mathcal{B}_1 \subseteq \mathcal{B}_2$ implies that $\ker_{\mathcal{P}_e} R_1(s) \subseteq \ker_{\mathcal{P}_e} R_2(s)$, therefore, by theorem 4.25, there is an holomorphic matrix $X(s)$ such that $X(s) R_1(s) = R_2(s)$.

On the other hand, the existence of $X(s)$ implies by the same theorem that $\ker_{\mathcal{P}_e} R_1(s) \subseteq \ker_{\mathcal{P}_e} R_2(s)$; however, $\ker_{\mathcal{P}_e} R_2(s) \subseteq \ker_{\mathcal{E}} R_2(s) = \mathcal{B}_2$. Employing theorem 4.24 we have also

$$\mathcal{B}_1 = \overline{\ker_{\mathcal{P}_e} R_1(s)} \subseteq \overline{\ker_{\mathcal{P}_e} R_1(s)} \subseteq \mathcal{B}_2.$$

□

We can drop the hypothesis that $R_1(s)$ is a full row rank matrix, and instead assume that $\operatorname{im}_{\mathcal{E}} R_1(s)$ (or $\operatorname{im}_{\mathcal{A}} \circ R_1(s)$) is closed, condition that is true, for instance, when $R_1(s)$ admits a generalized inverse (lemma 5.8). We obtain, for convolutional behaviors, that

Theorem 7.3. *Let $\mathcal{B}_1 = \ker_{\mathcal{E}} R_1(s)$ and $\mathcal{B}_2 = \ker_{\mathcal{E}} R_2(s)$ with $R_1(s) \in \mathcal{A}^{p_1 \times q}$ and $R_2(s) \in \mathcal{A}^{p_2 \times q}$; suppose that $\operatorname{im}_{\mathcal{E}} R_1(s)$ is closed. Then*

$$\mathcal{B}_1 \subseteq \mathcal{B}_2 \Leftrightarrow \exists X(s) \in \mathcal{A}^{p_2 \times p_1} \text{ such that } X(s)R_1(s) = R_2(s).$$

Proof. If $X(s)R_1(s) = R_2(s)$ then every $w(t) \in \ker_{\mathcal{E}} R_1(s)$ is also in $\ker_{\mathcal{E}} R_2(s)$.

Conversely, as lemma 5.16 shows, $\operatorname{im}_{\mathcal{A}} \circ R_1(s)$ is closed hence it is equal to $(\ker_{\mathcal{E}} R_1(s))^{\perp}$ by lemma 5.14.

Since $\ker_{\mathcal{E}} R_1(s) \subseteq \ker_{\mathcal{E}} R_2(s)$, each row $r_i(s)$ of $R_2(s)$ annihilates every $w(t) \in \ker_{\mathcal{E}} R_1(s)$; so it belongs to $(\ker_{\mathcal{E}} R_1(s))^{\perp} = \operatorname{im}_{\mathcal{A}} \circ R_1(s)$, i.e. there are $x_i(s)$ such that $r_i(s) = x_i(s)R_1(s)$. □

7.2 Controllability in $\mathcal{H}_m$

For systems more generic than delay–differential systems with commensurate delays there is no analogous of theorem 6.6; this section is mainly devoted to investigate in $\mathcal{H}_m$ and $\mathcal{A}$ relations existing between conditions similar to those that in theorem 6.6 are equivalent.

7.2.1 Algebraic controllabilities

This first part deals with the definitions of controllability relative to the module theoretic approach that were introduced in section 3.2.2.

A fundamental theorem about controllable systems in this framework is proved by H. Mounier in [Mou98a]:

Theorem 7.4. *Let $R\left(\frac{d}{dt},\boldsymbol{\sigma}\right) \in \mathbb{R}\left[\frac{d}{dt},\boldsymbol{\sigma}\right]^{p\times q}$, with full row rank, be the presentation matrix of a delay–differential system $\mathcal{M} = \operatorname{coker}_{\mathbb{R}[\frac{d}{dt},\boldsymbol{\sigma}]} \circ R\left(\frac{d}{dt},\boldsymbol{\sigma}\right)$ which we suppose $\mathbb{R}\left[\frac{d}{dt},\boldsymbol{\sigma},\boldsymbol{\sigma}^{-1}\right]$-torsion free controllable. Then $\mathcal{M}$ is spectrally controllable if and only if it is $\mathcal{H}_m$-torsion free controllable.*

A brief explanation is necessary: $\mathcal{M}$ is supposed to be $\mathbb{R}\left[\frac{d}{dt},\boldsymbol{\sigma},\boldsymbol{\sigma}^{-1}\right]$-torsion free controllable to permit the existence of 'forward shifts' (see remark 3.18): this hypothesis is no longer necessary if we assume, as we do for behavioral systems, that the basic ring of operators for delay–differential systems is exactly $\mathbb{R}\left[\frac{d}{dt},\boldsymbol{\sigma},\boldsymbol{\sigma}^{-1}\right]$, the ring of delay–differential Laurent polynomials.

Theorem 7.4 may be extended in two ways: first of all we can remove the 'full rank' constraint but we cannot have a necessary and sufficient condition for spectral controllability:

Theorem 7.5. *Suppose that $R(s) \in \mathcal{A}^{p\times q}$ is the presentation matrix of the system $\mathcal{M}$. Then $\mathcal{M}$ is spectrally controllable if it is torsion free controllable.*

Proof. Suppose that $\mathcal{M}$ is not spectrally controllable, i.e. there is a value $s_0 \in \mathbb{C}$ such that $\operatorname{rank} R(s_0) < r$ where r is the rank of $R(s)$. We want to show that there is a torsion element in $\mathcal{M}$, i.e. an element $\mathsf{m} = m(s) + \mathcal{A}^p R(s)$ that is not zero, $m(s) \notin \mathcal{A}^p R(s)$, but vanishes when it is multiplied by an element $a(s) \in \mathcal{A}$: $a(s)\mathsf{m} = \mathsf{0}$, that is to say $a(s)m(s) \in \mathcal{A}^p R(s)$.

The existence of the Smith form of $R(s)$ over $\mathcal{O}$ implies that, since $\mathcal{O}$ is a domain, every element is injective as a multiplier in $\mathcal{O}$, thus by (4.5) there exist a left–injective matrix with full row rank $p-r$ $U(s) \in \mathcal{O}^{(p-r)\times p}$ such that $\ker_{\mathcal{O}} \circ R(s) = \operatorname{im}_{\mathcal{O}} \circ U(s)$. We know that

$$\operatorname{rank}_{\mathbb{C}} U(s_0) = \operatorname{rank} U(s) = p - \operatorname{rank} R(s) < p - \operatorname{rank}_{\mathbb{C}} R(s_0)$$

therefore $\operatorname{rank}_{\mathbb{C}} U(s_0) + \operatorname{rank}_{\mathbb{C}} R(s_0) < p$ which implies that

$$\exists c \in \mathbb{C}^p,\ c \notin \operatorname{im}_{\mathbb{C}} \circ U(s_0) \text{ such that } cR(s_0) = 0. \tag{7.1}$$

The vector $cR(s)/(s-s_0)$ is holomorphic so by corollary 5.5, if we let

$$m(s) = \frac{cR(s)}{s-s_0} \in \mathcal{A}^q, \text{ we have that } \mathsf{m} = m(s) + \mathcal{A}^p R(s) \in \mathcal{M}.$$

Then clearly

$$(s-s_0)m(s) = cR(s) \in \mathcal{A}^p R(s) \ \Rightarrow\ (s-s_0)\mathsf{m} = 0$$

therefore if we can prove that $\mathsf{m} \neq 0$ then this is the torsion element we are searching for.

If m were zero then, by definition, $m(s) \in \mathcal{A}^p R(s)$ hence there is a $b(s) \in \mathcal{A}^p$ such that

$$m(s) = \frac{cR(s)}{s-s_0} = b(s)R(s) \ \Rightarrow\ (c-(s-s_0)b(s))\,R(s) = 0;$$

the kernel of $\circ R(s)$ is the image over $\mathcal{O}$ of $\circ U(s)$ therefore

$$\exists c(s) \in \mathcal{O}^{p-r} :\ c-(s-s_0)b(s) = c(s)U(s)$$

and evaluating at s_0, $c = c(s_0)U(s_0)$ contradicting the initial hypothesis (7.1). □

Corollary 7.6. *Suppose that $R(s) \in \mathcal{H}_m^{p\times q}$ is the presentation matrix of the system $\mathcal{M}$. Then $\mathcal{M}$ is spectrally controllable if it is torsion free controllable.*

Proof. This result follows trivially from the preceding proof *mutatis mutandi*. □

Another way to extend theorem 7.4 removing the 'full rank' condition but still mantaining necessary and sufficient conditions for spectral controllability consists in testing torsion freeness over a larger ring.

Theorem 7.7. *Suppose that $R(s) \in \mathcal{H}_m^{p\times q}$ (or $R(s) \in \mathcal{A}^{p\times q}$) is the presentation matrix of the system $\mathcal{M}$. Then $\mathcal{M}$ is spectrally controllable if and only if it is $\mathcal{O}$ (torsion free, projective and free) controllable.*

Proof. We remind that torsion free, projective and free controllability are equivalent over $\mathcal{O}$, an elementary divisor domain, as we showed in the proof of theorem 6.6. We consider here the weakest: torsion free controllability.

If $\mathcal{M}$ is spectrally controllable, then $R(s)$ does not loose rank: the matrix $\breve{R}(s)$ of the Smith form (4.2) of $R(s)$ over $\mathcal{O}$ cannot have zeros; it is invertible and, without loss of generality, we can suppose that it is the identity I. So we can write as in (4.3) $R(s) = P_1(s)Q_1(s)$ where $P_1(s)$ is left-injective and $Q_1(s)$ right invertible: $Q_1(s)V_1(s) = I$.

If given an $a(s) \in \mathcal{O}^q$ there is a $h(s) \in \mathcal{O}$ such that

$$\exists b(s) \in \mathcal{O}^p : \ h(s)a(s) = b(s)R(s) \ \Rightarrow \ h(s)a(s)V(s) = b(s)P(s);$$

if $h(s_0) = 0$ then $b(s_0)P(s_0) = 0$ but by injectivity of $\circ P(s)$, $b(s_0) = 0$. So, since every zero of $h(s)$ is a zero of $b(s)$ then $c(s) = b(s)/h(s)$ is holomorphic and $a(s) = c(s)R(s)$.

In a more abstract language,

$$\forall \mathsf{a} = a(s) + \mathcal{O}^p R(s) \in \mathcal{O} \otimes \mathcal{M}, \ h(s)\mathsf{a} = 0 \ \Rightarrow \ \mathsf{a} = 0$$

hence the system is $\mathcal{O}$–torsion free controllable.

The proof of the converse statement follows, with only slight modifications, the proof of theorem 7.5. □

A condition strictly related to (behavioral) controllability is the existence of a generalized inverse of the presentation matrix, as we shall show in corollary 7.13: we prove that for convolutional systems this is equivalent to $\mathcal{A}$–projective controllability.

Theorem 7.8. *Let $R(s) \in \mathcal{A}^{p\times q}$ be the presentation matrix of $\mathcal{M} = \operatorname{coker}_{\mathcal{A}} \circ R(s)$. Then $\mathcal{M}$ is projective controllable if and only if $R(s)$ admits a generalized inverse $G(s) \in \mathcal{A}^{q\times p}$.*

Proof. By lemma 3.25 $\mathcal{M}$ is projective if and only if there are n elements $\mathsf{x}_i \in \mathcal{M}$ and n $\mathcal{A}$–homomorphisms $\mathsf{f}_i : \mathcal{M} \to \mathcal{A}$ such that

$$\forall \mathsf{m} \in \mathcal{M}, \ \mathsf{m} = \sum_{i=1}^{n} \mathsf{f}_i(\mathsf{m})\mathsf{x}_i$$

or, if we let $\mathsf{m} = m(s) + \mathcal{A}^p R(s)$ and $\mathsf{x}_i = x_i(s) + \mathcal{A}^p R(s)$,

$$m(s) - \sum_{i=1}^{n} \mathsf{f}_i(\mathsf{m}) x_i(s) \in \mathcal{A}^p R(s). \tag{7.2}$$

Now, if $R(s) = R(s)G(s)R(s)$ with $G(s) \in \mathcal{A}^{q\times p}$ and we let $Y(s) = I - G(s)R(s)$, then $R(s)Y(s) = 0$: the columns of $Y(s)$, $y_i(s)^\top$, are homomorphisms of $\mathcal{M}$ into $\mathcal{A}$:

$$\mathsf{m}y_i(s)^\top = (m(s) + \mathcal{A}^p R(s))y_i(s)^\top = m(s)y_i(s)^\top.$$

So, if we put $\mathsf{f}_i(\mathsf{m}) \triangleq \mathsf{m}y_i(s)^\top$ and $x_i(s) \triangleq e_i$, the canonical basis of $\mathcal{A}^q$, then equation (7.2) becomes

$$m(s) - \sum_{i=1}^{q} m(s)y_i(s)^\top e_i = m(s) - m(s)Y(s) = m(s)G(s)R(s) \in \mathcal{A}^p R(s).$$

On the converse, given the n homomorphisms f_i of equation (7.2), we can define homomorphisms $f_i : \mathcal{A}^q \to \mathcal{A}$ fixing their values at e_j, i.e.

$$\forall i = 1, \dots, n \;\forall j = 1, \dots, q \quad f_i(e_j) \triangleq \mathsf{f}_i(\mathsf{e}_j).$$

Since, moreover, $\mathrm{Hom}_{\mathcal{A}}(\mathcal{A}^q, \mathcal{A}) \cong \mathcal{A}^q$ as we pointed out in section 5.2, every f_i corresponds to a vector $y_i(s) \in \mathcal{A}^q$ and $f_i(m(s)) = m(s)y_i(s)^\top$; we will denote by $Y(s) \in \mathcal{A}^{q\times n}$ the matrix having $y_i(s)^\top$ as its columns. We note also that since f_i are homomorphisms, then $\mathsf{f}_i(\mathsf{0}) = 0$, i.e.

$$m(s) = a(s)R(s) \;\Rightarrow\; m(s)y_i(s)^\top = a(s)R(s)y_i(s) = 0 \;\Rightarrow\; R(s)Y(s) = 0. \quad (7.3)$$

We can rewrite now equation 7.2 as

$$m(s) - \sum_{i=1}^{n} \mathsf{f}_i(\mathsf{m})x_i(s) = m(s)(I - Y(s)X(s)) \in \mathcal{A}^p R(s)$$

where the vectors $x_i(s)$ become the rows of $X(s) \in \mathcal{A}^{n\times q}$. If we take $m(s) = e_i$, then we can construct a matrix $G(s) \in \mathcal{A}^{q\times p}$ such that $I - Y(s)X(s) = G(s)R(s)$; therefore, by equation (7.3)

$$R(s) = R(s)(I - Y(s)X(s)) = R(s)G(s)R(s).$$

□

Corollary 7.9. *Let $R(s) \in \mathcal{H}_m^{p\times q}$ be the presentation matrix of the system $\mathcal{M} = \mathrm{coker}_{\mathcal{H}_m} \circ R(s)$. Then $\mathcal{M}$ is projective controllable if and only if $R(s)$ admits a generalized inverse $G(s) \in \mathcal{H}_m^{q\times p}$.*

7.2.2 Behavioral controllability in $\mathcal{A}$

There is no published result about controllability of delay–differential behaviors with non commensurate delays. In particular the link between rank of $R(\lambda)$ as $\lambda \in \mathbb{C}$ and controllability of $\ker_{\mathcal{E}} R(s)$ is not so clear.

Results given in this section permit to say something in particular cases but the full generality has not been arranged in a satisfactory theoretic framework.

Three main concepts have been introduced so far: existence of image representation, (behavioral) controllability and spectral controllability as defined in 3.27. Even for generic convolutional behaviors, as we shall see in following theorems, the first one implies the second one and controllability implies spectral controllability.

Unfortunately spectral controllability does not always imply the existence of an image representation, as example 7.14 will show, and this is due to number-theoretic properties of the delays involved.

Precisely this intrinsic difficulty suggests that, being so scarcely robust with respect to the value of time delays, the notions of controllability we are using may be improper.

Theorem 7.10. *A convolutional behavior which has an image representation is controllable.*

Proof. Let $\mathcal{B}$ a behavior that admits an image representation $\mathcal{B} = \operatorname{im}_{\mathcal{E}} M(s)$ with $M(s) = \hat{\mu}(s)$, Laplace transform of $\mu \in \mathcal{E}'^{q\times d}$. Given two trajectories $w_i = M(s)v_i \in \mathcal{B}$ we want to find a $T \geq 0$ and a trajectory $w(t) \in \mathcal{B}$ such that $w(t) = w_1(t)$ for $t \leq 0$ and $w(t) = w_2(t-T)$ for $t \geq T$.

The distribution μ has compact support: suppose that for every i and j $\operatorname{supp} \mu_{ij} \subseteq [a,b]$. Since for every $w(t) \in \mathcal{B}$ we have

$$w(\tau) = (\tilde{\mu}v)(\tau) = \langle \check{\mu}, \sigma_{-\tau} v\rangle = \langle \sigma_\tau \check{\mu}, v\rangle \tag{7.4}$$

and $\cup_{i,j} \operatorname{supp} \sigma_\tau \check{\mu}_{ij} \subseteq [\tau - b, \tau - a]$, then $w(\tau)$ depends only on the values of $v(t)$, $t \in [\tau - b, \tau - a]$.

Given any $T > b - a \geq 0$ let us consider the smooth function ψ of lemma 3.15 and let

$$\chi(t) = \psi\left(\frac{t+a}{T+a-b}\right);$$

$\chi(t)$ is zero for $t \leq -a$ and 1 for $t \geq T-b$. Let

$$v(t) = v_1(t) + \chi(t)[v_2(t-T) - v_1(t)] = \begin{cases} v_1(t) & t \leq -a \\ v_2(t-T) & t \geq T-b \end{cases} \tag{7.5}$$

and $w = M(s)v$. Then, when $\tau \leq 0$, the function $w(\tau)$ depends only on the values of $v(t)$ for $t \leq -a$ and in this interval $v(t) = v_1(t)$ by definition (7.5); that is to say: when $\tau \leq 0$, $w(\tau)$ coincides with $w_1(\tau)$.

When $\tau \geq T$, the function $w(\tau)$ depends on $v(t)$ for $t \geq T-b$; analogously then (7.5) implies that $w(\tau) = w_2(\tau - T)$ in the interval $\tau \geq T$. □

Theorem 7.11. *If the convolutional behavior $\mathcal{B} = \ker_{\mathcal{E}} R(s)$ with $R(s) \in \mathcal{A}^{p\times q}$ is controllable then it is also spectrally controllable.*

Proof. Let us consider the Smith form of $R(s)$ over the ring of holomorphic functions (4.1); we suppose its rank over $\mathcal{A}$ equal to r and denote $U_{1*}(s)$ and $U_{2*}(s)$ simply as $U_1(s)$ and $U_2(s)$. Then

$$\ker_{\mathcal{O}} \circ R(s) = \operatorname{im}_{\mathcal{O}} \circ U_2(s) = \mathcal{O}^{p-r} U_2(s), \tag{7.6}$$

by lemma 4.4, and

$$U(s) = \begin{bmatrix} U_1(s) \\ U_2(s) \end{bmatrix}$$

is a square, invertible matrix in $\mathcal{O}^{p\times p}$; therefore

$$\operatorname{im}_{\mathbb{C}} \circ U_1(s_0) \oplus \operatorname{im}_{\mathbb{C}} \circ U_2(s_0) = \mathbb{C}^p \ \forall s_0 \in \mathbb{C}. \tag{7.7}$$

Indeed, if $c = c_1 U_1(s_0) = c_2 U_2(s_0)$, then

$$0 = c_1 U_1(s_0) - c_2 U_2(s_0) = [c_1 \quad -c_2] \begin{bmatrix} U_1(s_0) \\ U_2(s_0) \end{bmatrix} \Leftrightarrow [c_1 \quad -c_2] = 0 \Leftrightarrow c = 0.$$

Note that the dimension of the kernel of a matrix over a field gets bigger as its rank lowers; thus the above relation, together with equation (7.6), shows that

$$R(s) \text{ does not loose rank at } s_0 \Leftrightarrow \ker_{\mathbb{C}} \circ R(s_0) \cap \operatorname{im}_{\mathbb{C}} \circ U_1(s_0) = \{0\}. \tag{7.8}$$

We want to prove that controllability implies condition (7.8): let us take

$$s_0 \in \mathbb{C} \text{ and } c \in \ker_{\mathbb{C}} \circ R(s_0) \cap \operatorname{im}_{\mathbb{C}} \circ U_1(s_0) \tag{7.9}$$

and construct the operator (well defined by corollary 5.5 because $cR(s_0) = 0$)

$$h(s) = \frac{1}{s - s_0} cR(s) \in \mathcal{A}^q.$$

Let $\mathcal{B}_h = \ker_{\mathcal{E}} h(s)$: we show that $\mathcal{B} \subseteq \mathcal{B}_h$.

Indeed: take a $w(t) \in \mathcal{B}$; the way $h(s)$ operates (see remark 5.6) implies that

$$f = h(s)w \Leftrightarrow (s - s_0)f = cR(s)w = 0 \Leftrightarrow \tfrac{d}{dt} f = s_0 f \Leftrightarrow f(t) = ke^{s_0 t}$$

so the image of $\mathcal{B}$ through $h(s)$ consists only in exponentials (that are zero everywhere if and only if they vanish at a single point $t_0 \in \mathbb{R}$).

Let us take any $w(t) \in \mathcal{B}$; since $\mathcal{B}$ is controllable there is a trajectory $\bar{w}(t)$ equal to $w(t)$ as $t \leq 0$ and zero as $t \geq \tau$ for some $\tau \geq 0$. Let $f = h(s)w$ and $\bar{f} = h(s)\bar{w}$.

We know that $h(s) = \hat{\theta}(s)$ is the Laplace transform of $\theta \in \mathcal{E}'$, a distribution with compact support, say $[a, b]$. This implies, as we saw more precisely in theorem 7.10, that $f(t) = \bar{f}(t)$ when $t \leq a$ and $\bar{f}(t) = 0$ when $t \geq \tau + b$: we see at once that $\bar{f} = 0$ and so $f = 0$, proving that $\mathcal{B} \subseteq \mathcal{B}_h$.

Now, in particular, $\ker_{\mathcal{P}_e} R(s) \subseteq \ker_{\mathcal{P}_e} h(s)$ hence, by theorem 4.25,

$$\exists a(s) \in \mathcal{O}^p : \; h(s) = a(s)R(s);$$

by definition of $h(s)$ we also have $(s - s_0)h(s) = cR(s)$ and

$$(c - (s - s_0)a(s))R(s) = 0 \;\Leftrightarrow\; \exists b(s) \in \mathcal{O}^{p-r} : \; c - (s - s_0)a(s) = b(s)U_2(s)$$

by equation (7.6); we deduce that $c = b(s_0)U_2(s_0)$ but, as supposed in (7.9), also $c = c_0 U_1(s_0)$ for some $c_0 \in \mathbb{C}^r$, hence, since the sum in (7.7) is direct, $c = 0$. This result, together with equation (7.8), proves that the system is spectrally controllable. □

If we require that the minors of $R(s)$ not only do not have common zeros, but also satisfy a Bézout equation, then it is easy to see that

Theorem 7.12. *Let $R(s) \in \mathcal{A}^{p\times q}$ have rank r. If its minors $r_i(s)$ of dimension r satisfy a Bézout equation*

$$\sum_i r_i(s) g_i(s) = 1, \; g_i(s) \in \mathcal{A}$$

then $\mathcal{B} = \ker_{\mathcal{E}} R(s)$ admits an image represention that can be constructed using the elements $g_i(s)$.

Proof. By theorem 4.12 $R(s)$ has a generalized inverse $G(s) \in \mathcal{A}^{q\times p}$; the proof in [BR83] is constructive and shows precisely how $G(s)$ is built up starting from $g_i(s)$.

In the end lemma 5.8 shows that if $M(s) = I - G(s)R(s)$, then $\ker_{\mathcal{E}} R(s) = \operatorname{im}_{\mathcal{E}} M(s)$. □

Corollary 7.13. *If $R(s) \in \mathcal{A}^{p\times q}$ admits a generalized inverse $G(s) \in \mathcal{A}^{q\times p}$, then $\mathcal{B} = \ker_{\mathcal{E}} R(s)$ admits an image representation $\mathcal{B} = \operatorname{im}_{\mathcal{E}}(I - G(s)R(s))$.*

If a behavior is only spectrally controllable then, in general, it does not admit an image representation.

Example 7.14. Let $\mathcal{B} = \ker_{\mathcal{E}} R(s)$ be the behavior in kernel representation with

$$R(s) = [h_1(s) \; h_2(s)] = \begin{bmatrix} \frac{e^s - e^{-s}}{2s} & \frac{e^{as} - e^{-as}}{2} \end{bmatrix} = \begin{bmatrix} \frac{i\sin(-is)}{s} & i\sin(-ias) \end{bmatrix}. \quad (7.10)$$

We have already showed in example 4.17 that if $a \notin \mathbb{Q}$, $h_1(s)$ and $h_2(s)$ do not have common zeros, therefore $\mathcal{B}$ is spectrally controllable. Moreover, if a is a Liouville number (4.14), $R(s)$ does not satisfy the hypothesis of theorem 7.12. Is there anyway an image representation of $\mathcal{B}$?

Let us suppose that a is a Liouville number therefore $h_i(s)$ do not have common zeros and do not satisfy a Bézout equation. If $\ker_{\mathcal{E}} R(s) = \operatorname{im}_{\mathcal{E}} N(s)$, with $N(s) = [n_{ij}(s)] \in \mathcal{A}^{2\times d}$, since $R(s)N(s) = 0$, it must be $h_1(s)n_{1j}(s) = -h_2(s)n_{2j}(s)$. We know also that if $h_1(s_0) = 0$, $h_2(s_0)$ cannot be zero hence $n_{2j}(s_0) = 0$; by symmetry

$$y_j(s) \triangleq \frac{n_{1j}(s)}{h_2(s)} = -\frac{n_{2j}(s)}{h_1(s)} \in \mathcal{O}.$$

By corollary 5.5, $y_j(s) \in \mathcal{A}$ and if we let $M(s)^\top \triangleq [h_2(s) \; -h_1(s)]$ (we have trivially

$R(s)M(s) = 0$), and $Y(s) \triangleq [y_j(s)] \in \mathcal{A}^d$, then $N(s) = M(s)Y(s)$ therefore

$$\ker_{\mathcal{E}} R(s) = \operatorname{im}_{\mathcal{E}} N(s) = \operatorname{im}_{\mathcal{E}} M(s)Y(s) \subseteq \operatorname{im}_{\mathcal{E}} M(s) \subseteq \ker_{\mathcal{E}} R(s).$$

This is a contradiction: $\operatorname{im}_{\mathcal{E}} M(s) = \ker_{\mathcal{E}} R(s)$ implies that the operator $M(s)$ has a closed image; the Spectral analysis theorem 4.26 states that it is also injective therefore, by theorem 5.9, $M(s)$ has a left inverse $X(s)$, i.e. there are two elements $x_i(s) \in \mathcal{A}$ such that $x_1(s)h_1(s) + x_2(s)h_2(s) = 1$, impossible by the assumption on a. ♣

7.2.3 A preliminary result

The following results on controllability of delay–differential behaviors need a preliminary theorem.

Theorem 7.15. *Suppose that $R(s) \in \mathcal{A}^{p\times q}$ with rank r admits a generalized inverse over $\mathcal{O}$, i.e.*

$$\exists X(s) \in \mathcal{O}^{q\times p} \text{ such that } R(s)X(s)R(s) = R(s).$$

Then there is a matrix $M(s) \in \mathcal{A}^{q\times d}$ with rank $q - r$ such that

$$\ker_{\mathcal{O}} R(s)\circ = \operatorname{im}_{\mathcal{O}} M(s)\circ \text{ and } \operatorname{im}_{\mathcal{O}} \circ R(s) = \ker_{\mathcal{O}} \circ M(s); \tag{7.11}$$

moreover also $M(s)$ admits a generalized inverse over $\mathcal{O}$.

Remark 7.16. *As the proof will show, theorem 7.15 employs only the fact that $\mathcal{A}$ is a subring of $\mathcal{O}$, hence this result is still valid if we replace $\mathcal{A}$ with any operator ring among the ones we encountered.*

The proof of this theorem needs a rather involved notation, so two simple examples will help in understanding it.

Example 7.17. Given $R(s) = [a\ b\ c\ d] \in \mathcal{A}^{1\times 4}$ (we omit to explicit the dependence of the elements on s), we satisfy the hypotheses of theorem 7.15 with $r = p = 1$, $q = 4$ as soon as $R(\lambda) \neq 0$ for all $\lambda \in \mathbb{C}$: $\mathcal{O}$ is a Bézout domain and in this case the elements of $R(s)$ satisfy the equation $R(s)X(s) = 1$ with $X(s) \in \mathcal{O}^{4\times 1}$.

We show how we can construct a matrix $M(s) \in \mathcal{A}^{4\times 6}$ of rank $q-r = q-p = 3$ such that $R(s)M(s) = 0$: let us consider the matrix

$$\tilde{R}(s) = \begin{bmatrix} w & x & y & z \\ a & b & c & d \end{bmatrix}$$

where the vector $[w\ x\ y\ z]$ belongs to the image of $\circ R(s)$, i.e. it is a scalar multiple of $R(s)$. In this case $\tilde{R}(s)$ has still rank $p = 1$ therefore its minors of order $p+1$ are zero. Note that the minors of $\tilde{R}(s)$ are linear functions of w, x, y or z: we will show that we can write them as particular row-column products.

We have $6 = \binom{4}{2} = \binom{q}{p+1}$ different (ordered) sets of $2 = p+1$ columns of $R(s)$:

$$\rho_1 = \{1,4\}\,,\ \rho_2 = \{2,4\}\,,\ \rho_3 = \{3,4\}\,,\ \rho_4 = \{1,3\}\,,\ \rho_5 = \{2,3\}\,,\ \rho_6 = \{1,2\}$$

that correspond to every minor of order $p+1$ of $\tilde{R}(s)$. The minor given by ρ_1 is $wd - za = 0$ that can be written (up to multiplication by -1):

$$\begin{bmatrix} w & x & y & z \end{bmatrix} \begin{bmatrix} -d \\ 0 \\ 0 \\ a \end{bmatrix}$$

i.e. the first element in the column is $-d$, minor of $R(s)$ corresponding to the set of columns $\{4\} = \rho_1 \setminus \{1\}$ (we consider it with the opposite sign if the row we are considering, 1, occupies an odd position in ρ). Second and third elements are zero ($2, 3 \notin \rho_1$) and the fourth element is a, the minor of column $\{1\} = \rho_1 \setminus \{4\}$ of $R(s)$.

From every ρ_i we can construct such a row and obtain

$$M(s) = \begin{bmatrix} -d & 0 & 0 & -c & 0 & -b \\ 0 & -d & 0 & 0 & -c & a \\ 0 & 0 & -d & a & b & 0 \\ a & b & c & 0 & 0 & 0 \end{bmatrix};$$

it is easy to check that it satisfies equation $R(s)M(s) = 0$.

Note that $-d^3$ is a minor of order $3 = q-p$ of $M(s)$; browsing patiently we could find every other third power of the minors of $R(s)$; some of them is not zero, so $M(s)$ has at least rank 3; it cannot have (full row) rank 4 because otherwise equation $R(s)M(s) = 0$ would imply that $R(s) = 0$. ♣

Example 7.18. Without being so detailed, we show what happens when $p = 2$ and $q = 4$:

$$R(s) = \begin{bmatrix} a & b & c & d \\ \alpha & \beta & \gamma & \delta \end{bmatrix}.$$

Matrix $\tilde{R}(s)$ is now

$$\tilde{R}(s) = \begin{bmatrix} w & x & y & z \\ a & b & c & d \\ \alpha & \beta & \gamma & \delta \end{bmatrix}$$

where $[w\ x\ y\ z]$ is a linear combination of the rows of $R(s)$; its minors of order $p+1 = 3$ correspond to columns in the $d = \binom{q}{p+1} = \binom{4}{3} = 4$ sets

$$\rho_1 = \{1,3,4\}\,,\ \rho_2 = \{2,3,4\}\,,\ \rho_3 = \{1,2,4\}\,,\ \rho_4 = \{1,2,3\}$$

and permit to construct $M(s)$ in the following way:

$$M(s) = \begin{bmatrix} d\gamma - c\delta & 0 & d\beta - b\delta & c\beta - b\gamma \\ 0 & d\gamma - c\delta & a\delta - d\alpha & a\gamma - c\alpha \\ a\delta - d\alpha & b\delta - d\beta & 0 & b\alpha - a\beta \\ c\alpha - a\gamma & c\beta - b\gamma & b\alpha - a\beta & 0 \end{bmatrix}.$$

$M(s)$ has obviously rank $2 = p$ since among its minors of order $q - p = 2$ there are the squares of the minors of order $p = 2$ of $R(s)$. ♣

Lemma 7.19. *Let $R(s) \in \mathcal{A}^{p\times q}$ with full row rank and*

$$\exists X(s) \in \mathcal{O}^{q\times p} \text{ such that } R(s)X(s) = I.$$

Then there is a matrix $M(s) \in \mathcal{A}^{q\times d}$ with rank $q - p$ admitting a generalized inverse over $\mathcal{O}$ and such that

$$\ker_{\mathcal{O}} R(s)\circ = \operatorname{im}_{\mathcal{O}} M(s)\circ \text{ and } \operatorname{im}_{\mathcal{O}} \circ R(s) = \ker_{\mathcal{O}} \circ M(s).$$

Proof. Since this proof is quite long, we divide it into smaller steps.

First step: Existence of $M(s)$ such that $R(s)M(s) = 0$.
If $p = q$ there is only a minor, the non zero determinant, that is a unit being $R(s)$ invertible on $\mathcal{O}$. Thus the kernel of $R(s)$ is trivial and is the image of a zero matrix.

Let us suppose that $p < q$. If $r(s)$ is any element in $\mathcal{O}^p R(s)$, we can build the matrix

$$\tilde{R}(s) = \begin{bmatrix} r(s) \\ R(s) \end{bmatrix}. \tag{7.12}$$

$\tilde{R}(s)$ has rank p, so every minor of order $p+1$ is zero and is a linear combination of $p+1$ elements of $r(s)$, the coefficients being $p+1$ minors of order p of $R(s)$.

More precisely: let $\rho \subseteq \mathbb{N}$ be a subset of $p+1$ elements of the set $\{1, 2, \ldots, q\}$; we write $\rho(i)$ to indicate the i-th element of ρ and suppose that ρ is ordered, i.e. $\rho(i) < \rho(j)$ whenever $1 \leq i < j \leq q$. Further we denote by

$$\bar{\rho}(i) = \rho \setminus \rho(i) \tag{7.13}$$

the ordered set with p elements that has the elements of ρ except $\rho(i)$ and let

$$n_\rho(i) \triangleq \begin{cases} 0 & \text{if } i \notin \rho \\ (-1)^k & \text{if } \rho(k) = i \end{cases} \tag{7.14}$$

that is to say: if $i \in \rho$ then $n_\rho(i)$ is equal to 1 when i occupies an even 'position' in ρ.

We know by basic combinatorics that there are exactly $d = \binom{q}{p+1}$ different sets ρ_j, so we can construct a matrix $M(s) \in \mathcal{A}^{q \times d}$ with elements $m_{ij}(s)$ defined as

$$m_{ij}(s) \triangleq n_{\rho_j}(i) R(s)_{\bar{\rho}_j(i)} \tag{7.15}$$

where $R(s)_\rho$ is a minor of $R(s)$, the determinant of the matrix formed by columns of $R(s)$ indexed by elements in the set ρ.

We see that if $r(s)$ in (7.12) has elements $r_i(s)$ and $M_j(s)$ is the j-th column of $M(s)$ then, remembering that by definition (7.14) $n_{\rho_j}(i) = 0$ when $i \notin \rho_j$,

$$0 = \tilde{R}(s)_{\rho_j} = \sum_{i \in \rho_j} r_i(s) n_{\rho_j}(i) R(s)_{\bar{\rho}_j(i)} = \sum_{i=1}^{q} r_i(s) n_{\rho_j}(i) R(s)_{\bar{\rho}_j(i)} = r(s) M_j(s).$$

Since $r(s)$ may be any row in $\mathcal{O}^p R(s)$, thus any row of $R(s)$, this proves that

$$R(s)M(s) = 0. \tag{7.16}$$

Second step: The matrix $M(s)$ has rank $q - p$.
Let $\rho = \{q-p+1, q-p+2, \dots, q-1, q\}$, the last p columns of $R(s)$ and let ρ_j, $j = 1, \dots, d$ be such that

$$\rho_j = \{j\} \cup \rho, \ \forall j \in \{1, 2, \dots, q-p\}.$$

By definition (7.13) $\bar{\rho}_j(j) = \rho$ thus by (7.15)

$$|m_{jj}(s)| = |R(s)_\rho|, \ \forall j \in \{1, 2, \dots, q-p\}.$$

Again the definition (7.15) of $m_{ij}(s)$ and (7.14) imply that

$$\forall i, j \in \{1, 2, \dots, q-p\}, i \neq j \ \Rightarrow \ n_{\rho_j}(i) = 0 \text{ therefore } m_{ij}(s) = 0$$

so the submatrix containing the first $q-p$ rows and columns of $M(s)$ is the identity matrix multiplied by the minor $R(s)_\rho$ of $R(s)$.

It is obvious, by the symmetric structure of the problem, that the set of minors of order $q-p$ of $M(s)$ contains every $(q-p)$–th power of the maximal minors of $R(s)$. This shows that $M(s)$ has at least rank $m \geq q-p$.

If we consider the Smith form (4.1) over $\mathcal{O}$ of $M(s)$, then

$$0 = R(s)M(s) = R(s)P(s)\bar{M}(s)Q(s) \ \Rightarrow \ \bar{R}(s)\bar{M}(s) = 0$$

where $\bar{R}(s) = R(s)P(s) \in \mathcal{O}^{p \times q}$ is still a full row rank matrix. $\bar{M}(s)$ has only m non zero diagonal entries, that is to say: the first m columns of $\bar{R}(s)$ are zero hence $q - m \geq p \Leftrightarrow m \leq q-p$. Together with the previously found relation $m = q-p$.

Third step: $\operatorname{im}_{\mathcal{O}} \circ R(s) = \ker_{\mathcal{O}} \circ M(s)$.
Equation (7.16) implies that $\operatorname{im}_{\mathcal{O}} \circ R(s) \subseteq \ker_{\mathcal{O}} \circ M(s)$; we want to prove that also the converse is true.

We know by theorem 4.12 that since $R(s)$ admits a generalized inverse, its $p \times p$ minors satisfy a Bézout equation. This implies that they have no common zero thus $R(s)$ is the presentation matrix of a spectrally controllable system and, by theorem 7.7, $\mathcal{O}^q/\mathcal{O}^p R(s)$ is torsion free.

By construction of $M(s)$ we know that if $r(s)M(s) = 0$ then matrix $\tilde{R}(s)$ defined in (7.12) does not have full rank:

$$\exists a(s) \in \mathcal{O},\ v(s) \in \mathcal{O}^p \text{ such that } a(s)r(s) = v(s)R(s).$$

So, if $\mathsf{r} = r(s) + \mathcal{O}^p R(s)$ is the equivalence class of $r(s)$ in $\mathcal{O}^q/\mathcal{O}^p R(s)$, then previous equations say that $a(s)\mathsf{r} = \mathsf{0}$ and by (torsion) freeness also $\mathsf{r} = \mathsf{0}$, i.e. $r(s) \in \operatorname{im}_{\mathcal{O}} \circ R(s)$.

Fourth step: $\ker_{\mathcal{O}} R(s)\circ = \operatorname{im}_{\mathcal{O}} M(s)\circ$.
As we pointed out, the maximal minor of $R(s)$ have no common zeros and their $(q–p)$-th power is contained in the set of minors of order $q-p$ of $M(s)$. This shows that neither these minors have common zeros; therefore, again by theorem 4.12, also $M(s)$ has a generalized inverse:

$$\exists G(s) \in \mathcal{O}^{d\times q} \text{ such that } M(s)G(s)M(s) = M(s).$$

Condition (7.16) may be restated also as

$$R(s)M(s) = 0 \;\Rightarrow\; \operatorname{im}_{\mathcal{O}} M(s)\circ \subseteq \ker_{\mathcal{O}} R(s)\circ. \tag{7.17}$$

On the other hand, using the result proved in the last step and lemma 5.8,

$$\operatorname{im}_{\mathcal{O}} \circ (G(s)M(s) - I) = \ker_{\mathcal{O}} \circ M(s) = \operatorname{im}_{\mathcal{O}} \circ R(s) = \ker_{\mathcal{O}} \circ (X(s)R(s) - I)$$

which implies that $(G(s)M(s) - I)(X(s)R(s) - I) = 0$ and therefore, 'changing side', $\operatorname{im}_{\mathcal{O}}(X(s)R(s) - I)\circ \subseteq \ker_{\mathcal{O}}(G(s)M(s) - I)\circ$. Thus, by (7.17)

$$\ker_{\mathcal{O}} R(s)\circ = \operatorname{im}_{\mathcal{O}}(X(s)R(s) - I)\circ \subseteq \ker_{\mathcal{O}}(G(s)M(s) - I)\circ = \operatorname{im}_{\mathcal{O}} M(s)\circ$$

and so $\operatorname{im}_{\mathcal{O}} M(s)\circ = \ker_{\mathcal{O}} R(s)\circ$. □

Proof of theorem 7.15. Let us suppose that $R(s)$ has rank r and that $F(s) \in \mathcal{A}^{r\times q}$ is a full row rank submatrix of $R(s)$ (i.e. every row of $F(s)$ is a row of $R(s)$). By theorem 4.12, the $r \times r$ minors of $R(s)$ satisfy a Bézout equation with coefficients in $\mathcal{O}$ or, in other words, have no common zeros; the minors of $F(s)$ are a subset of the minors of $R(s)$, hence also $F(s)$ has a generalized inverse and, by lemma 7.19, we can construct $M(s) \in \mathcal{A}^{q\times d}$ with $d = \binom{q}{r+1}$ such that

$$\ker_{\mathcal{O}} F(s)\circ = \operatorname{im}_{\mathcal{O}} M(s)\circ \text{ and } \operatorname{im}_{\mathcal{O}} \circ F(s) = \ker_{\mathcal{O}} \circ M(s).$$

It is easy to verify that $F(s) = Y(s)R(s)$ where $Y(s) \in \{0,1\}^{r\times p}$ only chooses the suitable rows of $R(s)$. This implies that

$$\ker_{\mathcal{O}} R(s)\circ \subseteq \ker_{\mathcal{O}} F(s)\circ \text{ and } \operatorname{im}_{\mathcal{O}} \circ R(s) \supseteq \operatorname{im}_{\mathcal{O}} \circ F(s).$$

Vice versa we know that given any row $r(s)$ of $R(s)$ there are a scalar $a(s) \in \mathcal{O}$ and a vector $z(s) \in \mathcal{O}^r$ such that

$$a(s)r(s) = z(s)F(s).$$

By definition 3.27, $F(s)$ is the presentation matrix of a spectrally controllable system; then, by theorem 7.7, $\mathcal{O}^q/\mathcal{O}^p R(s)$ is torsion free: that is to say that

$$r(s) = \tilde{z}(s)F(s) \text{ with } \tilde{z}(s) = \frac{z(s)}{a(s)} \in \mathcal{O}^r.$$

This proves that also $R(s) = Z(s)F(s)$ whence

$$\ker_{\mathcal{O}} R(s)\circ = \ker_{\mathcal{O}} F(s)\circ = \operatorname{im}_{\mathcal{O}} M(s)\circ \text{ and } \operatorname{im}_{\mathcal{O}} \circ R(s) = \operatorname{im}_{\mathcal{O}} \circ F(s) = \ker_{\mathcal{O}} \circ M(s).$$

□

Remark 7.20. *We note that theorem 7.15 gives a necessary and sufficient condition: actually we could start with a matrix $M(s) \in \mathcal{A}^{q\times d}$ with rank m that admits a generalized inverse over $\mathcal{O}$ and construct $R(s) \in \mathcal{A}^{p\times q}$ with rank $q - m$ which also admits a generalized inverse over $\mathcal{O}$ and satisfies relations (7.11). To prove this fact we only have apply theorem 7.15 to $M(s)^\top$.*

7.2.4 Behavioral controllability in $\mathcal{H}_m$

In this section we present results concerning more specifically delay–differential behaviors.

A condition that assures controllability of a behavior once we know that it is spectrally controllable, is the following:

Theorem 7.21. *Let $\mathcal{B} = \ker_{\mathcal{E}} R(s)$, $R(s) \in \mathcal{A}^{p\times q}$ of rank r, be spectrally controllable and suppose that a linear combination of the minors of order r with coefficients in $\mathcal{A}$ is equal to a polynomial, i.e. if $x_i(s)$ are the $r \times r$ minors of $R(s)$ and*

$$\exists h_i(s) \in \mathcal{A} : \sum_i x_i(s)h_i(s) \in \mathbb{R}[s].$$

Then $R(s)$ admits a generalized inverse $G(s) \in \mathcal{A}^{q\times p}$.

Proof. Let $X(s) \in \mathcal{A}^{t\times 1}$ be the row vector containing all the $r \times r$ minors of $R(s)$ and $H(s) \in \mathcal{A}^{1\times t}$ such that we have

$$a(s) \triangleq X(s)H(s) = \sum_i x_i(s)h_i(s) \in \mathbb{R}[s].$$

If s_0 is a zero of $a(s)$, then also $X(s_0)H(s_0) = 0$. Since $\mathcal{B}$ is spectrally controllable, $x_i(s)$ have no common zeros hence $X(s_0) \neq 0$.

By theorem 7.15 we know that there is a matrix $M(s) \in \mathcal{A}^{t\times d}$ such that $\ker_{\mathcal{O}} X(s)^\circ = \operatorname{im}_{\mathcal{O}} M(s)^\circ$ with constant $\operatorname{rank}_{\mathbb{C}} M(\lambda)$ for every $\lambda \in \mathbb{C}$, therefore $\ker_{\mathbb{C}} X(s_0)^\circ = \operatorname{im}_{\mathbb{C}} M(s_0)^\circ$: there exist a constant column $c \in \mathbb{C}^d$ such that $H(s_0) = M(s_0)c$.

Since $X(s)M(s)c = 0$ we can write

$$a(s) = X(s)(H(s)-M(s)c) \;\Rightarrow\; \tilde{a}(s) = \frac{a(s)}{s-s_0} = X(s)\frac{H(s)-M(s)c}{s-s_0} = X(s)\tilde{H}(s)$$

where $\tilde{a}(s)$ is a polynomial with lower degree than $a(s)$ and $\tilde{H}(s)$ is a vector in $\mathcal{A}$; iterating this procedure we get a Bézout equation for the minors of $R(s)$, with coefficients in $\mathcal{A}$: this permits to construct, by theorem 4.12 and lemma 5.8, a convolutional image representation of $\mathcal{B}$. □

Corollary 7.22. *Let $\mathcal{B} = \ker_{\mathcal{E}} R(s)$, $R(s) \in \mathcal{H}_m^{p\times q}$ of rank r, be spectrally controllable and suppose that a linear combination with coefficients in $\mathcal{H}_m$ of the minors of order r is equal to a polynomial, i.e. if $x_i(s)$ are the $r \times r$ minors of $R(s)$ and*

$$\exists h_i(s) \in \mathcal{H}_m : \sum_i x_i(s)h_i(s) \in \mathbb{R}[s]. \tag{7.18}$$

Then $R(s)$ admits a generalized inverse $G(s) \in \mathcal{H}_m^{q\times p}$.

Remark 7.23. *The above result is an algebraic-geometric criterion that, moreover, is computationally realizable by standard Gröbner bases algorithms: we must check equation (7.18) with $h_i(s) \in \mathcal{H}_m$. Actually it can be written using delay–differential polynomials multiplying by a common denominator*

$$a(s) = \sum_i \frac{r_i(s,\boldsymbol{z})}{\tilde{r}_i(s)\boldsymbol{z}^{\boldsymbol{n}_i}} \frac{h_i(s,\boldsymbol{z})}{\tilde{h}_i(s)\boldsymbol{z}^{\boldsymbol{k}_i}} \Leftrightarrow \tilde{a}(s)\boldsymbol{z}^{\boldsymbol{n}} = \sum_i r_i(s,\boldsymbol{z})h_i(s,\boldsymbol{z})$$

that employes only polynomials in $m+1$ variables. Anyway, the same result may be achieved directly in $\mathbb{R}[s,\boldsymbol{z},\boldsymbol{z}^{-1}]$ [BW93].

In general, a delay–differential behavior may admit a convolutional image representation; the following results show that once we know that a behavior admits a convolutional image representation, we can always find suitable delay–differential representations.

Lemma 7.24. *Let $R(s) \in \mathcal{A}^{p\times q}$ such that $\ker_{\mathcal{E}} R(s) = \operatorname{im}_{\mathcal{E}} N(s)$, $N(s) \in \mathcal{A}^{q\times d}$, and $\ker_{\mathcal{O}} R(s)^{\circ} = \operatorname{im}_{\mathcal{O}} M(s)^{\circ}$ for some $M(s) \in \mathcal{H}^{q\times d}$; then $\ker_{\mathcal{E}} R(s) = \operatorname{im}_{\mathcal{E}} M(s)$.*

Proof. We know that $R(s)N(s) = 0$ and that $M(s)^{\circ}$ generates over $\mathcal{O}$ every holomorphic vector belonging to $\ker_{\mathcal{O}} R(s)^{\circ}$, hence $N(s) = M(s)X(s)$ with $X(s)$ a holomorphic function. By theorem 4.14 there are $G(s) \in \mathcal{H}_m^{d\times q}$ and $a(s) \in \mathcal{H}_m$ such that $M(s)G(s)M(s) = a(s)M(s)$. Then

$$M(s)G(s)N(s) = M(s)G(s)M(s)X(s) = a(s)M(s)X(s) = a(s)N(s)$$

and, since $a(s)$ is surjective by theorem 4.29, $\operatorname{im}_{\mathcal{E}} N(s) = \operatorname{im}_{\mathcal{E}} N(s)a(s)$, therefore:

$$\operatorname{im}_{\mathcal{E}} N(s) = \operatorname{im}_{\mathcal{E}} M(s)G(s)N(s) \subseteq \operatorname{im}_{\mathcal{E}} M(s) \subseteq \ker_{\mathcal{E}} R(s) = \operatorname{im}_{\mathcal{E}} N(s).$$

□

Combining the above lemma and what is said in remark 7.16 about theorem 7.15, we obtain the following useful proposition.

Proposition 7.25. *If $\mathcal{B} = \ker_{\mathcal{E}} R(s)$, with $R(s) \in \mathcal{H}_m^{p\times q}$, has an image representation, then $\mathcal{B} = \operatorname{im}_{\mathcal{E}} M(s)$ with some $M(s) \in \mathcal{H}_m^{q\times d}$.*

So, once we know that an image representation exists, we can find one in $\mathcal{H}_m$; furthermore, there is a smaller matrix than the one showed by proposition 7.25:

Theorem 7.26. *Let $R(s) \in \mathcal{H}_m^{p\times q}$ with rank r; if $\mathcal{B} = \ker_{\mathcal{E}} R(s)$, has an image representation, then $\mathcal{B} = \operatorname{im}_{\mathcal{E}} M(s)$ with $M(s) \in \mathcal{H}_m^{q\times(q-r)}$ with full column rank.*

Proof. We know, by proposition 7.25, that there is an image representation with matrix $N(s) \in \mathcal{H}_m^{q\times d}$. If we denote by $\mathcal{K}_m$ the field of fractions of elements in $\mathcal{H}_m$ (which coincides with the field of fractions of exponential polynomials), $R(s)$ admits a Smith form, therefore lemma 4.4 states that there is an injective matrix $\tilde{M}(s) \in \mathcal{K}_m^{q\times(q-r)}$ such that

$$\ker_{\mathcal{K}_m} R(s)\circ = \operatorname{im}_{\mathcal{K}_m} \tilde{M}(s)\circ.$$

So, since $R(s)N(s) = 0$, we have also $N(s) = \tilde{M}(s)\tilde{X}(s)$ with some matrix $\tilde{X}(s) \in \mathcal{K}_m^{(q-r)\times d}$ and collecting the common denominator $a(s)$, $a(s)N(s) = M(s)X(s)$, with $M(s) \in \mathcal{H}_m^{q\times(q-r)}$ and $X(s) \in \mathcal{H}_m^{(q-r)\times d}$.

By theorem 4.29 the polynomial $a(s)$ is surjective on $\mathcal{E}$; therefore

$$\operatorname{im}_{\mathcal{E}} N(s) = \operatorname{im}_{\mathcal{E}} a(s)N(s)$$

and also

$$\ker_{\mathcal{E}} R(s) = \operatorname{im}_{\mathcal{E}} N(s) = \operatorname{im}_{\mathcal{E}} a(s)N(s) = \operatorname{im}_{\mathcal{E}} M(s)X(s) \subseteq \operatorname{im}_{\mathcal{E}} M(s) \subseteq \ker_{\mathcal{E}} R(s) \tag{7.19}$$

since $R(s)\tilde{M}(s) = 0$: this proves that $\ker_{\mathcal{E}} R(s) = \operatorname{im}_{\mathcal{E}} M(s)$. □

7.2.5 The single input case

Theorem 6.6 states among other facts, that in $\mathcal{H}_1$ a behavior $\mathcal{B} = \ker_{\mathcal{E}} R(s)$ admits an image representation if and only if $R(s)$ admits a generalized inverse. We show that this holds in $\mathcal{H}_m$ in a particular case.

Definition 7.27. *An* ***input/output representation*** *of a behavior*[1] $\mathcal{B}$ *is a partition of its variables $w(t)$ in input variables $u(t)$ and output variables $y(t)$, such that the input is* ***free*** *(i.e. for every $u(t)$ there is a $y(t)$ such that the trajectory $w(t)$ consisting of $y(t)$ and $u(t)$ is in $\mathcal{B}$) and* ***maximal*** *(i.e. once $u(t)$ has been fixed, $y(t)$ does not contain other free variables).*

If we know that $\mathcal{B} = \ker_{\mathcal{E}} R(s)$, where $R(s) \in \mathcal{A}^{p\times q}$ with rank r, after a suitable permutation of its columns we can partition $w^\top = [y^\top \ u^\top]$ and $R(s)$ consinstently as $R(s) = [P(s) \ -Q(s)]$ such that $P(s)$ is a full column rank matrix, i.e. there is a matrix $Y(s) \in \mathcal{A}^{r\times p}$ and a scalar $a(s) \in \mathcal{A}$ such that $Y(s)P(s) = a(s)I_r$ and the behavioral equations are

$$w(t) \in \mathcal{B} \Leftrightarrow R(s)w = 0 \Leftrightarrow P(s)y = Q(s)u.$$

If we suppose that amongst all possible choices of $P(s)$ there is one such that $a(s)$ is surjective on $\mathcal{E}$ (and this always happens if $R(s) \in \mathcal{H}_m^{p\times q}$), we can obtain an input/output representation:

$$\forall u(t) \in \mathcal{E}^{q-r} \ \exists y(t) \in \mathcal{E}^r : \ P(s)y = Q(s)u.$$

Indeed we know that $\exists v(t) \in \mathcal{E}^p$ such that $a(s)v = Q(s)u$, so

$$y = Y(s)v \Rightarrow P(s)y = P(s)Y(s)v = a(s)v = Q(s)u;$$

moreover, fixing $u(t) = 0$ then $P(s)y = 0$ implies that $Y(s)P(s)y = a(s)y = 0$ hence $y(t)$ is not free: every component is a solution of the same delay–differential equation. Note that the number of inputs is equal to $q - r$.

The following theorem, valid with much more generality for differential behaviors, is still valid for delay–differential behaviors with one input:

Theorem 7.28. *Suppose that $\mathcal{B} = \ker_{\mathcal{E}} R(s)$, with $R(s) \in \mathcal{H}_m^{p\times q}$ with rank $q-1$ (i.e. $\mathcal{B}$ is a single–input behavior); then*

$\mathcal{B}$ admits image representation $\Leftrightarrow \exists X(s) \in \mathcal{A}^{q\times p}$ such that $R(s)X(s)R(s) = R(s)$.

Proof. If $R(s)X(s)R(s) = R(s)$, we can use theorem 7.12. On the converse,

[1] *We are disregarding properness issues for the sake of simplicity: see for more details [PW97, ch. 3.3].*

remark 7.16 and the proof of lemma 7.19 state that there is a column $M(s) \in \mathcal{H}_m^{q\times 1}$ made with minors of $R(s)$ such that

$$\ker_{\mathcal{O}} R(s)\circ = \operatorname{im}_{\mathcal{O}} M(s)\circ.$$

If we prove that $1 \in \operatorname{im}_{\mathcal{A}} \circ M(s)$ we have proven that the minors of $R(s)$ satisfy a Bézout equation over $\mathcal{A}$ and, by theorem 4.12, $R(s)$ admits a generalized inverse.

Since the behavior admits an image representation, say $\mathcal{B} = \operatorname{im}_{\mathcal{E}} N(s)$ with $N(s) \in \mathcal{A}^{q\times d}$, then $R(s)N(s) = 0$ thus each row of $N(s)$ is in $\ker_{\mathcal{O}} R(s)\circ$: we have that

$$\exists Y(s) \in \mathcal{O}^{1\times d} :\ M(s)Y(s) = N(s) \ \Rightarrow\ m_i(s)y_j(s) = n_{ij}(s).$$

By corollary 5.5 $n_{ij}(s)/m_i(s) = y_j(s) \in \mathcal{O}$ is a Paley–Wiener function, so $Y(s) \in \mathcal{A}^{1\times d}$; now since $M(s)Y(s) = N(s)$ where every matrix is an operator, $\operatorname{im}_{\mathcal{E}} N(s) \subseteq \operatorname{im}_{\mathcal{E}} M(s)$. Therefore

$$\ker_{\mathcal{E}} R(s) = \operatorname{im}_{\mathcal{E}} N(s) \subseteq \operatorname{im}_{\mathcal{E}} M(s) \subseteq \ker_{\mathcal{E}} R(s)$$

so $\operatorname{im}_{\mathcal{E}} M(s)$ is closed and, by lemmas 5.14 and 5.16, $\operatorname{im}_{\mathcal{A}} \circ M(s)$ is closed and equal to $(\ker_{\mathcal{E}} M(s))^{\perp}$.

We know that the minors of $R(s)$, that are the elements of $M(s)$, have no common zeros, thus by theorem 4.26 $\ker_{\mathcal{E}} M(s) = \{0\}$: so $1 \in \mathcal{A} = \operatorname{im}_{\mathcal{A}} \circ M(s)$. □

7.2.6 Behavior closure

We have already introduced two statements that are equivalent to behavioral controllability for differential systems: existence of image representation and spectral controllability (the proof of this equivalence for delay–differential behaviors with one delay is the subject of [RW97]). There is another well-known way to characterize controllable behaviors (see e.g. [Wil91, p. 266] for discrete-time systems): if we take only the set of trajectories of a differential behavior $\mathcal{B}$ that are zero for large $|t|$ and consider its closure in $\mathcal{E}$, this set is equal to $\mathcal{B}$ itself if and only it is controllable.

We shall prove that the aforementioned property is equivalent to spectral controllability for the class of convolutional behaviors that admit a full row rank kernel representation; as a corollary we obtain an extension of theorem 6.6 for delay–differential behaviors with commensurate delays.

Definition 7.29. *Given a behavior $\mathcal{B}$ we let $\mathcal{B}_{cs}$ be the subset of $\mathcal{B}$ containing only functions with compact support: if we recall the definition of $\mathcal{D}$ (the set of smooth functions with compact support introduced at page 14) and suppose that $\mathcal{B} \subseteq \mathcal{E}^q$ then*

$$\mathcal{B}_{cs} = \mathcal{B} \cap \mathcal{D}^q = \{w(t) \in \mathcal{B} \text{ such that } w(t) \text{ has compact support}\}.$$

We have seen that the existence of an image representation is a rather strong property for convolutional behaviors; the main problem is the following: while kernels are always closed sets, images are not closed in general. So we can weaken the former property and ask for the equality of a given behavior and the *closure* of some behavior in image representation, fact that has still interesting consequences as next proposition will show.

Remark 7.30. *We note that if $M(s) \in \mathcal{A}^{q\times d}$; then*

$$\operatorname{im}_{\mathcal{E}} M(s) \subseteq \overline{\operatorname{im}_{\mathcal{D}} M(s)}. \tag{7.20}$$

Indeed, one of the equivalent definitions of continuity (see appendix B.1), is the following: a function $f : \mathcal{V} \to \mathcal{W}$ is continuous if and only if for every subspace $\mathcal{U} \subset \mathcal{V}$, $\operatorname{im}_{\overline{\mathcal{U}}} f \subseteq \overline{\operatorname{im}_{\mathcal{U}} f}$ (see [Dug66, thm. 8.3]).

Since $\mathcal{D}$ is dense in $\mathcal{E}$ (see e.g. [Tre67, cor. 1, p. 159]) then

$$\operatorname{im}_{\mathcal{E}} M(s) = \operatorname{im}_{\overline{\mathcal{D}}} M(s) \subseteq \overline{\operatorname{im}_{\mathcal{D}} M(s)}.$$

Proposition 7.31. *Suppose that the behavior $\mathcal{B}$ admits a kernel representation, $\mathcal{B} = \ker_{\mathcal{E}} R(s)$ with $R(s) \in \mathcal{A}^{p\times q}$, and let $M(s) \in \mathcal{A}^{q\times d}$. Then*

$$\mathcal{B} = \overline{\operatorname{im}_{\mathcal{E}} M(s)} \;\Rightarrow\; \mathcal{B} = \overline{\mathcal{B}_{cs}}.$$

Proof. Since $\mathcal{B}$ is closed and

$$\mathcal{B}_{\mathrm{cs}} = \mathcal{B} \cap \mathcal{D}^q = \ker_{\mathcal{E}} R(s) \cap \mathcal{D}^q = \ker_{\mathcal{D}} R(s)$$

then $\overline{\mathcal{B}_{\mathrm{cs}}} \subseteq \mathcal{B}$.

To prove the converse, we note first that $\operatorname{im}_{\mathcal{D}} M(s) \subseteq \ker_{\mathcal{D}} R(s)$; then, by equation (7.20), $\operatorname{im}_{\mathcal{E}} M(s) \subseteq \overline{\operatorname{im}_{\mathcal{D}} M(s)}$ and so

$$\mathcal{B} = \overline{\operatorname{im}_{\mathcal{E}} M(s)} \subseteq \overline{\operatorname{im}_{\mathcal{D}} M(s)} \subseteq \overline{\ker_{\mathcal{D}} R(s)} = \overline{\mathcal{B}_{\mathrm{cs}}}.$$

□

Next theorem shows that $\mathcal{B} = \overline{\mathcal{B}_{\mathrm{cs}}}$ always implies spectral controllability.

Theorem 7.32. *Let $\mathcal{B} = \ker_{\mathcal{E}} R(s)$ with $R(s) \in \mathcal{A}^{p\times q}$. Then*

$$\mathcal{B} = \overline{\mathcal{B}_{cs}} \;\Rightarrow\; \mathcal{B} \text{ is spectrally controllable.}$$

Proof. This proof is similar to the proof of theorem 7.11: we omit some details that may be found therein.

We know that if $R(s)$ has rank r, there are two matrices $U_1(s) \in \mathcal{O}^{r\times p}$ and $U_2(s) \in \mathcal{O}^{p-r\times p}$ such that

$$\ker_{\mathcal{O}} \circ R(s) = \operatorname{im}_{\mathcal{O}} \circ U_2(s), \tag{7.21}$$

$$\operatorname{im}_{\mathbb{C}} \circ U_1(s_0) \oplus \operatorname{im}_{\mathbb{C}} \circ U_2(s_0) = \mathbb{C}^p \;\forall s_0 \in \mathbb{C} \tag{7.22}$$

and that

$$R(s) \text{ does not loose rank at } s_0 \;\Leftrightarrow\; \ker_{\mathbb{C}} \circ R(s_0) \cap \operatorname{im}_{\mathbb{C}} \circ U_1(s_0) = \{0\}\,. \tag{7.23}$$

We remind that $\mathcal{B}$ is spectrally controllable if and only if the above conditions are satisfied for every $s_0 \in \mathbb{C}$.

So, given an $s_0 \in \mathbb{C}$, let us take $c \in \ker_{\mathbb{C}} \circ R(s_0) \cap \operatorname{im}_{\mathbb{C}} \circ U_1(s_0)$ and define

$$h(s) \triangleq \frac{1}{s - s_0} c R(s) \in \mathcal{A}^p.$$

As pointed out in remark 5.6, $f = h(s)w$ if and only if $(s - s_0)f = cR(s)w$: if $w(t) \in \mathcal{B} = \ker_{\mathcal{E}} R(s)$ then $f(t) = ke^{s_0 t}$ and hence the set $h(\mathcal{B})$ contains only exponentials.

If we take $w(t) \in \mathcal{B}_{\mathrm{cs}} \subseteq \mathcal{B}$ and $f = h(s)w$, then, for large $|\tau|$, $f(\tau) = 0$ since $h(s)$ is the Laplace transform of a distribution with compact support; being $f(t)$ an exponential, it must be zero for every $t \in \mathbb{R}$. Employing the hypothesis on $\mathcal{B}$, by equation (7.20), we can write that

$$h(\mathcal{B}) = h(\overline{\mathcal{B}_{\rm cs}}) \subseteq \overline{h(\mathcal{B}_{\rm cs})} = \overline{\{0\}} = \{0\} \;\Rightarrow\; \ker_{\mathcal{E}} R(s) \subseteq \ker_{\mathcal{E}} h(s).$$

This implies, by theorem 4.25, that $h(s) = a(s)R(s)$ for some $a(s) \in \mathcal{O}^q$. So, by definition of $h(s)$ and equation (7.21),

$$\exists b(s) \in \mathcal{O}^{p-r} :\; c - (s - s_0)a(s) = b(s)U_2(s) \;\Rightarrow\; c = b(s_0)U_2(s_0);$$

since $c \in \operatorname{im}_{\mathbb{C}} U_1(s_0)$ by hypothesis and the sum in (7.22) is direct, $c = 0$ and so $R(s_0)$ does not loose rank for any $s_0 \in \mathbb{C}$. □

Dealing with delay–differential systems with one delay it is very easy to prove that the conditions we have introduced in this section are equivalent to behavioral controllability:

Corollary 7.33. *Given the delay–differential behavior $\mathcal{B} = \ker_{\mathcal{E}} R(s)$, $R(s) \in \mathcal{H}_1^{p\times q}$ with commensurate delays, controllability, and the conditions listed in theorem 6.6, are equivalent to the following ones:*

- $\mathcal{B} = \overline{\mathcal{B}_{cs}}$;
- *there is a matrix $M(s) \in \mathcal{H}_1^{q\times d}$ such that $\mathcal{B} = \overline{\operatorname{im}_{\mathcal{E}} M(s)}$.*

Proof. Theorem 6.6 states that if $\mathcal{B}$ is spectrally controllable, it admits an image representation: $\mathcal{B} = \operatorname{im}_{\mathcal{E}} M(s)$. Since $\mathcal{B}$ is the kernel of an operator, the image of $M(s)$ is closed: the hypothesis of proposition 7.31 are satisfied and this, together with theorem 7.32, ends the proof. □

Convolutional behaviors do not share with delay–differential behaviors with one delay such elegant theorems but at least in a particular case, when $\mathcal{B} = \ker_{\mathcal{E}} R(s)$ and $R(s)$ is a full row rank matrix, it is possible to reverse the implications of proposition 7.31 and theorem 7.32:

Theorem 7.34. *Let $R(s) \in \mathcal{A}^{p\times q}$ with full row rank and $\mathcal{B} = \ker_{\mathcal{E}} R(s)$. Then*

$$\mathcal{B} \text{ is spectrally controllable } \;\Rightarrow\; \exists M(s) \in \mathcal{A}^{q\times d} \text{ such that } \mathcal{B} = \overline{\operatorname{im}_{\mathcal{E}} M(s)}.$$

Remark 7.35. *Since this theorem is based on theorem 7.15, we may substitute any other operator ring of holomorphic functions for the ring $\mathcal{A}$ within the statement, as we noticed in remark 7.16. Therefore if e.g. $R(s)$ is a delay–differential polynomial matrix, the matrix $M(s)$ we can construct belongs to the same class.*

Proof of theorem 7.34. By theorem 7.15 we know that there is a matrix $M(s) \in \mathcal{A}^{q\times d}$ such that

$$\mathrm{im}_{\mathcal{O}} \circ R(s) = \ker_{\mathcal{O}} \circ M(s); \tag{7.24}$$

since this implies that $R(s)M(s) = 0$, we have that also

$$\mathrm{im}_{\mathcal{A}} \circ R(s) \subseteq \ker_{\mathcal{A}} \circ M(s). \tag{7.25}$$

We want to prove that the converse inclusion holds for the closure of these sets.

Let $x(s) \in \ker_{\mathcal{A}} \circ M(s)$; by (7.24) there is a $y(s) \in \mathcal{O}^p$ such that

$$x(s) = y(s)R(s). \tag{7.26}$$

Let us suppose, without loss of generality, that $R(s) = [R_1(s)\ R_2(s)]$ where $R_1(s)$ is a square full rank submatrix of $R(s)$. Let $x(s)$ be partitioned in the same way: we obtain

$$x_1(s){=}y(s)R_1(s) \Rightarrow x_1(s)\mathrm{adj}\, R_1(s){=}y(s)R_1(s)\mathrm{adj}\, R_1(s){=}y(s)\det R_1(s). \tag{7.27}$$

We need now the following important result: if we let

$$H(r) \triangleq \{a(s) \in \mathcal{A} :\ b(s) \in \mathcal{A} \cap \mathcal{O}r(s) \Rightarrow a(s)b(s) \in \mathcal{A}r(s)\} \tag{7.28}$$

then, as stated in [Mal56, p. 308], for any non zero Paley–Wiener function

$$\forall r(s) \in \mathcal{A},\ r(s) \neq 0,\ \text{we have that } \overline{H(r)} = \mathcal{A}. \tag{7.29}$$

In other words: we know that if $b(s)$ and $r(s)$ are Paley–Wiener functions and $c(s) = b(s)/r(s)$ is holomorphic then, by proposition 5.4, $c(s) \in \mathcal{A}$ if and only if $\mathcal{A}r(s)$ is closed. However, even if $c(s) \notin \mathcal{A}$, we may find a function $a(s) \in \mathcal{A}$ such that $a(s)c(s) \in \mathcal{A}$. Malgrange's theorem, i.e. equation (7.29), tells us that the set $H(r)$ of functions $a(s)$ that map, by multiplication, *every* holomorphic fraction with denominator $r(s)$ into $\mathcal{A}$, is not only non trivial, but also dense in $\mathcal{A}$.

So, if we set $r(s) \triangleq \det R_1(s)$ and $\bar{x}_1(s) = x_1(s) \operatorname{adj} R_1(s) \in \mathcal{A}^p$, then in particular equation 7.27 becomes

$$\bar{x}_1(s) = y(s) r(s) \tag{7.30}$$

therefore $\bar{x}_1(s) \in \mathcal{A}^p \cap \mathcal{O}^p r(s)$; so, by definition (7.28) of $H(r)$,

$$\text{if } a(s) \in H(r) \text{ then } a(s)\bar{x}_1(s) \in \mathcal{A}^p r(s). \tag{7.31}$$

Actually $H(r)$ was defined only for scalar equations, but we remark that $H(r)$ depends only on $r(s)$, hence equation (7.31) is true componentwise.

Since $H(r)$ is dense in $\mathcal{A}$ by (7.29), there must be a sequence $a_n(s) \in H(r)$ converging to $1 \in \mathcal{A}$; from equation (7.31) we obtain that

$$\exists h_n(s) \in \mathcal{A}^p : \; a_n(s)\bar{x}_1(s) = h_n(s) r(s)$$

and by equation (7.30)

$$a_n(s)\bar{x}_1(s) = a_n(s) y(s) r(s) = h_n(s) r(s) \;\Rightarrow\; a_n(s) y(s) = h_n(s).$$

If we multiply by $a_n(s)$ both members of equation (7.26) we obtain

$$a_n(s) x(s) = a_n(s) y(s) R(s) = h_n(s) R(s) \in \mathcal{A}^p R(s).$$

Therefore $a_n(s)x(s)$ is a sequence in $\mathcal{A}^p R(s)$ and its limit lies in the closure; since $a_n(s)$ converges to 1, $a_n(s)x(s) \to x(s) \in \overline{\mathcal{A}^p R(s)}$. In other words:

$$\ker_{\mathcal{A}} \circ M(s) \subseteq \overline{\operatorname{im}_{\mathcal{A}} \circ R(s)}.$$

Considering also equation (7.25), if we take the orthogonals, we obtain

$$\left(\overline{\operatorname{im}_{\mathcal{A}} \circ R(s)}\right)^{\perp} \subseteq (\ker_{\mathcal{A}} \circ M(s))^{\perp} \subseteq (\operatorname{im}_{\mathcal{A}} \circ R(s))^{\perp}.$$

by proposition 5.13; the same proposition, together with lemma 5.14, states that

$$\left(\overline{\operatorname{im}_{\mathcal{A}} \circ R(s)}\right)^{\perp} = (\operatorname{im}_{\mathcal{A}} \circ R(s))^{\perp} = \ker_{\mathcal{E}} R(s),$$

thus, since also $(\ker_{\mathcal{A}} \circ M(s))^{\perp} = \overline{\operatorname{im}_{\mathcal{E}} M(s)}$, we get the desired result: $\ker_{\mathcal{E}} R(s) = \overline{\operatorname{im}_{\mathcal{E}} M(s)}$. □

Furthermore there is another implication that holds for convolutional behaviors: if $\mathcal{B}$ is controllable then $\mathcal{B} = \overline{\mathcal{B}_{\text{cs}}}$.

By the way, definition 3.11 of controllability is not the only possible: as a matter of fact, researchers in the behavioral framework proposed different types of controllability that are still based on trajectories like the one we have considered throughout this thesis, (see e.g. [ZM96] where relations between various concepts are showed). Next proposition actually employs, as intermediate step, the notion of **symmetric controllability**: a behavior is symmetric controllable if it is possible to steer every trajectory to zero both in the future and in the past.

Proposition 7.36. *Suppose that the behavior $\mathcal{B}$ is controllable. Then $\mathcal{B} = \overline{\mathcal{B}_{\text{cs}}}$.*

Proof. First of all we prove that a controllable behavior is symmetric controllable: we know that every trajectory may be steered to zero; we have to show that for every $w(t) \in \mathcal{B}$ we can find a $\tilde{w}(t) \in \mathcal{B}$ such that

$$\exists T > 0 \text{ such that } \tilde{w}(t) = 0 \; \forall t < -T \text{ and } \tilde{w}(t) = w(t) \; \forall t > 0.$$

This is rather simple: by definition 3.11, given $w_1(t) = 0$ and $w_2(t) = w(t)$, there is a $\tau > 0$ and a $\bar{w}(t)$ that is zero for $t \leq 0$ and coincides with $\sigma_\tau w(t)$ for $t \geq \tau$. If we take $T = \tau$, then $\tilde{w}(t) = \sigma_{-\tau}\bar{w}(t)$ is the desired function.

Next we suppose that the behavior is symmetric controllable and $w(t) \in \mathcal{B}$; if we consider an increasing sequence of compact intervals

$$[t_i, \tau_i] = K_i \subset K_{i+1} \text{ such that } \cup K_i = \mathbb{R},$$

then for every i we can find a trajectory $u_i(t) \in \mathcal{B}$ that is equal to $w(t)$ for $t \geq t_i$ and zero in the 'past'; we can also find a trajectory $v_i(t) \in \mathcal{B}$ that is equal to $u_i(t)$ for $t \leq \tau_i$ and zero in the 'future'.

Clearly

$$v_i(t) \in \mathcal{B}_{\text{cs}} \text{ and } v(t) = w(t) \; \forall t \in K_i;$$

the sequence $v_n(t)$ converges to $w(t)$ on compacts, i.e. in the topology of $\mathcal{E}$, therefore $w(t)$ is a limit point of $\mathcal{B}_{\text{cs}}$. So $\mathcal{B} \subseteq \overline{\mathcal{B}_{\text{cs}}}$.

The converse inclusion, that completes the proof, trivially holds always. □

In conclusion we note that slight variations of lemma 7.24, proposition 7.25 and theorem 7.26 hold in this context too: as regards lemma 7.24 we remark that if

$$\ker_{\mathcal{E}} R(s) = \overline{\operatorname{im}_{\mathcal{E}} N(s)} \text{ with } R(s) \in \mathcal{A}^{p\times q} \text{ and } N(s) \in \mathcal{A}^{q\times d},$$

then clearly $\ker_{\mathcal{E}} R(s) \supseteq \operatorname{im}_{\mathcal{E}} N(s)$ therefore $R(s)N(s) = 0$. If there is an $M(s) \in \mathcal{H}^{q\times d}$ such that $\ker_{\mathcal{O}} R(s)^{\circ} = \operatorname{im}_{\mathcal{O}} M(s)^{\circ}$ then, as the proof of the lemma 7.24 shows, $\operatorname{im}_{\mathcal{E}} N(s) = \operatorname{im}_{\mathcal{E}} M(s)$ therefore $\ker_{\mathcal{E}} R(s) = \overline{\operatorname{im}_{\mathcal{E}} M(s)}$.

The matrix $M(s)$ always exists whenever $\mathcal{B}$ is a delay–differential behavior by theorem 7.15 and remark 7.16; we record in the following proposition the most important consequence.

Proposition 7.37. *Let $R(s) \in \mathcal{H}_m^{p\times q}$ with rank r and $\mathcal{B} = \ker_{\mathcal{E}} R(s)$; if $\mathcal{B} = \overline{\operatorname{im}_{\mathcal{E}} N(s)}$ with $N(s) \in \mathcal{A}^{q\times d}$ then $\mathcal{B} = \overline{\operatorname{im}_{\mathcal{E}} M(s)}$ with $M(s) \in \mathcal{H}_m^{q\times(q-r)}$ with full column rank.*

Proof. We have just proved that there is a matrix $L(s) \in \mathcal{H}_m^{q\times d}$ such that $\mathcal{B} = \overline{\operatorname{im}_{\mathcal{E}} L(s)}$. As proof of theorem 7.26 shows, we can find an element $a(s) \in \mathcal{H}_m$ and matrices $M(s) \in \mathcal{H}_m^{q\times(q-r)}$ with full column rank and $X(s) \in \mathcal{H}_m^{(q-r)\times d}$ such that $a(s)L(s) = M(s)X(s)$. We obtain

$$\operatorname{im}_{\mathcal{E}} L(s) = \operatorname{im}_{\mathcal{E}} a(s)L(s) = \operatorname{im}_{\mathcal{E}} M(s)X(s) \subseteq \operatorname{im}_{\mathcal{E}} M(s) \subseteq \ker_{\mathcal{E}} R(s)$$

and, closing with respect to the topology of $\mathcal{E}$,

$$\ker_{\mathcal{E}} R(s) = \overline{\operatorname{im}_{\mathcal{E}} L(s)} \subseteq \overline{\operatorname{im}_{\mathcal{E}} M(s)} \subseteq \ker_{\mathcal{E}} R(s).$$

□

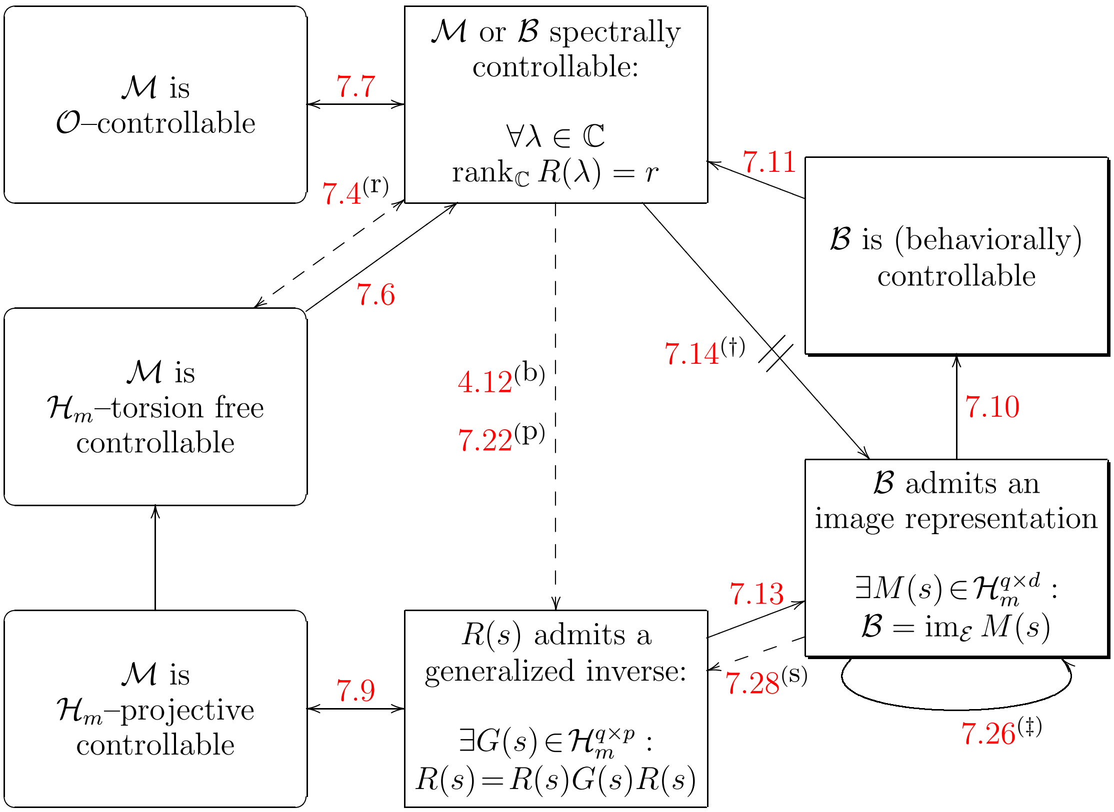
7.3 Summarizing pictures

Since for delay differential systems with non commensurate delays, or more generically, for convolutional systems there is no simple and unifying theorem on controllabilities like 6.6, we show three pictorial representations of the relations we have found between various conditions related to controllability.

The first two sum up, roughly speaking, the content of sections 7.2.1–7.2.5 regarding delay–differential systems (in $\mathcal{H}_m$) and, respectively, convolutional systems (in $\mathcal{A}$) both in the behavioral and in the Fliess' approach.

The third one is concerned only with behavioral notions and presents the results of section 7.2.6.

The following picture holds for delay–differential behaviors, i.e. we let $R(s) \in \mathcal{H}_m^{p\times q}$ with rank r, so $\mathcal{B} = \ker_{\mathcal{E}} R(s)$ and $\mathcal{M} = \mathrm{coker}_{\mathcal{H}_m} \circ R(s)$.



Legenda

Rounded boxes refer to the module and shaded to the behavioral approach. Numbers on arrows refer to propositions; continuous arrows ⟶ always hold; dashed arrows - - → need additional hypotheses indicated by small parenthesized letters (in the following list we denote by $x_i(s)$ the $r \times r$ minors of $R(s)$):

(b): the minors $x_i(s)$ of $R(s)$ satisfy a Bézout equation over $\mathcal{H}_m$;

(p): the ideal generated by the minors $x_i(s)$ of $R(s)$ in $\mathcal{H}_m$ contains a polynomial;

(r): $R(s)$ has full row rank (i.e. $p = r$);

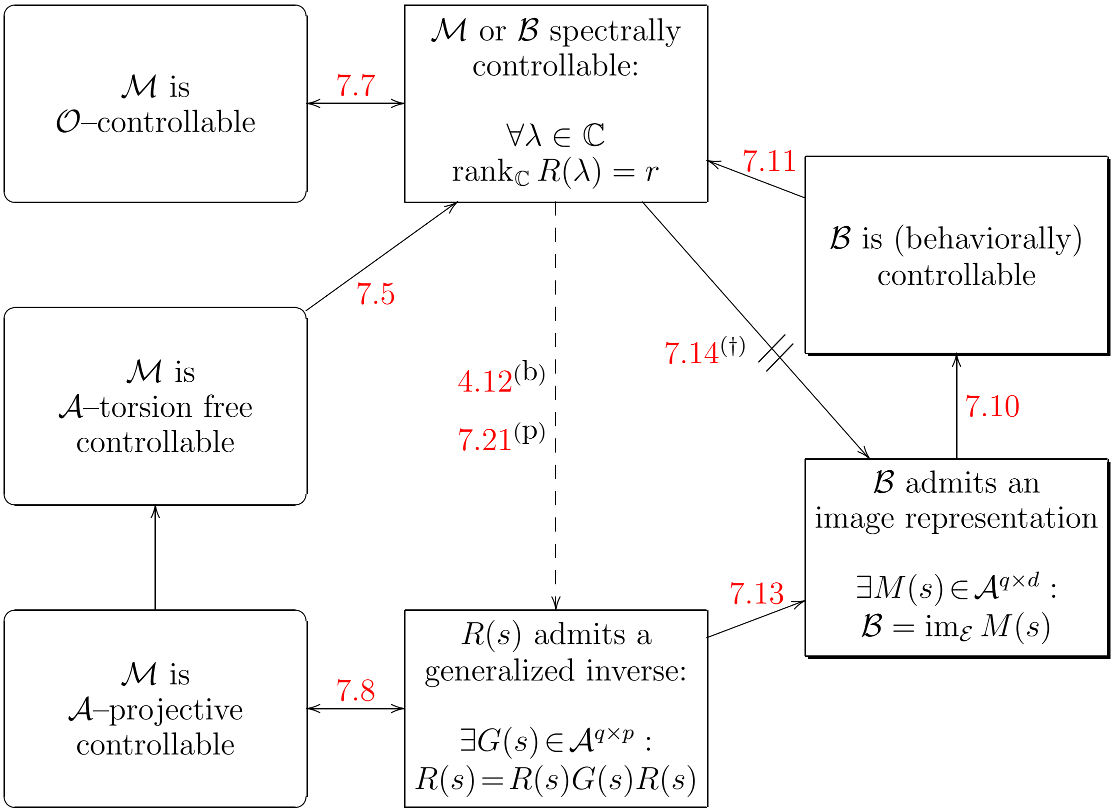
(s): the system has a single input.

Moreover:

(†): this is a counterexample;

(‡): we can always find a full column rank $M(s) \in \mathcal{H}_m^{q\times(q-r)}$.

This picture holds for convolutional systems: we let $R(s) \in \mathcal{A}^{p\times q}$ with rank r, so $\mathcal{B} = \ker_{\mathcal{E}} R(s)$ and $\mathcal{M} = \mathrm{coker}_{\mathcal{A}} \circ R(s)$.



Legenda

Rounded boxes refer to the module and shaded to the behavioral approach.
Numbers on arrows refer to propositions; continuous arrows $\longrightarrow$ always hold; dashed arrows $-\,-\!\!\succ$ need additional hypotheses indicated by small parenthesized letters (in the following list we denote by $x_i(s)$ the $r\times r$ minors of $R(s)$):

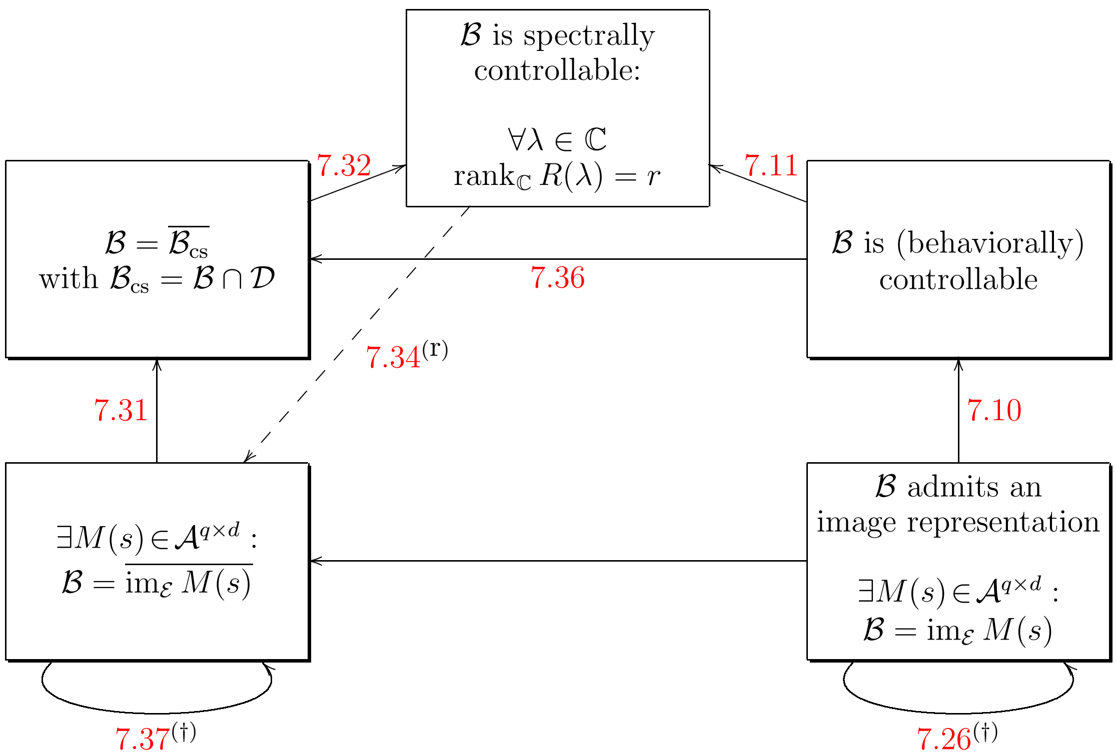
$^{(b)}$: the minors $x_i(s)$ of $R(s)$ satisfy a Bézout equation over $\mathcal{A}$;

$^{(p)}$: the ideal generated by the minors $x_i(s)$ of $R(s)$ in $\mathcal{A}$ contains a polynomial.

Moreover:

$^{(\dagger)}$: this is a counterexample.

This picture holds for convolutional and delay–differential systems: we let generically $R(s) \in \mathcal{A}^{p\times q}$ with rank r and $\mathcal{B} = \ker_{\mathcal{E}} R(s)$.



Legenda

Numbers on arrows refer to propositions; continuous arrows $\longrightarrow$ always hold; the dashed arrow $\dashrightarrow$ needs an additional hypothesis indicated by the small parenthesized letter:

(r): $R(s)$ has full row rank (i.e. $p = r$).

Moreover:

(†): if $\mathcal{B}$ is a delay–differential behavior, we can always find a full column rank $M(s) \in \mathcal{H}_m^{q\times(q-r)}$.

Appendix A

Basic Algebra

Many concepts of linear algebra were used without being defined throughout the thesis. This appendix offers a brief *glossary* of the necessary concepts and results.

A.1 Algebraic structures

A **semigroup** $(S, \circ)$ is a set S equipped with a mapping, or **operation**, $\circ : S^2 \to S$ that is **associative**, $x \circ (y \circ z) = (x \circ y) \circ z$. A **monoid** $(M, \circ)$ is a semigroup and admits a **unit** $e \in M$ such that $e \circ x = x \circ e = x \ \forall x \in M$. A **group** $(G, \circ)$ is a monoid such that there is an **inverse** for every element: $\forall x \in G \ \exists y \in G$ such that $x \circ y = y \circ x = e$. A subgroup of $(G, \circ)$ is a subset $H \subseteq G$ such that $(H, \circ)$ is a group.

The operation is **commutative** if $x \circ y = y \circ x$. Groups that are commutative are also called **Abelian**.

We will use mainly a 'multiplicative' operation ($x \circ y = xy$, the unit is 1 and the inverse of x is x^{-1}) or sometimes 'additive' operation ($x \circ y = x + y$, the unit is 0 and the inverse of x is $-x$).

A group **homomorphism** is a mapping $f : G_1 \to G_2$ such that $f(xy) = f(x)f(y)$. The **kernel** of a homomorphism is the subgroup $K \subseteq G_1$ such that $f(x) = 0$; it is denoted by $\ker f$. The **image** of a homomorphism is the subgroup $R \subseteq G_2$ such that $\forall y \in R$ there is an $x \in G_1$ and $f(x) = y$; it is denoted by $\operatorname{im} f$. A homomorphism is **injective** if its kernel is $\ker f = \{0\}$; it is surjective if $\operatorname{im} f = G_2$; it is **bijective** if it is both injective and surjective.

A **ring** $(R, +, \cdot)$ is a set equipped with two operations such that $(R, +)$ is an abelian group, $(R, \cdot)$ is a monoid and the operations satisfy **distributivity**:

$(x+y)z = xz + yz$. A ring is commutative if '$\cdot$' is commutative. A subring is a set $S \subseteq R$ such that $1 \in S$, $(S,+)$ is subgroup of $(R,+)$, and is **multiplicatively closed**: $x, y \in S \Rightarrow xy \in S$.

An **invertible** element or **unit** of the ring is an element that admits inverse with respect to '$\cdot$'. The set of units is a subring. A **domain** is a ring such that $xy = 0$ implies $x = 0$ or $y = 0$. A **field** is a ring such that $(R \setminus \{0\}, \cdot)$ is an abelian group.

A **left** (or **right**) **ideal** of a ring R is a subset $I \subseteq R$ such that $(I,+)$ is a subgroup of $(R,+)$ and $I = RI \triangleq \{ri : r \in R, i \in I\}$ (or $I = IR$); if we say only ideal, we mean left and right ideal. An ideal I is called **principal** if $I = Rx = R\{x\}$ for some $x \in R$: x is a **generator** for I. If I is generated by $\{a_1, \dots, a_n\}$ we write $I = (a_1, \dots, a_n)_R$.

A.2 Modules and sequences

A **left** or **right module** over the ring $\mathcal{R}$ is an abelian group M together with an operation of $\mathcal{R}$ on M such that if $a, b \in \mathcal{R}$ and $x, y \in M$: $(a+b)x = ax + bx$ and $a(x+y) = ax + ay$ (for a right module the order is inverted: $xa \in M$. A submodule is a subgroup that is still an $\mathcal{R}$-module.

M is **generated** by elements $\{x_i\}$ if every element in M is a linear combination, with coefficients in $\mathcal{R}$ of $\{x_i\}$. A set $\{x_i\}$ is **linearly dependent** if there is a linear combination with non zero coefficient that is zero. If M is generated by a set that is not linearly dependent, then the set is called **basis** and the module is **free**.

A **vector space** is a module over a field; it has a basis and therefore is always free.

A module is **torsion** if there are non zero elements $a \in \mathcal{R}$ and $x \in M$, such that $ax = 0$; in this case x is called **torsion element**. The set of torsion elements is a submodule, called **torsion submodule**.

If M is finitely generated over $\mathcal{R}$ a principal ideal ring, then $M = M_f \oplus M_t$ where M_f is free, M_t is the torsion submodule and the sum is direct, meaning that only $0 \in M_t \cap M_f$.

A **homomorphism** of two modules (M into N) over the same ring $\mathcal{R}$ is a group homomorphism such that $\forall a \in \mathcal{R}$, $\forall x \in M$ $f(ax) = af(x)$. Kernels and images are defined in the same way as for group homomorphisms.

An **sequence of modules** $\{M_i, \phi_i\}$ is a chain of maps $\phi_i : M_i \to M_{i+1}$ such that $\operatorname{im}\phi_i \subseteq \ker\phi_{i+1}$, i.e. $\phi_{i+1}\phi_i = 0$; it is **exact** if $\operatorname{im}\phi_i = \ker\phi_{i+1}$. Usually an exact sequence is represented graphically as

$$\cdots \longrightarrow M_{i-1} \xrightarrow{\phi_{i-1}} M_i \xrightarrow{\phi_i} M_{i+1} \longrightarrow \cdots .$$

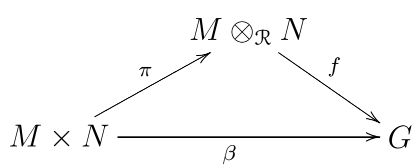
Using this notation the homomorphism $\phi : M \to N$ is injective or surjective if

$$0 \longrightarrow M \xrightarrow{\phi} N \quad \text{or} \quad M \xrightarrow{\phi} N \longrightarrow 0 .$$

The set of $\mathcal{R}$-linear maps of M into N, denoted by $\operatorname{Hom}_{\mathcal{R}}(M, N)$ is an abelian group.

Another important algebraic structure is the tensor product: given a right $\mathcal{R}$-module M and a left $\mathcal{R}$-module N, we say that $\beta : M \times N \to G$, where G is an abelian group, is **balanced** if $\beta(\cdot, n)$ and $\beta(m, \cdot)$ are $\mathcal{R}$-linear for every fixed $m \in M$ and $n \in N$ and $\beta(mr, n) = \beta(m, rn)$ for every $r \in \mathcal{R}$.

Then the pair $(M \otimes_{\mathcal{R}} N, \pi)$, where $M \otimes_{\mathcal{R}} N$ is an abelian group (unique to within isomorphisms) and $\pi : M \times N \to M \otimes_{\mathcal{R}} N$ is an $\mathcal{R}$-balanced map such that every other $\mathcal{R}$-balanced map $\beta : M \times N \to G$, determines a unique group homomorphism $f : M \otimes_{\mathcal{R}} N \to G$ such that $\beta = f \circ \pi$, i.e. the diagram



commutes. Since the map π is canonical, we indicate only $M \otimes_{\mathcal{R}} N$ and call it the **tensor product** of M and N.

A.3 Functors

See, for an incredibly brief summary of category theory [AF92, 0.11-0.13].

If $\mathcal{S}$ is a ring, then the set $\operatorname{Hom}_{\mathcal{R}}(M, N)$ of homomorphisms of left $\mathcal{R}$-modules is a left (right) $\mathcal{S}$-module if and only if M (N) is also a right (left) $\mathcal{S}$-module. So, if we denote by ${}_{\mathcal{R}}M_{\mathcal{S}}$ the fact that M is both a left $\mathcal{R}$-module and a right $\mathcal{S}$-module, and call it a $\mathcal{R},\mathcal{S}$-module, then

$$f \in {}_{\mathcal{S}}H_{\mathcal{T}} = \mathrm{Hom}_{\mathcal{R}}({}_{\mathcal{R}}M_{\mathcal{S}}, {}_{\mathcal{R}}N_{\mathcal{T}}) \;\Rightarrow\; (sft)(m) = (f(ms))t \;\; \forall s \in \mathcal{S}, t \in \mathcal{T}, m \in {}_{\mathcal{R}}M_{\mathcal{S}}.$$

We say, in this case that $\mathrm{Hom}_{\mathcal{R}}(\cdot, {}_{\mathcal{R}}N_{\mathcal{T}})$ maps $\mathcal{R},\mathcal{S}$-modules into $\mathcal{S},\mathcal{T}$-modules; moreover, if $\phi : M_1 \to M_2$ there is a map

$$\mathrm{Hom}_{\mathcal{R}}(\phi, N) : \mathrm{Hom}_{\mathcal{R}}(M_2, N) \to \mathrm{Hom}_{\mathcal{R}}(M_1, N), \; \lambda_2 \mapsto \lambda_1 = \lambda_2 \circ \phi$$

where $\lambda_1 : M_1 \to N$ and $\lambda_2 : M_2 \to N$ or also

$$\begin{array}{ccc} M_1 & \xrightarrow{\phi} & M_2 \\ & \underset{\lambda_1=\lambda_2\circ\phi}{\searrow} & \downarrow{\scriptstyle \lambda_2} \\ & & N \end{array}$$

with a more intuitive representation.

$\mathrm{Hom}_{\mathcal{R}}(\cdot, N)$ is a particular type of **functor** [Lan93]), a pair of maps, in our case, one between different classes of modules and the other one between their homomorphisms. This functor is **contravariant** since it *reverses* arrows:

$$\begin{array}{ccc} M_1 & \xrightarrow{\qquad\phi\qquad} & M_2 \\ \mathrm{Hom}_{\mathcal{R}}(M_1, N) & \xleftarrow{\mathrm{Hom}_{\mathcal{R}}(\phi,N)} & \mathrm{Hom}_{\mathcal{R}}(M_2, N) \end{array} \qquad \Big\downarrow \mathrm{Hom}_{\mathcal{R}}(\cdot, N)$$

Also the tensor product gives rise to a functor: in fact it is possible to show [AF92, p. 221] that

$${}_{\mathcal{S}}P_{\mathcal{T}} = {}_{\mathcal{S}}M_{\mathcal{R}} \otimes_{\mathcal{R}} {}_{\mathcal{R}}N_{\mathcal{T}}.$$

In this case the functor $M \otimes_{\mathcal{R}} \cdot$ maps ${}_{\mathcal{R}}N_{\mathcal{T}} \mapsto {}_{\mathcal{S}}P_{\mathcal{T}}$; this tensor is **covariant**, i.e. it preserves the direction of arrows:

$$\begin{array}{ccc} N_1 & \xrightarrow{\qquad\phi\qquad} & N_2 \\ M \otimes_{\mathcal{R}} N_1 & \xrightarrow{M\otimes_{\mathcal{R}}\phi} & M \otimes_{\mathcal{R}} N_2 \end{array} \qquad \Big\downarrow M \otimes_{\mathcal{R}} \cdot$$

A.4 Behaviors, homomorphisms and tensors

Two functors F and G, F transforming $\mathcal{R}$-modules into $\mathcal{S}$-modules and G vice versa, are an **adjoint pair** whenever $\mathrm{Hom}_{\mathcal{S}}(F(M), N)$ and $\mathrm{Hom}_{\mathcal{R}}(M, G(N))$ are $\mathbb{Z}$-isomorphic.

The two functors we have introduced so far are an adjoint pair: as shown in [AF92, pr. 20.6]), if we have modules ${}_{\mathcal{R}}M$, ${}_{\mathcal{S}}W_{\mathcal{R}}$ and $N_{\mathcal{S}}$ then the following is an isomorphism of abelian groups

$$\Psi : \mathrm{Hom}_{\mathcal{R}}(M, \mathrm{Hom}_{\mathcal{S}}(W, N)) \to \mathrm{Hom}_{\mathcal{S}}(W \otimes_{\mathcal{R}} M, N),\ \Psi(\phi)(w \otimes m) = \phi(m)(w). \quad \text{(A.1)}$$

If moreover M is an $\mathcal{R}, \mathcal{T}$-module, Ψ is an isomorphism of (left) $\mathcal{T}$-modules: indeed for any $t \in \mathcal{T}$ and $\phi, \gamma \in \mathrm{Hom}_{\mathcal{R}}(M, \mathrm{Hom}_{\mathcal{S}}(W, N))$ we have

$$\begin{aligned}\Psi(t\phi + \gamma)(w \otimes m) &= [t\phi(m) + \gamma(m)](w)\\ &= [\phi(mt) + \gamma(m)](w) = \Psi(\phi)(w \otimes mt) + \Psi(\gamma)(w \otimes m)\\ &= t\Psi(\phi)(w \otimes m) + \Psi(\gamma)(w \otimes m) = [t\Psi(\phi) + \Psi(\gamma)](w \otimes m).\end{aligned}$$

We can employ this property to prove an invariance property of behaviors. As we have shown in proposition 5.2, if we put $\mathcal{M} \triangleq \mathrm{coker}_{\mathcal{R}} \circ R$ with $R \in \mathcal{R}^{p \times q}$ and the ring $\mathcal{R}$ operates on $\mathcal{E}$, then the behavior $\mathcal{B} = \ker_{\mathcal{E}} R$ is isomorphic to $\mathrm{Hom}_{\mathcal{R}}(\mathcal{M}, \mathcal{E})$. In section 3.2.2 we have introduced the module $\mathcal{S} \otimes_{\mathcal{R}} \mathcal{M}$ where $\mathcal{S}$ is an overring of $\mathcal{R}$. If even the elements of $\mathcal{S}$ operate on $\mathcal{E}$, then the relationship between $\mathcal{B}$ and $\mathrm{Hom}_{\mathcal{S}}(\mathcal{S} \otimes_{\mathcal{R}} \mathcal{M}, \mathcal{E})$ is very simple: they are (isomorphic to) the same module.

Actually, since $\mathcal{M}$ is both a left and right $\mathcal{R}$-module and $\mathcal{S}$ is, among others, an $\mathcal{S}, \mathcal{R}$-module, by equation A.1 we can write

$$\mathrm{Hom}_{\mathcal{R}}(\mathcal{M}, \mathrm{Hom}_{\mathcal{S}}(\mathcal{S}, \mathcal{E})) \cong \mathrm{Hom}_{\mathcal{S}}(\mathcal{S} \otimes_{\mathcal{R}} \mathcal{M}, \mathcal{E}). \quad \text{(A.2)}$$

Finally $\mathrm{Hom}_{\mathcal{S}}(\mathcal{S}, \mathcal{E})$, as a left $\mathcal{R}$-module, is isomorphic to ${}_{\mathcal{R}}\mathcal{E}$: to any $w(t) \in \mathcal{E}$ we can associate a left $\mathcal{R}$-homomorphism $\theta_w : \mathcal{S} \to \mathcal{E}$, $s \mapsto \theta_w(s) = sw$, since $r\theta_w = \theta_{rw}$; moreover $\theta_w(s) = sw = 0\ \forall s \in \mathcal{S}$ if and only if $w(t) = 0$ and for every $\xi \in \mathrm{Hom}_{\mathcal{S}}(\mathcal{S}, \mathcal{E})$, we have $\xi = \theta_{\xi(1)}$.

Therefore, from equation A.2 we obtain the following isomorphism of left $\mathcal{R}$-

modules:

$$\mathcal{B} \cong \mathrm{Hom}_{\mathcal{R}}(\mathcal{M}, \mathcal{E}) \cong \mathrm{Hom}_{\mathcal{S}}(\mathcal{S} \otimes_{\mathcal{R}} \mathcal{M}, \mathcal{E}).$$

Appendix B

Basic Topology

This appendix contains mainly topological definitions: fundamental theorems are listed in chapter 2

B.1 Topological spaces

A **topological space** (S, τ) is a set S equipped with a topology τ, i.e. a set containing: S itself and the empty set $\emptyset$; the intersection of the elements of every finite subset; the union of the elements of every countable subset.

Every $A \in \tau$ is a subset of S and is said **open**. Every open set that contains $x \in S$ is a **neighborhood** of x. Complements of open sets are **closed**.

$E \subseteq S$ is **dense** in $F \subseteq S$ if $\bar{E} \supseteq F$; if $\bar{E} = V$ then E is **dense**.

(S, τ) is **Hausdorff** if distinct points have disjoint neighborhoods.

$x \in S$ is a **limit point** of $E \subseteq S$ if for every neighborhood A of x, $A \cap E \neq \emptyset$.

The **closure** of $E \subseteq S$ is the set $\bar{E}$ of its limit points and also the intersection of all closed sets that contain E. The **interior** of E is the set $\mathring{E}$, union of every open subset of E.

An **open cover** of a set $E \subseteq S$ is a collection of open sets whose union is a superset of E. The set $E \subseteq S$ is **compact** if every open cover admits a finite subcover.

A collection $\tau' \subseteq \tau$ is a **base** for τ if every open set in τ is a union of members of τ'.

A collection γ of neighborhoods of $x \in S$ is a **local base** at x if every neighborhood of x contains a member of γ.

A sequence $\{x_n\} \subseteq S$ converges to $x \in S$ if every neighborhood of x contains all but finitely many of the points x_n.

If (V, τ) and (W, ν) are topological spaces, then the map $\alpha : V \to W$ is **continuous** at $x \in V$ if for every neighborhood $U \ni \alpha(x)$ there is a neighborhood $V \ni x$ such that $\alpha(V) \subseteq U$.

A function $d : S^2 \to \mathbb{R}^+$ is a metric if $d(x,y) = 0$ iff $x = y$; $d(x,y) = d(y,x)$; $d(x,z) \leq d(x,y) + d(y,z)$. d defines a topology whose open sets are $B_r(x) = \{y \in S : d(x,y) < r\}$. τ is compatible with the metric d if τ is compatible with the topology defined by d.

A set $C \subseteq V$ is **convex** if $tC + (1-t)C \subseteq C$, $0 \leq t \leq 1$. A set $B \subseteq V$ is **bounded** if for every neighboorhood of zero A there is a $t > 0$ such that $B \subseteq tA$.

B.2 Topological vector spaces

A **topological vector space** (V, τ) is a vector space V equipped with a topology τ, such that every $x \in V$ is closed and the operation of V are continuous. These conditions together imply that (V, τ) is Hausdorff.

A topological vector space is **translation-invariant**: a set $E \subseteq V$ is open if and only if $x + E$ is open for every $x \in V$. Thus τ is completely determined by one local base γ, usually the local base at 0.

A metric d on a topological vector space is invariant if $d(x+z, y+z) = d(x,y)$ forall $x, y, z \in V$.

A topological vector space is **locally convex** if there is a base γ whose members are convex.

A topological vector space is **metrizable** if its topology is compatible with some metric.

A **seminorm** on a vector space is a function $p : V \to \mathbb{R}$ such that $p(x+y) \leq p(x) + p(y)$; $p(\alpha x) = |\alpha| p(x)$, α a (real or complex) scalar.

A family P of seminorms is called **separating** if $\forall x \in V$ there is at least one $p \in P$ such that $p(x) \neq 0$

The dual space of the dual space, is called **bidual**; a topological vector space is **reflexive** if it is isomorphic to its bidual. The isomorphism is given by the 'evaluation at a point': for every $x \in V$ it is the map $\phi_x : V' \to \mathbb{F}$ that maps $\alpha \mapsto \phi_x(\alpha) = \langle \alpha, x \rangle$.

B.3 Smooth functions

Just a few words about the properties that make $C^\infty(\mathbb{R},\mathbb{R})$ into the Fréchet space (metrizable, complete and locally convex topological vector space) $\mathcal{E}$.

- $\mathcal{E}$ is metrizable because it has a countable separating family of seminorms defined for every $f \in \mathcal{E}$

  $$p_n(f) = \max_{x\in K_n} \left\{|f^{(i)}(x)|,\ i \le n\right\}$$

  where K_n is an increasing sequence of compacts such that $\cup K_n = \mathbb{R}$.

- The local convex basis induced by these seminorms is

  $$V_n = \left\{ f \in \mathcal{E} : p_n(f) < \frac{1}{n} \right\}.$$

- With this topology a sequence $\{f_j\}$ converges to f when every derivative converges uniformly on any compact subset of $\mathbb{R}$.

Appendix C

Notations and symbols

This appendix offers a reference to the page where some symbols or concepts have been defined or used for the first time.

Proofs and examples are terminated by the symbols $\square$ and respectively $\clubsuit$.

C.1 Linear systems

$\mathcal{B}$	Behavior, set of trajectories of a dynamical system, 24.
$\mathcal{B}_c$	Subset of trajectories with compact support of $\mathcal{B}$, 101.
$\mathcal{M}$	System in the module theoretic approach, 35.
ker, im	Kernel or image of an operator, 23.
coker	Cokernel of an operator, 35.
Hom	Group of homomorphisms, 60, 114.
$\otimes$	Tensor product, 37, 114.
$^\perp$	Orthogonal module, 67.

C.2 Matrices

I_d	Identity matrix with $d \times d$ elements; d may be omitted.
$\circ R$, $R\circ$	Matrices may act, as operators, on the left or on the right, 23.
$\bar{R}$, $\breve{R}$	See the Smith form of matrix R, 42.
$C_r(R)$	Compound matrix of order r of the matrix R, 45.

Generalized inverses: $R(s) = R(s)G(s)R(s)$, 45

C.3 Spaces of functions and distributions

$\mathcal{D}$	Smooth functions with compact support, *test functions*, 14.
$\mathcal{S}$	Smooth functions rapidly decreasing at infinity, 18.
$\mathcal{E}$	Smooth functions, 12.
$\mathcal{P}_e$	Polynomial exponential functions, 47.
$\mathcal{D}'$	Distributions, 14.
$\mathcal{S}'$	Tempered distributions.
$\mathcal{E}'$	Distributions with compact support, 17.

C.4 Distributions and operators

σ_τ	Shift (delay) operator on functions and distributions, 15, 15.
$\langle \cdot, \cdot \rangle$	Evaluation of distributions at functions, e.g. $\langle \cdot, \cdot \rangle : \mathcal{E}' \times \mathcal{E} \to \mathbb{R}$, 14.
$\check{}$	Symmetric of functions and distributions, 15, 16.
supp	Support of a distribution, 16.
$\star$	Convolution between functions and/or distributions, 16, 16, 16.
$\tilde{}$	As operator on distributions, $\tilde{\alpha}$ is the adjoint of $\star\check{\alpha}$, 18.
$\mathcal{L}, \hat{}$	Laplace operator $\mathcal{L} : \mathcal{E}' \to \mathcal{A}$, $\alpha \mapsto \hat{\alpha}(s)$, 19.

C.5 Operator rings

$\mathcal{R}$	A generic operator ring; it may be any one of the following.
$\mathbb{R}\left[\frac{d}{dt}\right]$	Ring of polynomial differential operators, 26.
$\mathbb{R}\left[\frac{d}{dt}, \boldsymbol{\sigma}\right]$	Polynomial delay–differential operators, 32.
$\mathbb{R}\left[\frac{d}{dt}, \boldsymbol{\sigma}, \boldsymbol{\sigma}^{-1}\right]$	Laurent delay–differential operators, 33.
$\mathbb{R}[s]$	Differential polynomials, 26.
$\mathbb{R}\left[s, e^{-s\boldsymbol{\tau}}\right]$	Exponential polynomials, 54.
$\mathbb{R}\left[s, e^{-s\boldsymbol{\tau}}, e^{s\boldsymbol{\tau}}\right]$	Laurent exponential polynomials, 54.
$\mathcal{H}_m$	Holomorphic fractions of delay–differential polynomials, 57.
$\mathcal{A}$	Paley–Wiener functions, 21.
$\mathcal{O}$	Holomorphic functions, 47.

Bibliography

[AF92] F. W. Anderson and K. R. Fuller. *Rings and Categories of Modules*. Springer-Verlag, New York, 2^{nd} edition, 1992.

[AM69] M. F. Atiyah and I. G. Macdonald. *Introduction to Commutative Algebra*. Addison–Wesley Publishing Company, Reading, MA, 1969.

[BBRMP90] R. B. Bapat, K. P. S. Bhaskara Rao and K. Manjunatha Prasad. Generalized inverses over integral domains. *Linear Algebra Appl.*, 140:181–196, 1990.

[BW93] T. Becker and V. Weispfenning. *Gröbner Bases*. Springer-Verlag, Berlin, 1993.

[BIG74] A. Ben-Israel and T. N. E. Greville. *Generalized Inverses: Theory and Applications*. John Wiley & Sons Ltd., New York, 1974.

[BPDM93] A. Bensoussan, G. da Prato, M. C. Delfour and S. K. Mitter. *Representation and Control of Infinite-Dimensional Systems, Vol I, II*. Birkhauser, Inc., Boston MA, 1992,1993.

[BD74] C. A. Berenstein and M. A. Dostal. The Ritt theorem in several variables. *Ark. Mat.*, 12:267–280, 1974.

[BS93] C. A. Berenstein and D. C. Struppa. Complex analysis and convolution equations. In *Several Complex Variables*, volume 54 of *Encyclopaedia of Mathematical Sciences*, chapter I, pages 5–108. Springer–Verlag, Berlin, 1993.

[BT79] C. A. Berenstein and B. A. Taylor. A new look at interpolation theory for entire functions of one variable. *Adv. in Math.*, 33:109–143, 1979.

[BR83] K. P. S. Bhaskara Rao. On generalized inverses of matrices over integral domains. *Linear Algebra Appl.*, 49:179–189, 1983.

[BFL97] L. Bitauld, M. Fliess and J. Lévine. A flatness based control synthesis of linear systems and application to windshield wipers. In *Proceedings of 4^{th} European Control Conference*, Brussels, July 1997.

[Bro92] W. C. Brown. *Matrices over Commutative Rings.* Marcel Dekker, Inc., 270 Madison Avenue, New York, 1992.

[BBV86] J. W. Brewer, J. W. Bunce and F. S. Van Vleck. *Linear Systems over Commutative Rings.* Marcel Dekker, Inc., New York, 1986.

[Coh95] P. M. Cohn. *Algebra*, volume 2. John Wiley & Sons Ltd., 2^{nd} edition, 1995.

[CZ95] R. F. Curtain and H. Zwart. *An Introduction to Infinite–Dimensional Linear Systems Theory.* Springer-Verlag, Berlin, 1995.

[Dug66] J. Dugundji. *Topology.* Allyn and Bacon, Inc., Boston, 1966.

[Ehr54] L. Ehrenpreis. Solution of some problem of division. Part I. Division by a polynomial of derivation. *Amer. J. Math.*, 76:883–903, 1954.

[Ehr55] L. Ehrenpreis. Solution of some problem of division. Part II. Division by a punctual distribution. *Amer. J. Math.*, 77:286–292, 1955.

[Ehr60] L. Ehrenpreis. Solution of some problem of division. Part IV. Invertible and elliptic operators. *Amer. J. Math.*, 82:522–588, 1960.

[Ehr70] L. Ehrenpreis. *Fourier Analysis in Several Complex Variables.* Wiley-Interscience Publishers, New York, 1970.

[EP97] G. R. Everest and A. J. van der Poorten. Factorisation in the ring of exponential polynomials. *Proc. Amer. Math. Soc.*, 125(5):1293–1298, 1997.

[Fli90a] M. Fliess. Generalized controller canonical forms for linear and nonlinear dynamics. *IEEE Trans. Automat. Control*, 35(9):994–1001, 1990.

[Fli90b] M. Fliess. Some basic structural properties of generalized linear systems. *Systems Control Lett.*, 15(5):391–396, 1990.

[Fli92] M. Fliess. Reversible linear and nonlinear discrete-time dynamics. *IEEE Trans. Automat. Control*, 37(8):1144–1153, August 1992.

[Fli93] M. Fliess. Invertibility of causal discrete time dynamical systems. *J. Pure Appl. Algebra*, 86(2):173–179, 1993.

[FG93] M. Fliess and S. T. Glad. An algebraic approach to linear and nonlinear control. *Essays on Control: Perspectives in the Theory and its Applications (Proceedings of ECC93, Groningen)*, pages 223–267. Birkhauser Inc., Boston, MA, 1993.

[FLMOR97] M. Fliess, J. Lévine, P. Martin, F. Ollivier and P. Rouchon. Controlling nonlinear systems by flatness. In *Systems and Control in the Twenty-First Century (Proceedings MTNS-96, St. Louis, MO)*, pages 137–154. Birkhauser Inc., Boston, MA, 1997.

[FLMR95] M. Fliess, J. Lévine, P. Martin and P. Rouchon. Flatness and defect of non-linear systems: Introductory theory and examples. *Internat. J. Control*, 61(6):1327–1361, 1995.

[FLMR97] M. Fliess, J. Lévine, P. Martin and P. Rouchon. On a new differential geometric setting in nonlinear control. *J. Math. Sci.*, 83(4):524–530, 1997.

[FM95] M. Fliess and H. Mounier. Interpretation and comparison of various types of delay system controllabilities. In *IFAC Conference, System Structure and Control*, pages 330–335, Nantes, France, 1995.

[FM98] M. Fliess and H. Mounier. Tracking control and π-freeness of infinite dimensional linear systems. Submitted for pubblication, 1998.

[FMRR95] M. Fliess, H. Mounier, P. Rouchon and J. Rudolph. Controllability and motion planning for linear delay systems with an application to a flexible rod. In *Proceedings of the 34th Conference on Decision and Control*, pages 2046–2051, New Orleans, 1995.

[GL97a] H. Glüsing-Lüerssen. A behavioral approach to delay-differential systems. *SIAM J. Control Optim.*, 35(2):480–499, March 1997.

[GL97b] H. Glüsing-Lüerssen. First-order representations of delay-differential systems in a behavioral setting. *European J. Combin.*, 3:137–149, 1997.

[Hab94] L. C. G. J. M. Habets. *Algebraic and Computational Aspects of Time-Delay Systems.* PhD thesis, Department of Mathematics and Computing Science, Eindhoven University of Technology, Eindhoven, the Netherlands, 1994.

[Hab98] L. C. G. J. M. Habets. System equivalence for AR-systems over rings with an application to delay-differential systems. Technical Report RANA 98-12, Eindhoven University of Technology, Department of Mathematics and Computing Science, June 1998.

[Hal77] J. Hale. *Theory of Functional Differential Equations.* Springer–Verlag, New York, 1977.

[Hel43] O. Helmer. The elementary divisor theorem for certain rings without chain condition. *Bull. Amer. Math. Soc. (N.S.)*, 49:225–236, 1943.

[Hör67] L. Hörmander. Generators for some rings of analytic functions. *Bull. Amer. Math. Soc. (N.S.)*, 73:943–949, 1967.

[Hör73] L. Hörmander. *An Introduction to Complex Analysis in Several Variables.* North-Holland Publishing Company, Amsterdam, 1973.

[HF98] R. Hotzel and M. Fliess. On linear systems with a fractional derivation: Introductory theory and examples. *Math. Comput. Simulation*, 45(3–4):385–395, 1998.

[KKT86] E. W. Kamen, P. P. Khargonekar and A. Tannenbaum. Proper stable Bezout factorizations and feedback control of linear time-delay systems. *Internat. J. Control*, 43(3):837–857, 1986.

[Lan93] S. Lang. *Algebra.* Addison-Wesley, Reading, MA, 3^{rd} edition, 1993.

[Mal56] B. Malgrange. Existence et approximation des solutions des équations aux dérivées partielles et des équations de convolution. *Ann. Inst. Fourier (Grenoble)*, 6:271–355, 1955–1956.

[MPBR96] K. Manjunatha Prasad and K. P. S. Bhaskara Rao. On bordering of regular matrices. *Linear Algebra Appl.*, 234:245–259, 1996.

[MW95] J. Michalík and J. C. Willems. A new setting for differential algebraic systems. In *Proceedings of 3^{rd} European Control Conference*, pages 218–223, Rome, 1995.

[Mou98a] H. Mounier. Algebraic interpretations of the spectral controllability of a linear delay system. *Forum Math.*, 10(1):39–58, 1998.

[Mou98b] H. Mounier. Stabilization of a class of linear delay systems. *Math. Comput. Simulation*, 45(3–4):329–338, 1998.

[MRPF95] H. Mounier, J. Rudolph, M. Petitot and M. Fliess. A flexible rod as a linear delay system. In *Proceedings of 3^{rd} European Control Conference*, pages 3676–3681, Rome, 1995.

[Niv56] I. Niven. *Irrational Numbers.* John Wiley and Sons, Inc., New York, 1956.

[Obe90] U. Oberst. Multidimensional constant linear systems. *Acta Appl. Math.*, 20(1–2):1–175, 1990.

[OF98] U. Oberst and S. Fröhler. Continuous time-varying linear systems. *Systems Control Lett.*, 35(2):97–110, 1998.

[Pal70] V. P. Palamodov. *Linear Differential Operators with Constant Coefficients*, volume 168 of *Die Grundlehren der mathematischen Wissenschaften*. Springer–Verlag, Berlin, 1970.

[Pol97] J. W. Polderman. Proper elimination of latent variables. *Systems Control Lett.*, 32(5):261–269, 1997.

[PW97] J. W. Polderman and J. C. Willems. *Introduction to Mathematical Systems Theory: A Behavioral Approach*. Springer-Verlag, Berlin, 1997.

[RM71] C. R. Rao and S. K. Mitra. *Generalized Inverse of Matrices and its Applications*. John Wiley & Sons, Inc., New York, 1971.

[Rap98] P. Rapisarda. *Linear Differential Systems*. PhD thesis, Mathematics Institute, University of Groningen, Groningen, the Netherlands, June 1998.

[RW97] P. Rocha and J. C. Willems. Behavioral controllability of delay-differential systems. *SIAM J. Control Optim.*, 35(1):254–264, January 1997.

[Rud73] W. Rudin. *Functional Analysis*. McGraw-Hill, New York, 1973.

[Rud87] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, New York, 3^{rd} edition, 1987.

[Sch47] L. Schwartz. Théorie générale des fonctions moyenne périodiques. *Ann. of Math. (2)*, 48:857–929, 1947.

[Son80] E. D. Sontag. On generalized inverses of polynomial and other matrices. *IEEE Trans. Automat. Control*, 25(3):514–517, 1980.

[Str83] D. C. Struppa. *The Fundamental Principle of Convolution Equations*. Number 273 in Memoirs of the American Mathematical Society. AMS, Providence, USA, 1983.

[TSM98] S. Tatikonda, A. Sahai and S. K. Mitter Control of LQG systems under communication constraints. In *Proceeding of the 37^{th} Conference on Decision and Control, CDC98*, pages 1165–1170, Tampa, 1998.

[Tre67] F. Treves. *Topological Vector Spaces, Distributions and Kernels.* Academic Press, New York, London, 1967.

[Val98] M. E. Valcher. Delay-differential systems in the behavioral framework. Submitted for pubblication, 1998.

[Wil86] J. C. Willems. From time seris to linear system — part I. finite dimensional linear time invariant systems. *Automatica J. IFAC*, 22(5):561–580, 1986.

[Wil89] J. C. Willems. Models for dynamics. In *Dynamics Reported*, volume 2, pages 171–269. John Wiley & Sons Ltd., 1989.

[Wil91] J. C. Willems. Paradigms and puzzles in the theory of dynamical systems. *IEEE Trans. Automat. Control*, 36(3):259–294, March 1991.

[YP84] D. C. Youla and P. F. Pickel. The Quillen–Suslin theorem and the structure of n–dimensional elementary polynomial matrices. *IEEE Trans. Circuits and Systems*, 31(6):513–518, June 1984.

[ZM96] S. Zampieri and S. K. Mitter. Linear systems over Noetherian rings in the behavioural approach. *J. Math. Systems Estim. Control*, 6:235–238, 1996. Summary; the full paper, file `15711.ps`, is available via ftp from the publisher or at http://www.birkhauser.com/journals/jmsec/download.html.

[ZWZ97] Y. Zheng, J. C. Willems and C. Zhang. Common factors and controllability of nonlinear systems. In *Proceeding of the* 36^{th} *Conference on Decision and Control, CDC97*, pages 2584–2589, San Diego, 1997.