Chapter 2

Dynamical Systems

In this second chapter we introduce the essential mathematical concepts underlying the approach to modeling systems that emerged in the previous chapter. This provides us with the essential background needed in the following to develop algorithms to solve some of the basic questions which have started emerging in the previous pages. At the same time, as is often the case when one sets out to formalize concepts, the definitions we make turn out to allow us to deal with an even wider range of applications than the one described in the previous chapter.

The starting point of our discussion is the definition of a system as the set of feasible trajectories of the variables whose dynamics we are interested in modeling (the manifest variables of chapter 1). Such a definition has two crucial aspects: on the one hand it abandons the idea of a system as an input/output map, as a signal processor, and simply makes it a dynamic relation between quantities for which no hierarchical or cause/effect structure is a priori given. On the other hand it clearly distinguishes between the system (i.e. the feasible trajectories) and its representations (e.g. equations, graphs, grammar rules, etc...). Very relevant system properties such as linearity and shift invariance can already be defined at such an abstract level.

At the level of representations we concentrate on linear differential systems, i.e. systems whose trajectories can be described as solutions to a set of linear differential equations; more in particular, a central role is played by linear, ordinary and partial, differential equations with constant coefficients.
In the spirit of most of this work, we discuss ODE’s and PDE’s as much as possible in parallel, in such a way as to highlight analogies and differences between the two cases.

Great attention will be payed to showing how the same system admits many different representations; this is, actually, the central result of the whole chapter. In order to establish this, we first introduce some concepts from commutative algebra, namely ideals and modules over polynomial rings, and then proceed to establish a connection with systems of partial differential equations.

Detailed references on the concepts presented in this chapter are [67] and [54] for the general framework and systems modeled by ordinary differential or difference equations. References [52], [70] and [51] are exhaustive presentations of systems described by partial differential equations, while [59] and [51] address the partial difference case.

2.1 Mathematical Models

In the previous chapter we saw that, when modeling a dynamical system, one essentially tries to describe how a set of variables of interest, call them \( w \), evolve as a function of another set of independent variables, call them \( x \).

We denote by \( \mathbb{I} \subseteq \mathbb{R}^n \) the set in which the \( x \)'s take their value (e.g. \( \mathbb{I} = \mathbb{R} \) for models that depend on continuous time, \( \mathbb{I} = \mathbb{R}^4 \) for models depending on continuous time and space, \( \mathbb{I} = \mathbb{Z} \) for those that evolve as function of discrete time, etc. ) and by \( \mathcal{W} \) the space in which the variables of interest take on their values (e.g. \( \mathbb{R}^w \) if there are \( w \) real valued variables). The \( w \)'s are then elements of \( \mathcal{W} \), with \( \mathcal{W} \) denoting the set of maps from \( \mathbb{I} \) to \( \mathcal{W} \). The model of the system tells us that only a subset of such trajectories can actually happen, namely the subset that complies with the laws of the system. We indicate this set of admissible trajectories as \( \mathfrak{B} \) and refer to it as the behavior of the dynamical system.

Formalizing the above discussion, we have:

**Definition 16** : A dynamical system is a triple \( \Sigma = (\mathcal{W}, \mathbb{I}, \mathfrak{B}) \) with \( \mathcal{W} \) the signal space, \( \mathbb{I} \) the index set and \( \mathfrak{B} \subseteq \mathcal{W}^\mathbb{I} \) the behavior of the system.

The above definition is the cornerstone of the rest of this work, and in essence
2.1. Mathematical Models

it defines a model as an exclusion law, a rule that allows us to pick a subset of feasible trajectories out of a set of possible ones. The whole previous chapter is full of examples of systems presented in the spirit of definition 16; given its crucial importance, however, we want to further illustrate it with a few more examples which will also point out how generally applicable the above definition is.

Example 17: **Kepler’s laws** describe the possible motions of the planets in the solar system. This defines a dynamical system with $T = \mathbb{R}$, $W = \mathbb{R}^3$, and $\mathcal{B}$ the set of maps $w : \mathbb{R} \to \mathbb{R}^3$ that satisfy Kepler’s laws: the paths $w$ must be ellipses in $\mathbb{R}^3$ with the sun (assumed in fixed position, say the origin of $\mathbb{R}^3$) in one of the foci; the radius vector from the sun to the planet must sweep out equal areas in equal time, and the ratio of the period of revolution around the ellipse to the major axis must be the same for all $w$’s in $\mathcal{B}$.

**Example 18:** **Maxwell’s equations** describe the possible realizations of the fields $\vec{E} : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$, $\vec{B} : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$, $\vec{j} : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$, and $\rho : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}$. They are

\[
\nabla \cdot \vec{E} = \frac{1}{\varepsilon_0} \rho, \\
\nabla \times \vec{E} = -\frac{\partial}{\partial t} \vec{B}, \\
\nabla \cdot \vec{B} = 0, \\
c^2 \nabla \times \vec{B} = \frac{1}{\varepsilon_0} \vec{j} + \frac{\partial}{\partial t} \vec{E},
\]

with $\varepsilon_0$ the dielectric constant of the medium and $c^2$ the speed of light in the medium. The defines the system $(\mathbb{R} \times \mathbb{R}^3, \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}, \mathcal{B})$, with $\mathcal{B}$ the set of all fields $(\vec{E}, \vec{B}, \vec{j}, \rho) : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}$ that satisfy Maxwell’s equations.

**Example 19:** A code $\mathcal{C}$ is a subset of $\mathbb{A}^d$ with $\mathbb{A}$ the code alphabet (e.g. $\mathbb{A} = \mathbb{F}^* \times \mathbb{F}^*$ with $\mathbb{F}$ a finite field) and $\mathbb{I}$ an indexing set (e.g. $\mathbb{I} = (1, \ldots, n)$ in block codes, $\mathbb{I} = \mathbb{N}$ or $\mathbb{Z}$ in convolutional codes). Codes are thus dynamical systems according to the abstract definition 16.
Classical notions such as linearity and shift-invariance are very naturally introduced starting from the above formal definition of a system. In particular we talk about a linear system if \( \mathbb{W} \) is a vector space and \( \mathcal{B} \) a linear subspace of \( \mathbb{W}^\mathbb{I} \), and of a shift-invariant one if \( \mathbb{I} \) is a semigroup under addition and \( \sigma^x \mathcal{B} = \mathcal{B} \) for all \( x \in \mathbb{I} \), where \( \sigma^x \) denotes the \( x \)-shift operator defined by \( (\sigma^x f)(x') := f(x' + x) \). In case \( \mathbb{I} = \mathbb{R} \) or \( \mathbb{I} = \mathbb{Z} \), the above definition of shift-invariance corresponds to the familiar concept of a time-invariant system.

### 2.2 Differential systems

As discussed in the above section, when it comes to modeling a dynamical system, what we really are after is the behavior \( \mathcal{B} \), the set of admissible trajectories. Of course such a set can be described in many possible ways, for example through differential equations as in Newton’s second law, or through formal descriptions, such as in Kepler’s laws. It is therefore conceptually misleading to identify the idea of system with that of a set of differential equations, because, as we just pointed out, equations are just one of many possible instruments which can be used to specify behaviors (as further examples think of finite state automata whose behavior is typically described through graphs, or non-linear electronic components, often described through characteristics in the I-V plane).

Although identifying systems with equations is not appropriate, the class of systems whose behavior is specified by differential equations plays such a prominent role in physical and engineering applications that they deserve special attention. We define a differential system as a system with \( \mathbb{I} = \mathbb{R}^n \) and \( \mathbb{W} = \mathbb{R}^p \) whose behavior \( \mathcal{B} \) consists of all solutions of a finite set of partial differential equations of the form:

\[
f(x, w(x), \ldots, \frac{\partial^k}{\partial x^k} w(x), \ldots) = 0 \tag{2.1}
\]

Here we use the multi-index notation \( \frac{\partial^k}{\partial x^k} = \frac{\partial}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}} \) and assume that only a finite number, say \( d \), of such partial derivatives are involved in the above equations. \( f \) is, therefore, a function from \( \mathbb{R}^{dn} \) to \( \mathbb{R}^p \), with \( p \) the number of differential equations defining the system and \( w \) the cardinality of the vector \( w \).
2.2. Differential systems

A special subclass of differential systems is represented by ordinary differential systems, i.e. system with \( I = \mathbb{R} \) whose behavior \( \mathcal{B} \) consists of all solutions of a finite set of ordinary differential equations of the form:

\[
f(t, w(t), \frac{d}{dt}w(t), \ldots, \frac{d^L}{dt^L}w(t)) = 0. \tag{2.2}\]

with \( f \) a function from \( \mathbb{R}^{L+1} \) to \( \mathbb{R}^p \), with \( p \) the number of differential equations defining the system, and \( w \) the cardinality of the vector \( w \). Notice that, in this case, the independent variable is indicated by \( t \) instead of \( x \) because time is the independent variable for most systems described by ordinary differential equations.

This work actually concentrates on the special subclass of (2.1) given by linear constant coefficient partial differential systems, namely systems with \( I = \mathbb{R}^n \), \( \mathcal{W} = \mathbb{R}^p \) and behavior \( \mathcal{B} \) consisting of all functions \( w : x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \rightarrow w(x) = (w_1(x), w_2(x), \ldots, w_N(x)) \in \mathbb{R}^N \) that are a solution of equations of the form:

\[
R(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n})w = 0 \tag{2.3}\]

where \( R \in \mathbb{R}^{N \times N}[\xi_1, \xi_2, \ldots, \xi_n] \) is a polynomial matrix in \( n \) indeterminates with \( w \) columns and an arbitrary but finite number of rows. Therefore \( R(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}) \) is a linear constant coefficient partial differential operator and (2.3) is a system of partial differential equations.

Of course, one has to specify what kind of solutions one is looking for when considering equations such as 2.3. In the following, unless otherwise stated, we will present results which are valid if one considers \( \mathcal{B} \) either as the set of smooth solutions of the above set of equations, in other words

\[
\mathcal{B} = \{ w \in C^\infty(\mathbb{R}^n, \mathbb{R}^p) \mid R(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n})w = 0 \}\]

or as the set of distributional solutions, with differentiation and equality taken in the appropriate sense, therefore

\[
\mathcal{B} = \{ w \in \mathcal{D}'(\mathbb{R}^n, \mathbb{R}^p) \mid R(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n})w = 0 \}.
\]

Here \( C^\infty(\mathbb{R}^n, \mathbb{R}^p) \) is the set of infinitely differentiable functions from \( \mathbb{R}^n \) to \( \mathbb{R}^p \) while \( \mathcal{D}'(\mathbb{R}^n, \mathbb{R}^p) \) is the set of all \( \mathbb{R}^p \) valued distributions over \( \mathbb{R}^n \).
It is easily seen that in both cases \( \mathcal{B} \) is linear and shift invariant. For obvious reasons we refer to (2.3) as a kernel representation of \( \mathcal{B} \) and also write \( \mathcal{B} = \ker(R(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n})) \). We often use the notation \( \mathcal{L}_n^k \) to indicate the set of all linear shift-invariant linear partial differential behaviors with \( n \) independent and \( w \) dependent variables. Thus \( (\mathcal{B} \in \mathcal{L}_n^k) \Leftrightarrow (\exists R \in \mathbb{R}^{k \times n}[\xi_1, \xi_2, \ldots, \xi_n], \text{such that } \mathcal{B} = \ker(R(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}))) \).

**Example 20**: To the matrix

\[
R = \begin{pmatrix}
0 & -\xi_3 & \xi_2 \\
\xi_3 & 0 & -\xi_1 \\
-\xi_2 & \xi_1 & 0
\end{pmatrix}
\]

we associate the system of linear partial differential equations

\[
\begin{align*}
\frac{\partial w_0}{\partial x_3} &= \frac{\partial w_3}{\partial x_2} \\
\frac{\partial w_1}{\partial x_3} &= \frac{\partial w_3}{\partial x_2} \\
\frac{\partial w_2}{\partial x_3} &= \frac{\partial w_1}{\partial x_2}
\end{align*}
\]

(2.4)

The behavior \( \mathcal{B} \in \mathcal{L}_3^3 \) defined by (2.4) is given by all functions \( w = (w_1, w_2, w_3) \in C^\infty(\mathbb{R}^3, \mathbb{R}^3) \) that solve (2.4). It is not difficult to recognize that the partial differential operator \( R(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}) \) is the three dimensional curl, \( \nabla \times \).

**Example 21**: Setting \( w = (u_x, u_y, F_x, F_y) \), the behavior of the elastic membrane from example 7 can be described as \( \ker((R(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}))) \) with

\[
R = \begin{bmatrix}
\rho \xi_3^2 - (\lambda + 2\mu)\xi_1^2 - \mu \xi_2^2 & -(\lambda + \mu)\xi_1 \xi_2 & -1 & 0 \\
-(\lambda + \mu)\xi_1 \xi_2 & \rho \xi_3^2 - (\lambda + 2\mu)\xi_1^2 - \mu \xi_2^2 & 0 & -1
\end{bmatrix}
\]

and the association \( \xi_1 \rightarrow \frac{\partial}{\partial x}, \xi_2 \rightarrow \frac{\partial}{\partial y}, \xi_3 \rightarrow \frac{\partial}{\partial t} \). It follows that \( \mathcal{B} \in \mathcal{L}_3^4 \).

**Example 22**: Maxwell’s equations as presented in example 18 define a behavior in \( \mathcal{L}_4^{10} \). They can, in fact, be written in the form (2.3) by setting
2.2. Differential systems

\[ w = (E, B, \gamma, \rho) \in \mathbb{R}^{10} \] and

\[
R = \begin{bmatrix}
\xi_1 & \xi_2 & \xi_3 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{c_0} \\
0 & -\xi_3 & \xi_2 & -\xi_4 & 0 & 0 & 0 & 0 & 0 \\
\xi_3 & 0 & -\xi_1 & 0 & -\xi_4 & 0 & 0 & 0 & 0 \\
-\xi_2 & \xi_1 & 0 & 0 & 0 & -\xi_4 & 0 & 0 & 0 \\
0 & 0 & 0 & \xi_1 & \xi_2 & \xi_3 & 0 & 0 & 0 \\
\xi_4 & 0 & 0 & 0 & -c^2 \xi_3 & c^2 \xi_2 & \frac{1}{c_0} & 0 & 0 \\
0 & \xi_4 & 0 & c^2 \xi_3 & 0 & -c^2 \xi_1 & 0 & \frac{1}{c_0} & 0 \\
0 & 0 & \xi_4 & -c^2 \xi_2 & c^2 \xi_1 & 0 & 0 & 0 & \frac{1}{c_0} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

In this case \( R \in \mathbb{R}^{8 \times 10} [\xi_1, \xi_2, \xi_3, \xi_4] \) with the association \( \xi_1 \rightarrow \frac{\partial}{\partial x}, \xi_2 \rightarrow \frac{\partial}{\partial y} \), \( \xi_3 \rightarrow \frac{\partial}{\partial z}, \xi_4 \rightarrow \frac{\partial}{\partial \bar{w}} \).

A very important special case of behaviors in \( \mathcal{L}_n^x \) is obtained when \( n = 1 \), in which case we use the notation \( \mathcal{L}^x \). These systems, in fact, are defined by ordinary differential equations. Formally, we now consider a one variable polynomial matrix \( R = R_0 + R_1 \xi + \cdots + R_L \xi^L \in \mathbb{R}^{x \times y}[\xi] \) with \( m \) columns and an arbitrary but finite number of columns. To \( R \) we associate the set of ordinary differential equations

\[
R_0 \dot{w} + R_1 \frac{d}{dt} \dot{w} + \cdots + R_L \frac{d^L}{dt^L} \dot{w} = 0
\]

often also written as

\[
R \left( \frac{d}{dt} \right) \dot{w} = 0 \] (2.5)

Notice how algebraic constraints (i.e. differential equations of order 0) are automatically included in equations (2.5). To (2.5) corresponds the behavior

\[
\mathcal{B} = \{ w \in C^\infty(\mathbb{R}, \mathbb{R}^y) \mid R \left( \frac{d}{dt} \right) \dot{w} = 0 \}
\]

or

\[
\mathcal{B} = \{ w \in \mathcal{D}(\mathbb{R}, \mathbb{R}^y) \mid R \left( \frac{d}{dt} \right) \dot{w} = 0 \}
\]

Again, the variable with respect to which differentiation is taken is indicated by \( t \) corresponding to the fact that in most models described by ordinary differential equations, time is the independent variable. Systems in \( \mathcal{L}^x \) are referred to as linear time-invariant differential systems. Again, we often refer to the above as a kernel representation of the behavior of our linear time-invariant differential system, and write \( \mathcal{B} = \ker \left( R \left( \frac{d}{dt} \right) \right) \).
Example 23: Setting \( w = (V_0, I_0, V_{12}, I_{12}) \), the manifest behavior of the electrical circuit in example 1.13 corresponding to equations (1.4), can be written in the form \( \mathcal{B} = \ker(R(\frac{d}{dt})) \) for

\[
R = \begin{bmatrix}
\frac{E_C}{R_C} + (1 + \frac{E_C}{R_C})CRc \xi + CRc \frac{1}{R_C} \xi^2 & -(1 + CRc + \frac{1}{R_C})Re \xi + \frac{CLR^2}{R_C} \xi^2 \\
-(\frac{E_C}{R_C} + (1 + \frac{E_C}{R_C})CRc \xi + CRc \frac{1}{R_C} \xi^2) & 0
\end{bmatrix}
\]

Thus \( R \in \mathbb{R}^{2 \times 4}[\xi] \) and \( \mathcal{B} \in \mathcal{L}^4 \).

A property which turns out to be of crucial importance in the following is that a one-to-one relationship can be established between behaviors in \( \mathcal{L}^n \) and submodules of \( \mathbb{R}[\xi_1, \xi_2, \ldots, \xi_n] \). In order to keep this work self-contained we thus dedicate a section to recalling some basic algebraic concepts before investigating how they relate to our description of differential systems.

### 2.3 Algebraic intermezzo

In the best tradition of algebraic literature, this section mostly contains a series of definitions and examples, yielding the building blocks of what we discuss in the following. Our aim is not to provide a thorough treatment of the concepts we introduce, but rather to single out what are the main tools we use in the rest of this work. For more details, the reader is referred to the vast literature on algebra, among which we cite [3] as a good introductory reference to the subject, and [46] as a more complete and advanced one.

#### 2.3.1 Rings and ideals

The very first building block we introduce is a ring. Notice that the definition we give actually corresponds to the subclass of commutative rings with unit element; this is however the subclass we shall be interested into in the rest of this work and therefore, with slight abuse of notation, we refer to it simply as ring.
2.3. Algebraic intermezzo

**Definition 24**: A ring is a set $\mathcal{R}$ endowed with two binary composition laws $+: \mathcal{R} \times \mathcal{R} \rightarrow \mathcal{R}$ (addition) and $\cdot: \mathcal{R} \times \mathcal{R} \rightarrow \mathcal{R}$ (multiplication) such that

1. $\forall x, y, z \in \mathcal{R} \ (x + y) + z = x + (y + z)$ and $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ (Associativity)
2. $\forall x, y \in \mathcal{R} \ x + y = y + x$ and $x \cdot y = y \cdot x$ (Commutativity)
3. $\forall x, y, z \in \mathcal{R} \ x \cdot (y + z) = x \cdot y + x \cdot z$ (Addition distributive over multiplication)
4. $\exists 0 \in \mathcal{R}$ s.t. $\forall x \in \mathcal{R} \ x + 0 = x$ (Identity element for addition)
5. $\exists 1 \in \mathcal{R}$ s.t. $\forall x \in \mathcal{R} \ 1 \cdot x = x$ (Identity element for multiplication)
6. $\forall x \in \mathcal{R} \exists -x \in \mathcal{R}$ s.t. $x + (-x) = 0$ (Inverse for addition)

In the following, we drop the explicit indication of $\cdot$ for multiplication and replace the notation $x \cdot y$ by $xy$. Notice that in a ring every element has an inverse with respect to addition, whereas this is not true with respect to multiplication. If a given $x \in \mathcal{R}$ admits an inverse element with respect to multiplication, in other words if there exists $x^{-1} \in \mathcal{R}$ such that $xx^{-1} = 1$, we will say that $x$ is *unit* in $\mathcal{R}$. A ring in which every element except 0 is a unit is called a **field**.

If we define in the logical way subtraction ($-$) and division ($/$) of two elements in a ring by $x - y = x + (-y)$ and $x/y = xy^{-1}$, we see that we can always add, subtract or multiply two elements of a ring, but not necessarily divide them, whereas in a field also this last operation is always well defined unless the divisor is 0.

**Example 25**: By defining addition and product defined in the usual way it is easily seen that the set $\mathbb{Z}$ of integers is a ring, while the sets of real ($\mathbb{R}$), rational ($\mathbb{Q}$) and complex ($\mathbb{C}$) numbers are all fields.

**Example 26**: The ring which will be ubiquitous in the rest of this work is the set of polynomials with real coefficients in the $n$ indeterminates $\xi_1, \ldots, \xi_n$ with addition and multiplication defined in the usual way. Such a set is indicated as $\mathbb{R}[\xi_1, \ldots, \xi_n]$. Polynomials in one indeterminate are indicated by $\mathbb{R}[\xi]$. The real numbers, excluding 0, are the set of units in $\mathbb{R}[\xi_1, \ldots, \xi_n]$.
With the notion of ring at hand we are ready to define the second building block of our construction, namely ideals. These are subsets of a ring, endowed with the special property of being closed under addition and multiplication with arbitrary elements of the ring. More formally

**Definition 27:** An ideal of a ring \( \mathcal{R} \) is a subset \( \mathfrak{I} \subseteq \mathcal{R} \) such that

1. \((a \in \mathfrak{I}, x \in \mathcal{R}) \Rightarrow (ax \in \mathfrak{I})\)
2. \((a, b \in \mathfrak{I}) \Rightarrow (a + b \in \mathfrak{I})\)

**Example 28:** The set of even integers is an ideal of \( \mathbb{Z} \); it is actually the ideal generated by 2, in the sense which we explain next.

Given any subset \( \mathcal{I} \) of a ring \( \mathcal{R} \) we indicate by \(< \mathcal{I}>\) the smallest ideal of \( \mathcal{R} \) which contains \( \mathcal{I} \), and say that \( \mathcal{I} \) is a generating set for \(< \mathcal{I}>\) or, equivalently, that \(< \mathcal{I}>\) is generated by \( \mathcal{I} \). It can be shown that \(< \mathcal{I}>\) = \(\{ \sum_{j \in J} h_j t_j, \ t_j \in \mathcal{I}, \ h_j \in \mathcal{R} \}\) for \( J \) a finite index set. In other words, the ideal generated by a (possibly infinite) set \( \mathcal{I} \subseteq \mathcal{R} \) is the set of all possible finite linear combinations of elements of the generating set with coefficients taken in the ring. If \( \mathcal{I} = \{ t_1, \ldots, t_s \} \subseteq \mathcal{R} \) is a finite subset of \( \mathcal{R} \) we say that \( \mathfrak{I} = < t_1, \ldots, t_s > = \{ \sum_{i=1}^{s} h_i t_i, \ h_i \in \mathcal{R} \} \) is a finitely generated ideal. In the special case that \( \mathcal{I} = t \) is a singleton we say that \( \mathfrak{I} = < t > \) is a principal ideal.

**Example 29:** The principal ideal \( \mathfrak{I} = < \xi > = \{ f \in \mathbb{R}[\xi] \text{ s.t. } f = \xi g, \ g \in \mathbb{R}[\xi] \} \) contains all polynomials in one variable which have \( \xi \) as a factor. More in general, if \( \mathcal{R} = \mathbb{R}[\xi_1, \ldots, \xi_n] \) and \( \mathcal{I} = \{ t_1, \ldots, t_s \} \subseteq \mathbb{R}[\xi_1, \ldots, \xi_n] \) then \( < t_1, \ldots, t_s > = \{ \sum_{i=1}^{s} h_i t_i, \ h_i \in \mathbb{R}[\xi_1, \ldots, \xi_n] \}\). In other words the ideal generated by \( s \) polynomials \( t_1, \ldots, t_s \) is the set of all possible linear combinations with polynomial coefficients of \( t_1, \ldots, t_s \).

The cardinality of the generating sets of ideals entails a very important classification of rings. A ring in which every ideal is finitely generated is called Noetherian, whereas one in which every ideal is principal is called a principal ideal domain (often abbreviated PID). It can be shown that
$$\mathbb{R}[\xi_1, \ldots, \xi_n]$$ is a Noetherian ring (this is the famous Hilbert basis theorem) while $$\mathbb{R}[\xi]$$ is a principal ideal domain. Stated more explicitly this means that given any ideal $$\mathfrak{J} \subseteq \mathbb{R}[\xi_1, \ldots, \xi_n]$$ we can always find $$t_1, \ldots, t_s \in \mathbb{R}[\xi_1, \ldots, \xi_n]$$ such that $$\mathfrak{J} = \langle t_1, \ldots, t_s \rangle$$ as in example 29, whereas given any ideal $$\mathfrak{J} \subseteq \mathbb{R}[\xi]$$ we can always find a single element $$t \in \mathbb{R}[\xi]$$ such that $$\mathfrak{J} = \langle t \rangle$$.

A property of Noetherian rings which is often used in the following is presented in the following famous theorem, also known as the ascending chain condition.

**Theorem 30**: Let

$$\mathfrak{J}_1 \subseteq \mathfrak{J}_2 \subseteq \cdots \subseteq \mathfrak{J}_k \subseteq \cdots$$

be an ascending chain of ideals of a Noetherian ring $$\mathfrak{A}$$. Then there exists an index $$N \geq 1$$ such that $$\mathfrak{J}_i = \mathfrak{J}_{i+1}$$, for $$i \geq N$$.

### 2.3.2 Modules

We dedicate this section to defining the last of the basic building blocks which we need in the following, namely modules over a ring $$\mathfrak{A}$$. The special attention we dedicate to this algebraic structure will be justified in the next section, where we investigate the strong connections between modules over the ring $$\mathbb{R}[\xi_1, \ldots, \xi_n]$$ and dynamical systems described by linear differential equations with constant coefficients.

**Definition 31**: Given a ring $$\mathfrak{A}$$, a module over $$\mathfrak{A}$$ (or $$\mathfrak{A}$$-module) is a set $$\mathfrak{M}$$ for which maps $$+ : \mathfrak{M} \times \mathfrak{M} \to \mathfrak{M}$$ (addition) and $$\cdot : \mathfrak{A} \times \mathfrak{M} \to \mathfrak{M}$$ (multiplication with scalars) are defined, such that

1. $$\forall x, y, z \in \mathfrak{M} \quad (x + y)z = x + (y + z)$$ (Associativity of addition)
2. $$\forall x, y \in \mathfrak{M} \quad x + y = y + x$$ (Commutativity)
3. $$\exists 0 \in \mathfrak{M}$$ s.t. $$\forall x \in \mathfrak{M} \quad x + 0 = x$$ (Identity element for addition)
4. $$\forall x \in \mathfrak{M} \quad \exists -x \in \mathfrak{M}$$ s.t. $$x + (-x) = 0$$ (Inverse for addition)
5. $$\forall x, y \in \mathfrak{M}, \forall a, b \in \mathfrak{A} \quad a(x + y) = ax + ay, \quad (a + b)x = ax + bx$$ (Distributivity)
6. \( \forall x \in \mathcal{M}, \ \forall a, b \in \mathcal{A} \ (ab)x = a(bx) \) (Associativity of product)

7. If 1 is the unit element for the product in \( \mathcal{A} \) then \( \forall x \in \mathcal{M} \ 1x = x \) (Identity element for product)

Notice that we have again dropped the notation \( \cdot \) for multiplication and simply replaced \( a \cdot x \) by \( ax \).

The definition above might be very confusing at first sight. A good way of building up some intuition about it is noticing that familiar vector spaces satisfy exactly the same properties listed above except that the ring of scalars \( \mathcal{A} \) is assumed to be a field (which we know is a special case of a ring). The following example wishes to strengthen this intuition by showing that \( m \)-tuples of polynomials form a module over \( \mathbb{R}[\xi_1, \ldots, \xi_n] \) in much the same way as \( m \)-tuples of real numbers are a vector space over \( \mathbb{R} \).

**Example 32**: Consider the set \( \mathbb{R}^n[\xi_1, \ldots, \xi_n] \) of \( m \)-tuples of real polynomials in \( n \) indeterminates. Each element \( x \in \mathbb{R}^n[\xi_1, \ldots, \xi_n] \) has, therefore, the structure \( (x_1, \ldots, x_m) \), \( x_i \in \mathbb{R}[\xi_1, \ldots, \xi_n] \). Given a \( a \in \mathbb{R}[\xi_1, \ldots, \xi_n] \) we can define the product \( ax \) component-wise as \( ax = (ax_1, \ldots, ax_m) \); given \( x, y \in \mathbb{R}^n[\xi_1, \ldots, \xi_n] \) we can define the sum \( x + y \) again component-wise as \( x + y = (x_1 + y_1, \ldots, x_m + y_m) \). It is then not difficult to recognize that with these two operations \( \mathbb{R}^n[\xi_1, \ldots, \xi_n] \) becomes an \( \mathbb{R}[\xi_1, \ldots, \xi_n] \)-module.

Notice how any ideal \( \mathcal{J} \) of \( \mathcal{A} \) is naturally an \( \mathcal{A} \)-module.

A *submodule* \( \mathcal{S} \) of an \( \mathcal{A} \)-module \( \mathcal{M} \) is a subset \( \mathcal{S} \subseteq \mathcal{M} \) which is itself endowed with a module structure. Given any subset \( \mathcal{J} \) of an \( \mathcal{A} \)-module \( \mathcal{M} \) we indicate by \( < \mathcal{J} > \) the smallest submodule of \( \mathcal{M} \) which contains \( \mathcal{J} \) and say that \( \mathcal{J} \) is a *generating set* for \( < \mathcal{J} > \) or, equivalently, that \( < \mathcal{J} > \) is *generated* by \( \mathcal{J} \). Much as in the case of ideals it can be shown that \( < \mathcal{J} > = \{ \sum_{j \in J} h_jt_j, \ t_j \in \mathcal{J}, \ h_j \in \mathcal{A} \} \) for \( J \) a finite index set. In other words the submodule of \( \mathcal{M} \) generated by a (possibly infinite) set \( \mathcal{J} \subseteq \mathcal{M} \) is the set of all possible finite linear combinations of elements of the generating set with coefficients taken in the ring.

**Example 33**: Consider the module \( \mathbb{R}^n[\xi_1, \ldots, \xi_n] \) and let \( \mathcal{J} = \{t_1, \ldots, t_s\} \subseteq \mathbb{R}^n[\xi_1, \ldots, \xi_n] \). Then \( t_1, \ldots, t_s \geq \{ \sum_{j=1}^s h_jt_j, \ h_j \in \mathbb{R}[\xi_1, \ldots, \xi_n] \} \). The submodule of \( \mathbb{R}^n[\xi_1, \ldots, \xi_n] \) generated by the \( s \) vector polynomials \( t_1, \ldots, t_s \) is therefore the set of all possible linear combinations with polynomial coefficients of \( t_1, \ldots, t_s \).
2.3. Algebraic intermezzo

As in the case of ideals, an \( \mathfrak{R} \)-module \( \mathfrak{M} \) is said to be \textit{finitely generated} if there exists a finite set \( \mathcal{I} = \{t_1, \ldots, t_s\} \subseteq \mathfrak{M} \) such that \( \mathfrak{M} = \langle t_1, \ldots, t_s \rangle = \{ \sum_{i=1}^{s} h_i t_i, h_i \in \mathfrak{R} \} \). A result of crucial importance in the following is that, as a consequence of \( \mathbb{R}[\xi_1, \ldots, \xi_n] \) being a Noetherian ring, it can be shown that every submodule of \( \mathfrak{M} \subseteq \mathbb{R}^n [\xi_1, \ldots, \xi_n] \) is finitely generated.

A subset \( \mathcal{I} \) of an \( \mathfrak{R} \)-module is said to be \textit{independent} if, whenever \( J \) is a finite indexing set,

\[
\left( \sum_{j \in J} h_j t_j = 0, h_j \in \mathfrak{R}, t_j \in \mathcal{I} \right) \iff (h_j = 0, j \in J)
\]

in other words the only way in which \( 0 \) can be obtained as a linear combination of elements of \( \mathcal{I} \) is by choosing all the coefficients in \( \mathfrak{R} \) equal to zero. \( \mathfrak{R} \)-modules \( \mathfrak{M} \) generated by independent sets are said to be \textit{free}. It is easily seen that if \( \mathfrak{M} = \langle \mathcal{I} \rangle \) and \( \mathcal{I} \) is independent, every element \( m \in \mathfrak{M} \) can be written in a unique way as a linear combination of elements of \( \mathcal{I} \). Given the paramount importance this plays in the sequel, we recall (see [46] for a proof) that as a consequence of \( \mathbb{R}[\xi] \) being a PID it follows that every submodule of \( \mathbb{R}[\xi]^m \) is free (and also finitely generated as seen above).

This is not the case in the general case of \( \mathbb{R}[\xi_1, \xi_2, \ldots, \xi_n] \)-modules which are, in general, not free. This difference between the ring of polynomials in one and several variables is crucial and accounts for many of the differences in the treatment of ordinary and partial differential equations presented in the following. Given its relevance for our purposes we illustrate it with an example.

\textbf{Example 34 :} The two polynomials \( t_1 = \xi_1 \) and \( t_2 = \xi_2 \) are not independent because \( h_1 t_1 + h_2 t_2 = 0 \Leftrightarrow h_1 = \xi_2 r(\xi_1, \xi_2), h_2 = -\xi_1 r(\xi_1, \xi_2), \forall r(\xi_1, \xi_2) \in \mathbb{R}[\xi_1, \xi_2] \). Simple computations show that there cannot exist a single polynomial \( p(\xi_1, \xi_2) \) such that \( < p > = < t_1, t_2 > \) and thus that \( < \xi_1, \xi_2 > \) cannot be generated by an independent set of elements.

A submodule of \( \mathfrak{M} \subseteq \mathbb{R}^n [\xi_1, \xi_2, \ldots, \xi_n] \) admits different sets of generators. Any such set, however, must contain a minimal number of elements, which we indicate by \( \pi(\mathfrak{M}) \). Any set of generators of \( \mathfrak{M} \) containing exactly \( \pi(\mathfrak{M}) \) elements is called a \textit{minimal} set of generators of \( \mathfrak{M} \). In the special case of submodules of \( \mathbb{R}[\xi] \), it is easily seen that sets of generators are minimal if and only if they are independent. As shown by example 34 this is not the
case for submodules of $\mathbb{R}^n[\xi_1, \xi_2, \ldots, \xi_n]$ in which case establishing $\pi(M)$ for a given $M$ is a very complex matter.

We close this section by presenting a last concept related to modules, namely that of quotient module, which turns out to be useful in the following chapters.

Any submodule $S$ of an $\mathcal{M}$-module $M$ induces in a natural way an equivalence relation on $M$. If $f, g \in M$ we say that they are equivalent modulo $S$ and write $g \equiv f \pmod{S}$ if $f - g \in S$. For any $f \in M$ the equivalence class of $f$ modulo $S$, indicated as customary by $[f]$, will therefore be the set of all elements of $M$ which can be obtained by adding to $f$ any element belonging to $S$. The set of all equivalence classes modulo $S$ is indicated by $M/S$. In other words:

$$M/S = \{[f] : f \in M\}.$$ 

On $M/S$ we can define addition as $[f] + [g] = [f + g]$ and product by a scalar as $a[f] = [af]$, $\forall a \in \mathcal{M}, f \in M$ in such a way that $M/S$ is itself endowed with a module structure over $\mathcal{M}$; it is therefore referred to as the Quotient Module of $M$ with respect to $S$.

### 2.3.3 Polynomial matrices

In the sequel, we often use the matrix notation $< T >$ to indicate the submodule of $\mathbb{R}^m[\xi_1, \ldots, \xi_n]$ generated by the columns of the polynomial matrix $T \in \mathbb{R}^{m \times s}[\xi_1, \xi_2, \ldots, \xi_n]$.

As a trivial consequence of submodules of $\mathbb{R}^m[\xi_1, \ldots, \xi_n]$ being finitely generated it follows that given any such submodule $M$, we can always find a matrix $T \in \mathbb{R}^{m \times s}[\xi_1, \xi_2, \ldots, \xi_n]$ such that $M = < T >$. Such a matrix is, of course, not unique; in particular, two polynomial matrices $T_1$ and $T_2$ are such that $< T_1 > = < T_2 >$ if and only if there exist other polynomial matrices $U_1$ and $U_2$ such that $T_1 = T_2 U_2$ and $T_2 = T_1 U_1$. In the following, polynomial matrices $T_1$ and $T_2$ with the property that $< T_1 > = < T_2 >$ are called equivalent; it is easily seen that this is indeed an equivalence relation in the formal sense of the word (i.e., transitive, reflexive, and symmetric).

The rank of $T \in \mathbb{R}^{m \times s}[\xi_1, \xi_2, \ldots, \xi_n]$ is defined as the maximum number of independent columns of $T$. If $\text{rank}(T) = p$ then $T$ contains a $p \times p$ sub-matrix whose determinant is not the zero polynomial; square sub-matrices of
size greater than \( p \), instead, all have a zero determinant. As a consequence of submodules of \( \mathbb{R}[\xi] \) being free, it follows that given any such submodule \( \mathfrak{M} \), we can always find a matrix \( T \in \mathbb{R}^{m \times p}[\xi] \) of full column rank (i.e. \( \text{rank}(T) = p \)) such that \( \mathfrak{M} = \langle T \rangle \). Obviously \( p = \mathfrak{m}(\mathfrak{M}) \), the minimal number of generators of the submodule \( \mathfrak{M} \). As a further consequence, we obtain that given any polynomial matrix \( T_1 \) we can always find polynomial matrices \( T_2 \), \( U_1 \) and \( U_2 \) with \( T_2 \) of full column rank such that \( T_1 = T_2 U_2 \) and \( T_2 = T_1 U_1 \). In particular if \( T_1 \) is itself of full column rank, then \( U_1 U_2 = U_2 U_1 = I \) with \( I \) the identity matrix of suitable dimension. This is equivalent to saying that \( U_1 \) and \( U_2 \) are unimodular polynomial matrices, namely matrices that admit a polynomial inverse or, equivalently, such that their determinant is a real non-zero constant.

### 2.4 Modules and linear differential systems

The following result which establishes the correspondence between behaviors and submodules of \( \mathbb{R}^2[\xi_1, \xi_2, \ldots, \xi_n] \) was proven in [54] for behaviors in \( \mathcal{L}^2 \) and in [51] for the general case of \( \mathcal{L}_n^2 \) (see also [52] and [71]).

**Theorem 35**: Let \( R_1, R_2 \in \mathbb{R}^{n \times m}[\xi_1, \xi_2, \ldots, \xi_n] \). Then

\[
\mathfrak{B} = \ker(R_1(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n})) = \ker(R_2(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n})) \iff \langle R_1^T \rangle = \langle R_2^T \rangle
\]

The above theorem thus states that a same behavior admits many (actually infinitely many) different kernel representations. It formalizes at a very general and abstract level (at least for differential systems) the intuitive remark of section 1.2 on the possibility of describing the same behavior using different representations (i.e. different equations) involving different parameters (i.e. coefficients of the equations).

All kernel representations of a same behavior, though, have the property that the transposes of the rows of the corresponding polynomial matrices generate the same module. For the special case of behaviors in \( \mathcal{L}^2 \) this implies that we can always find a full row rank polynomial matrix \( R \) such that \( \mathfrak{B} = \ker(R(A)) \); any such matrix is said to induce a minimal kernel representation of the corresponding \( \mathfrak{B} \). At a more technical level it should
be remarked that the one-to-one correspondence between modules and behaviors established in theorem 35 depends crucially on the fact that we only consider $C^\infty$ or distributional solutions. As explained in detail in [62], this correspondence does not hold in case we consider other solution spaces, such as, for example, $C^\infty$ solutions with compact support.

The attention given to properties of solutions is what essentially differentiates the point of view followed in this work from that presented in [23], [57] and related publications. These references, in fact, provide very interesting applications of powerful tools from algebra, in particular differential and homological algebra, to the study of systems described by differential equations. The approach taken, however, is entirely formal, in other words properties of the equations are studied without reference to the solutions of the equations themselves. In this respect, these works are far in spirit from the point of view taken in this work, following the references cited in the introduction, where the solutions of given equations (i.e. the behavior) are the central object of study.

Given any $\mathcal{B} \in \mathcal{E}^*_{\mathcal{F}}$ we can define the set $\mathcal{N}_\mathcal{B}$ of consequences of $\mathcal{B}$ as

$$\mathcal{N}_\mathcal{B} = \{ u \in \mathbb{R}^n| \xi_1, \xi_2, \ldots, \xi_n | u^T((\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}))\mathcal{B} = 0 \}$$

It is easily seen that $\mathcal{N}_\mathcal{B}$ is actually a submodule of $\mathbb{R}^n| \xi_1, \xi_2, \ldots, \xi_n |$ and that $< R^T > \subseteq \mathcal{N}_\mathcal{B}$ for any $R$ such that $\mathcal{B} = \ker(R(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}))$. A crucial result from [51] is that $< R^T > = \mathcal{N}_\mathcal{B}$; this property is again fundamentally dependent on the choice of space of trajectories (infinitely differentiable functions or distributions).

Theorem 35 establishes the link between trajectories of dynamical systems on the one hand and algebraic objects on the other, that is crucial to the development of this work, in particular because it opens the way to exploiting algorithms from computer algebra for the analysis of systems. References [23], [57] and related publications also present very interesting applications of powerful tools from algebra, in particular differential and homological algebra, to the study of systems described by differential equations. The approach taken, however, is entirely formal, in other words properties of the equations are studied without reference to the solutions of the equations themselves. In this respect, these works are far in spirit from the point of view taken in this work, where the solutions of given equations (i.e. the behavior) are the central object.
2.5 Conclusions and further research

The main purpose of this chapter has been providing a rigorous mathematical framework to the language for modeling systems which had been introduced in the first chapter. To this aim we introduced the abstract concept of behavior of a system, defined as the set of all possible trajectories of the variables we are modeling. Through various examples we showed how behaviors can be specified in many ways; in particular we choose to concentrate on models whose trajectories are described as solutions to systems of constant coefficient linear partial differential equations. The main result we presented regarding this special class of models is the one-to-one correspondence between modules over the polynomial ring \( \mathbb{R}[\xi_1, \xi_2, \ldots, \xi_n] \) and linear shift invariant behaviors described in section 2.4. The connection thus established between dynamical systems on the one hand and algebraic objects on the other, provides the crucial link we need in the following chapters to exploit algorithms from computer algebra for the analysis of systems.

Further research on the problems discussed in this chapter could be to investigate how far a correspondence as the one discussed in section 2.4 can be established for other classes of systems. In particular, time varying systems, non linear systems and distributed systems with finite boundaries (i.e. systems described by partial differential equations for which the indexing set \( I \) is a proper subset of \( \mathbb{R}^n \)) are all classes of great interest in applications. For the case of time varying systems a start has been made in [26], where a one to one correspondence is established between modules over a skew polynomial ring and solutions to systems of ordinary linear differential equations with variable coefficients, for the case in which solutions belong to the space of hyperfunctions.