Chapter 4

The elimination problem

The framework for modeling systems introduced in chapter 1 naturally resulted in models which, in first instance, included both the variables one is interested in (called manifest), and extra variables arising as a byproduct of the modeling procedure (called latent). Already at that stage, the rather natural question arose whether one could start from a description including both manifest and latent variables and reduce it to one that only uses the manifest ones.

In this chapter we first give a formal definition of system with latent variables. For the class of systems discussed in this work (i.e. linear and differential) we then formulate and constructively solve the elimination problem, namely how differential equations that only involve the manifest variables can be obtained starting from the original equations involving both manifest and latent variables.

Before proceeding, let us remark that elimination theory for systems of algebraic equations is a classical and central topic in algebraic geometry (see e.g. [19], [20], [40]). In the case of linear partial differential equations, instead, the question has arisen, in a slightly different set up, as that of finding conditions for solvability of given systems of equations (see e.g. [37], [55]). Formulating elimination as a question naturally arising in modeling of physical systems, is one of the new problems introduced in systems' theory by the behavioral approach that we have been using all along.
4.1 Dynamical systems with latent variables

We begin by giving a general definition of a dynamical system with latent variables:

**Definition 72**: A dynamical system with latent variables is a quadruple $\Sigma_L = (\mathbb{I}, \mathbb{W}, \mathbb{L}, \mathbb{B}_{\text{full}})$, with $\mathbb{I}$ the index set, $\mathbb{W}$ the manifest signal space, $\mathbb{L}$ the latent variable space, and $\mathbb{B}_{\text{full}} \subseteq \mathbb{W} \times \mathbb{L}$ the full behavior of the system.

Consequently, the trajectories of a system with latent variables are pairs $(w, \ell)$, with $w$ the manifest variable trajectory, and $\ell$ the latent trajectory. A dynamical system with latent variables induces a dynamical system in the sense of definition 16 as follows.

**Definition 73**: Let $\Sigma_L = (\mathbb{I}, \mathbb{W}, \mathbb{L}, \mathbb{B}_{\text{full}})$ be a dynamical system with latent variables. The manifest dynamical system induced by $\Sigma_L$ is the dynamical system $\Sigma = (\mathbb{I}, \mathbb{W}, \mathbb{B})$, with the behavior $\mathbb{B}$ defined as

$$\mathbb{B} := \{w : \mathbb{I} \to \mathbb{W} | \exists \ell : \mathbb{I} \to \mathbb{L} \text{ such that } (w, \ell) \in \mathbb{B}_{\text{full}}\}$$

and can be regarded as the projection of the full behavior onto the manifest variables.

The notions of linearity and time invariance of a latent variable system are an obvious extension of those given in the case of dynamical systems, and we will not write them down explicitly. Chapter 1 presents a lot of cases exemplifying the above definitions, all of them belonging to the special class of differential systems. Given the central position occupied by such systems in the present work, we proceed to specialize definitions 72 and 73 for this class.

A differential system with latent variables is a system with $\mathbb{I} = \mathbb{R}^n$ whose full behavior $\mathbb{B}_{\text{full}}$ consists of all solutions $(w, \ell)$ of a set of partial differential equations of the form:

$$f(x, w(x), \ldots, \frac{\partial^k}{\partial x^k} w(x), \ldots \ell(x), \ldots, \frac{\partial^j}{\partial x^j} \ell(x)) = 0$$

(4.1)

with $k$ a multi-index as in chapter 2. It is assumed that only a finite number of partial derivatives are involved in the above equations, say $d_i$ derivatives
of the $w$ and $d_2$ derivatives of the $\ell$ variables. $f$ is, therefore, a function from $\mathbb{R}^{d_1+w+d_2+1+n}$ to $\mathbb{R}^p$, with $p$ the number of differential equations defining the system, $w$ and $1$ the cardinality of vectors $w$ and $\ell$ respectively.

An ordinary differential system with latent variables is, of course, a system with $I = \mathbb{R}$ and full behavior $\mathcal{B}_{\text{full}}$ consisting of all solutions $(w, \ell)$ of a finite set of ordinary differential equations of the form
\[
f(t, w(t), \frac{d}{dt} w(t), \ldots, \frac{d^d}{dt^d} w(t), \ell(t), \frac{d}{dt} \ell(t), \ldots, \frac{d^N}{dt^N} \ell(t)) = 0.
\]
with $f$ a function from $\mathbb{R}^{d_1+w+N+1}$ to $\mathbb{R}^p$, with $p$ the number of differential equations defining the system, $w$ and 1 the cardinality of vectors $w$ and $\ell$ respectively.

Instances of equations (4.1) are provided by the model of the solid bar in example 9 and the two link arm in example 13.

As in chapter 2, we concentrate on a special subclass of models (4.1), namely linear constant coefficient partial differential systems. In this case we deal with systems with $I = \mathbb{R}^n$, $W = \mathbb{R}^q$ $L = \mathbb{R}^1$ and
\[
\mathcal{B}_{\text{full}} = \{(w, \ell) \in C^\infty(\mathbb{R}^n, \mathbb{R}^{n+1}) \text{ such that } N(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}) w = M(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}) \ell \}
\]
for the distributional counterpart
\[
\mathcal{B}_{\text{full}} = \{(w, \ell) \in D'(\mathbb{R}^n, \mathbb{R}^{n+1}) \text{ such that } N(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}) w = M(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}) \ell \}
\]
with $N \in \mathbb{R}^{p \times n}$, $M \in \mathbb{R}^{p \times 1}$ polynomial matrices in $n$ indeterminates with the same, finite but otherwise arbitrary number of rows. $N(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n})$ and $M(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n})$ are, therefore, linear constant coefficient partial differential operators. Notice that $\mathcal{B}_{\text{full}} = \ker(N(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}) - M(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}))$, equivalently $\mathcal{B}_{\text{full}} \subset \mathcal{L}^{n+1}$.

Example 74: In example 22 we regarded Maxwell’s equations as defining a manifest behavior for variables $w = (E, B, \mathcal{J}, \rho) \in \mathbb{R}^8$. Another possible way of regarding them, however, is as a system with latent variables $\ell = (\mathcal{B}, \rho)$ and manifest $w = (E, \mathcal{J})$, corresponding to a situation in which we are interested in the evolution of electrical and current field and introduce
magnetic field and charge density to help us describe \( \vec{E} \) and \( \vec{j} \). In this case the full behavior can be written as in 4.2 with

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

**Example 75**: Considering, as done in example 14, \( w = (w_1, F_1) \) and \( \ell = (w_2, F_2) \), equations 1.8 describing the interconnection of two membranes can be written in the form 4.2 with

\[
N = \begin{bmatrix}
\rho_1 \xi_3^2 & -(\lambda_1 + 2\mu_1)\xi_1^2 - \mu_1 \xi_2^2 \\
-(\lambda_1 + \mu_1)\xi_1 \xi_2 & \rho_1 \xi_3^2 & -(\lambda_1 + 2\mu_1)\xi_1 \xi_2 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}, M = \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & -1
\end{bmatrix}
\]

As usual, a very important special case of 4.2 is obtained when considering linear constant coefficient ordinary differential systems with latent variables.
In this case we are dealing with equations

\[ N_0 w + N_1 \frac{d}{dt} w + \cdots + N_L \frac{d^L}{dt^L} w = M_0 \ell + M_1 \frac{d}{dt} \ell + \cdots + M_N \frac{d^N}{dt^N} \ell \]  

(4.4)

for suitable real matrices \( N_i \in \mathbb{R}^{p \times p} \) and \( M_i \in \mathbb{R}^{p \times 1} \). Defining \( N = N_0 + N_1 + \cdots + N_L \xi^L \in \mathbb{R}^{p \times p} \) and \( M = M_0 + M_1 + \cdots + M_N \xi^N \in \mathbb{R}^{p \times 1} \) we find as in (4.2)

\[ \mathfrak{B}_{\text{full}} = \{(w, \ell) \in C^\infty(\mathbb{R}, \mathbb{R}^{p+1}) \mid N \left( \frac{d}{dt} \right) w = M \left( \frac{d}{dt} \right) \ell \} \]

or the distributional version as in 4.3. Needless to say, in this case \( \mathfrak{B}_{\text{full}} \in \mathbb{D}^{\ast+1} \).

**Example 76:** The equations of the circuit in example 1.13 with the choice of manifest variables \( w = (V_9, I_9, V_{12}, I_{12}) \) correspond to

\[
N = \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}
\]
\[ M = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
1 & -R_c & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0
0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0
0 & 0 & 0 & 0 & 1 & -R_L & -1 & 0 & 0 & 0
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
0 & 0 & 0 & 0 & 0 & 0 & 0 & C\xi & -1 & -C\xi
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}\]
4.2 The fundamental principle

If we apply Definition 73 to the special case of partial differential systems we obtain that the manifest behavior \( \mathcal{B} \) associated to the system with latent variables (4.2) is defined as

\[
\mathcal{B} = \{ w \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}^N) \mid \exists \ell \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}^1) \text{ such that (4.2) holds} \}
\]

The distributional counterpart, as usual, follows from (4.3).

At a number of points in Chapter 1, the question arose whether we could start from a description of a system with latent variables and obtain from it a description of the manifest behavior which only involved the manifest variables. In the examples we always managed to do so, and it is reasonable to ask the question whether this is a general property. It is, in fact, at least for linear differential systems as we now show by providing a very crisp way in which equations for \( \mathcal{B} \) can be deduced from the equations defining \( \mathcal{B}_{\text{full}} \).

In order to do so, we first need to define the set \( \text{SYZ}(T) \) of syzygies of a polynomial matrix \( T \in \mathbb{R}^{p \times s}[\xi_1, \ldots, \xi_n] \) as:

\[
\text{SYZ}(T) = \{ h \in \mathbb{R}^{s}[\xi_1, \ldots, \xi_n] : Th = 0 \}
\]

\( \text{SYZ}(T) \) is, thus, the set of all polynomial vectors solving the equation \( Th = 0 \). It is easily seen that \( \text{SYZ}(T) \subseteq \mathbb{R}^{s}[\xi_1, \ldots, \xi_n] \) is a module over \( \mathbb{R}[\xi_1, \ldots, \xi_n] \) and is also referred to as the \textit{syzygy module} of \( T \) (see [46], [1]). We can then state the following result

**Theorem 77**: Given \( w \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}^N) \) (resp. \( w \in \mathcal{D}'(\mathbb{R}^n, \mathbb{R}^N) \)) and \( M \in \mathbb{R}^{s \times 1}[\xi_1, \ldots, \xi_n] \) there exists \( \ell \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}^1) \) (resp. \( \ell \in \mathcal{D}'(\mathbb{R}^n, \mathbb{R}^1) \)) such that \( w = M \left( \frac{\partial}{\partial \xi_1}, \frac{\partial}{\partial \xi_2}, \ldots, \frac{\partial}{\partial \xi_n} \right) \ell \) if and only if \( h^T \left( \frac{\partial}{\partial \xi_1}, \frac{\partial}{\partial \xi_2}, \ldots, \frac{\partial}{\partial \xi_n} \right) w = 0 \) for all \( h \in \text{SYZ}(M^T) \).

Theorem 77 is a central result in the theory of linear partial differential equations and is sometimes referred to as the \textit{fundamental principle}, hence the title of this section. The necessity part of the above theorem is rather straightforward to see; the proof of sufficiency for the general case of partial differential equations requires deep arguments from complex analysis and homological algebra, and amounts to showing that the signal spaces \( \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}^N) \) and \( \mathcal{D}'(\mathbb{R}^n, \mathbb{R}^N) \) are injective \( \mathbb{R}[\xi_1, \ldots, \xi_n] \)-modules (see [36], [51] and references therein).
As an easy consequence of theorem 77 and of the definition of manifest behavior given at the beginning of this section we obtain the following theorem that effectively answers the question of characterizing the manifest behavior corresponding to a given full behavior.

**Theorem 78:** Let $\mathcal{B}$ be the manifest behavior corresponding to the latent variable representation $N\left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right)w = M\left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right)\ell$. Then the following are equivalent

(i) $w \in \mathcal{B}$

(ii) $h \in SYZ(M^T) \Rightarrow h^T \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right) N\left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right)w = 0$

In section 4.4 we present a direct proof of theorem 78 for the special case of ordinary differential systems, using more elementary arguments than those needed to prove theorem 77. We wish, however, to stress how the result crucially depends on the chosen solution spaces ($C^\infty$ functions or distributions), already at the level of ODE’s. The theorem is, for example, not true if one looks at solutions which are locally integrable or compact support functions (see [53]).

As a consequence of the module structure of $SYZ(M^T)$, it follows from theorem 78 that

$$(w \in \mathcal{B}) \Leftrightarrow (H^T \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right) N\left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right)w = 0.$$ 

Here $H$ is any polynomial matrix such that $<H> = SYZ(M^T)$. In other words $\mathcal{B} = \ker(H^T \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right) N\left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right))$. The importance of this remark is crucial, showing that if we start with the solution set to a system of linear differential equations, and project it on a subset of the variables involved, then what we come up with can again be written as the solution set of a system of linear differential equations. Otherwise stated, the manifest behavior corresponding to a latent variable model which is linear and differential ($\mathcal{B}_{full} \in \mathcal{L}_n^{+1}$), is again linear and differential ($\mathcal{B} \in \mathcal{L}_n^{+1}$). This justifies the fact that theorem 78 is often referred to as the elimination theorem. Its importance from the point of view of modeling is crucial; in particular it shows that kernel representations are not only important as a way of defining what a linear system is, as done in chapter 2, but also as
the effective end result of the procedure of modeling a linear time invariant system, possibly after eliminating auxiliary variables.

Latent variables have emerged in chapter 1 as a useful instrument when modeling systems. In the sequel it will become clear that, together with the elimination theorem, they also provide a very powerful instrument when it comes to manipulation and analysis of linear differential equations and their solutions.

4.3 Algorithms for Elimination

The final remarks of the previous section have, hopefully, convinced the reader about the interest of theorem 78. As we now proceed to show, such interest is increased by the fact that it effectively provides an algorithm to build equations describing the manifest behavior.

4.3.1 Building Syzygies

We now describe an algorithm that exploits Gröbner bases in order to compute a set of generators for the module SYZ(T) of a given polynomial matrix $T \in \mathbb{R}^{p \times s}[\xi_1, \ldots, \xi_n]$. We first state a result of general interest regarding Gröbner bases with respect to product orderings (see example 44).

**Lemma 79**: Let $H \in \mathbb{R}^{[p+s] \times s}[\xi_1, \ldots, \xi_n]$ and define the module $\mathcal{H} = \{ q \in \mathbb{R}^s[\xi_1, \ldots, \xi_n] | \begin{pmatrix} 0 \\ q \end{pmatrix} \in < H > \}$. Given monomial orderings for $\mathbb{N}_n^s$ and $\mathbb{N}_n^s$, let $\mathcal{S}$ be a Gröbner basis for $< H >$ with respect to the induced product ordering for $\mathbb{N}_n^s$. Then $\mathcal{C} = \{ c \in \mathbb{R}^s[\xi_1, \ldots, \xi_n] | \begin{pmatrix} 0 \\ c \end{pmatrix} \in \mathcal{S} \}$ is a Gröbner basis for $\mathcal{H}$ with respect to the given ordering on $\mathbb{N}_n^s$.

We can now state the key result we need in the following

**Proposition 80**: Let $T \in \mathbb{R}^{[p+s] \times s}[\xi_1, \ldots, \xi_n]$ and $\hat{T} = \begin{pmatrix} T \\ I_{s \times s} \end{pmatrix} \in \mathbb{R}^{[p+s] \times s}[\xi_1, \ldots, \xi_n]$. If $G \in \mathbb{R}^{[p+s] \times s}[\xi_1, \ldots, \xi_n]$ is a matrix whose columns
are a Gröbner basis for $<\hat{T}>$ with respect to the product ordering on $M_n^{p+s}$ induced by given orderings on $M_n^p$ and $M_n^s$, then, modulo a permutation of columns

$$G = \begin{bmatrix} G_1 & 0 \\ G_2 & H \end{bmatrix}$$

with $H \in \mathbb{R}^{s \times \bullet}[\xi_1, \ldots, \xi_n]$ and $<H> = SYZ(T)$

Proposition 80 follows easily from proposition 79 and the trivial remark $h \in SYZ(T) \Leftrightarrow \begin{pmatrix} 0 \\ h \end{pmatrix} \in <\hat{T}>$. The content of proposition 80 is summarized in the following algorithm that computes a matrix $H$ generating $SYZ(T)$, with $T$ an input matrix

**Algorithm 81:**

1. $H = \text{Syzygy}(T)$;
2. $\hat{T} = \begin{pmatrix} T & \mathbf{I}_{s \times s} \end{pmatrix}$;
3. mo=Product ordering on $M_n^{p+s}$;
4. $G = \text{Gröbner}(\hat{T})$;
5. $H = \text{Select columns of } G$ with first $p$ components equal to zero;

By going through the proof of lemma 79 it is evident that the output $H$ is actually a Gröbner basis for $SYZ(T)$ with respect to the given ordering on $M_n^s$. Also notice that algorithm 81 reduces the computation of $SYZ(T)$ to a Gröbner basis computation. An alternative, somewhat more classical, way of computing $SYZ(T)$ using Gröbner basis is described in [1], [20], [40]. We chose not to present it because its derivation is more involved and because its performance also appears to be worse than that of algorithm 81. Both Singular and the upcoming version of CoCoA, for example, implement Syzygy computations based on the idea sketched in algorithm 81.

Using algorithm 81 and the elimination theorem, we can then design an algorithm that has polynomial matrices $N$ and $M$ as input and outputs a polynomial matrix $R$ such that $\mathfrak{B} = \ker(R(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}))$ is
the manifest behavior corresponding to the latent variable representation
\( N(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \cdots, \frac{\partial}{\partial x_n})w = M(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \cdots, \frac{\partial}{\partial x_n})T. \)

**Algorithm 82:**

\[ R = \text{Elimination}(N, M); \]
\[ H = \text{Syzygy}(M^T); \]
\[ R = H^T N; \]

**Example 83:** A set of generators for \( \text{SYZ}(M^T) \) with \( M \) given in example 74 is given

\[
H = \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & \xi_3 & -\xi_2 \\
0 & -\xi_3 & 0 & \xi_1 \\
0 & \xi_2 & -\xi_1 & 0 \\
0 & 0 & 0 & 0 \\
\xi_1 & \xi_4 & 0 & 0 \\
\xi_2 & 0 & \xi_4 & 0 \\
\xi_3 & 0 & 0 & \xi_4 \\
\end{bmatrix}
\]

\( H \) actually is the Gröbner basis for \( \text{SYZ}(M^T) \) with respect to TOP, lexicographic ordering of \( M^T \). One then obtains

\[
R = H^T N = \begin{bmatrix}
\epsilon_0 \xi_1 \xi_4 & \epsilon_0 \xi_2 \xi_4 & \epsilon_0 \xi_3 \xi_4 \\
-\epsilon_0 \epsilon_0 \xi_1 \xi_2 & \epsilon_0 \epsilon_0 \xi_1 \xi_3 + \epsilon_0 \xi_4^2 & \epsilon_0 \epsilon_0 \xi_2 \xi_3 \\
\epsilon_0 \epsilon_0 \xi_1 \xi_3 & \epsilon_0 \epsilon_0 \xi_2 \xi_3 \\
\epsilon_0 \xi_3 \xi_4 & \xi_1 & \xi_2 & \xi_3 \\
\epsilon_0 \xi_4 \xi_3 & \xi_4 & 0 & 0 \\
\epsilon_0 \epsilon_0 \xi_2 \xi_3 & \xi_4 & 0 & 0 \\
-\epsilon_0 \epsilon_0 \xi_1 \xi_2 + \epsilon_0 \xi_4^2 & 0 & 0 & \xi_4 \\
\end{bmatrix}
\]

Corresponding to equations

\[
\epsilon_0 \frac{\partial}{\partial r} \nabla \cdot \vec{E} + \nabla \cdot \vec{j} = 0 \\
\epsilon_0 \frac{\partial}{\partial r} \dot{\vec{E}} + \epsilon_0 \epsilon_0 \nabla \times \nabla \times \dot{\vec{E}} + \frac{\partial}{\partial r} \dot{\vec{j}} = 0
\]

describing the manifest behavior of \( w = (\vec{E}, \vec{j}) \).
Given a polynomial matrix \( T \in \mathbb{R}^{p \times s} [\xi_1, \ldots, \xi_n] \), alongside its Syzygy module it also makes sense to define its null space \( \text{NULL}(T) \) as:

\[
\text{NULL}(T) = \{ h \in \mathbb{R}^s (\xi_1, \ldots, \xi_n) \mid Th = 0 \}
\]

\( \text{NULL}(T) \) is, thus, the set of all vectors of rational functions solving the equation \( Th = 0 \) and is easily seen to be a vector space over the field \( \mathbb{R}(\xi_1, \ldots, \xi_n) \). It is also straightforward to realize that \( \text{SYZ}(T) \subseteq \text{NULL}(T) \) and, more in particular, that \( \text{NULL}(T) = \langle \text{SYZ}(T) \rangle_{\mathbb{R}(\xi_1, \ldots, \xi_n)} \), where the notation \( \langle f_1, \ldots, f_s \rangle_{\mathbb{R}(\xi_1, \ldots, \xi_n)} \) stands for the vector space over the field \( \mathbb{R}(\xi_1, \ldots, \xi_n) \) generated by vectors \( f_1, \ldots, f_s \).

As a consequence, a set of generators of \( \text{SYZ}(T) \) corresponds to a set of generators of \( \text{NULL}(T) \). In the case of one variable polynomial matrices, ad hoc algorithms to compute polynomial vectors that generate \( \text{NULL}(T) \) are discussed, for example, in [39], [24]. An implementation thereof is the command \texttt{axb} from Matlab Polynomial Toolbox which we used to perform the computations needed for the following example.

**Example 84**: Consider the electrical circuit of example 1.13 with values \( R_c = 1 \Omega, R_L = 1 \Omega, C = 1 \text{mF}, L = 5 \text{mH} \). For \( M \) as in example 76 and the given numerical values a matrix \( H \in \mathbb{R}^{26 \times 2} [\xi] \) generating \( \text{SYZ}(M^T) \) is

\[
H^T = \begin{bmatrix}
1 - 0.0002 \xi & 0.0007 \xi + 5 \times 10^{-7} \xi^2 & 0 & 1 + 0.0002 \xi & 0 & 1 + 0.0002 \xi & -1 - 0.0002 \xi & \ldots \\
1 + 0.0002 \xi & -1 - 0.0002 \xi & 5 \times 10^{-7} \xi^2 & 1 + 0.0002 \xi & 0 & -1 - 0.0002 \xi & -0.0002 \xi + 5 \times 10^{-7} \xi^2 & \ldots \\
1 + 0.0002 \xi & 5 \times 10^{-7} \xi^2 & 0.0007 \xi + 5 \times 10^{-7} \xi^2 & 1 + 0.0002 \xi & -1 - 0.0002 \xi & \ldots \\
1 + 0.0002 \xi & 5 \times 10^{-7} \xi^2 & -1 - 0.0002 \xi & 1 + 0.0002 \xi & 0.0007 \xi + 5 \times 10^{-7} \xi^2 & 0 & -1 & \ldots \\
1 + 0.0002 \xi & -1 - 0.0002 \xi & 5 \times 10^{-7} \xi^2 & 1 + 0.0002 \xi & 0 & 5 \times 10^{-7} \xi^2 & \end{bmatrix}
\]

and

\[
H^T N = \begin{bmatrix}
-1 - 0.0002 \xi & -5 \times 10^{-7} \xi^2 & 1 + 0.0002 \xi & 5 \times 10^{-7} \xi^2 & 0 & 1 + 0.0002 \xi & 5 \times 10^{-7} \xi^2 & 0 & \ldots \\
0 & -1 & \end{bmatrix}
\]

which is easily seen to be equal to matrix \( R \) from example 2.6 for the given parameter values.
4.3. Algorithms for Elimination

We close this section by showing how the idea of elimination also works on a non linear example.

Example 85: The manifest behavior of the two link robot arm from example 13 can be described by the following equations only involving the manifest variables:

\[ J_{1x} + J_{1y} = 0 \]
\[ V_{1x} - V_{1y} = L_2 \frac{d}{dt} \theta_{11} + R_2 J_{11} + R_2 \frac{K_{2x}}{K_{2y}} J_{11} + \frac{K_{2x}}{K_{2y}} J_{11} + \frac{K_{2x}}{K_{2y}} (V_5 + V_6) + K_1 \frac{d}{dt} \theta_1 - K_2 \frac{d}{dt} \theta_1 \]
\[ J_{2x} + \frac{m \Omega^2 z_1}{2} J_{0} J_{11} + \frac{m \Omega^2 z_2}{2} J_{0} J_{11} + \frac{m \Omega^2 z_3}{2} J_{0} J_{11} + \frac{m \Omega^2 z_4}{2} J_{0} J_{11} = \frac{m \Omega^2 z_1}{2} \sin(\theta_{14}) - \frac{m \Omega^2 z_2}{2} \sin(\theta_{14}) + N_2 J_{11} = \]
\[ T_{11} = H_2 \sin(\theta_{14}) + H_2 F_{11} \cos(\theta_{14}) - \frac{m \Omega^2 z_1}{2} \sin(\theta_{14}) + H_2 F_{11} \cos(\theta_{14}) = \frac{m \Omega^2 z_1}{2} \cos(\theta_{14}) \]
\[ J_{2x} + \frac{m \Omega^2 z_1}{2} J_{0} J_{11} + \frac{m \Omega^2 z_2}{2} J_{0} J_{11} + \frac{m \Omega^2 z_3}{2} J_{0} J_{11} + \frac{m \Omega^2 z_4}{2} J_{0} J_{11} = \frac{m \Omega^2 z_1}{2} \cos(\theta_{14}) \]
\[ J_{2x} = \frac{m \Omega^2 z_1}{2} J_{0} J_{11} + \frac{m \Omega^2 z_2}{2} J_{0} J_{11} + \frac{m \Omega^2 z_3}{2} J_{0} J_{11} + \frac{m \Omega^2 z_4}{2} J_{0} J_{11} = \]
\[ (\sin(\theta_{14}) - H_2 \cos(\theta_{14}))(y_1 + H_2 \sin(\theta_{14})) - y_1^2 + 2 H_1 (x_1 + H_2 \cos(\theta_{14}) - x_1 \cos(\theta_{14})) + 1 y_2 + H_2 \sin(\theta_{14}) - y_1 \sin(\theta_{14}) + H_2^2 = 0 \]

with

\[ T_{0} = K_2 J_{11} - \frac{J_{0} K_{1}}{K_{2}} J_{11} - \frac{J_{0} K_{1}}{K_{2}} J_{11} - \frac{J_{0} K_{1}}{K_{2}} (V_5 + V_6) - J_{0} \frac{d}{dt} \theta_1 \]
\[ F_{0} = F_{11} - m_2 H_1 \left( \begin{array}{c} \frac{d^2}{dt^2} \theta_{14} + H_2 \cos(\theta_{14}) \left( \frac{d^2}{dt^2} \theta_{14} \right)^2 + H_2 \sin(\theta_{14}) \left( \frac{d^2}{dt^2} \theta_{14} \right)^2 \right) - \left( m_{12} + m_{21} \right) \left( \begin{array}{c} H_2 \cos(\theta_{14}) \left( \frac{d^2}{dt^2} \theta_{14} \right)^2 + H_2 \sin(\theta_{14}) \left( \frac{d^2}{dt^2} \theta_{14} \right)^2 \right) \right) - m_2 H_2 g \]
\[ E_1 = m_1 H_1 \left( \begin{array}{c} \frac{d^2}{dt^2} \theta_{14} + H_2 \cos(\theta_{14}) \left( \frac{d^2}{dt^2} \theta_{14} \right)^2 + H_2 \sin(\theta_{14}) \left( \frac{d^2}{dt^2} \theta_{14} \right)^2 \right) \right) - \left( m_{12} + m_{21} \right) \left( \begin{array}{c} H_2 \cos(\theta_{14}) \left( \frac{d^2}{dt^2} \theta_{14} \right)^2 + H_2 \sin(\theta_{14}) \left( \frac{d^2}{dt^2} \theta_{14} \right)^2 \right) \right) - m_2 H_2 g \]
\[ E_2 = \frac{1}{2} \left( \begin{array}{c} y_1 + H_2 \sin(\theta_{14}) - y_1 \sin(\theta_{14}) + H_2 \cos(\theta_{14}) \left( \frac{d^2}{dt^2} \theta_{14} \right)^2 \right) - \left( m_{12} + m_{21} \right) \left( \begin{array}{c} H_2 \cos(\theta_{14}) \left( \frac{d^2}{dt^2} \theta_{14} \right)^2 + H_2 \sin(\theta_{14}) \left( \frac{d^2}{dt^2} \theta_{14} \right)^2 \right) \right) - m_2 H_2 g \]
\[ E_3 = \frac{1}{2} \left( \begin{array}{c} y_1 + H_2 \sin(\theta_{14}) - y_1 \sin(\theta_{14}) + H_2 \cos(\theta_{14}) \left( \frac{d^2}{dt^2} \theta_{14} \right)^2 \right) - \left( m_{12} + m_{21} \right) \left( \begin{array}{c} H_2 \cos(\theta_{14}) \left( \frac{d^2}{dt^2} \theta_{14} \right)^2 + H_2 \sin(\theta_{14}) \left( \frac{d^2}{dt^2} \theta_{14} \right)^2 \right) \right) - m_2 H_2 g \]
The computations for example 85 have been done manually because up to now no general theory of elimination for nonlinear systems has been developed. Results for the case of polynomial nonlinearities have been presented in [23], [21] and related works, where elimination is tackled with tools from differential algebra, in particular autoreduced sets and Ritt’s algorithm. The theory is, however, entirely formal, in the sense that it is uniquely based on equation manipulation and pays no attention to how the solutions (i.e. the system’s trajectories) are affected by these manipulations. Developing a nonlinear theory of elimination based on a strong trajectory characterization as the one provided in the linear case by theorem 78, remains an open problem, and one of crucial importance to the framework for modeling systems which we have been proposing so far.

4.3.2 Ordering Variables

Using Lemma 79 we can provide an alternative algorithm for elimination, whose proof follows easily from the remark that \((h \in SYZ(M^T)) \iff \begin{pmatrix} 0 \\ N^T h \end{pmatrix} \in \langle \begin{pmatrix} M^T \\ N^T \end{pmatrix} \rangle >.

**Algorithm 86:**

\[
R = \text{Elimination order}(N, M);
\]

\[
\hat{R} = \begin{pmatrix} M^T \\ N^T \end{pmatrix};
\]

\[
\text{mo=Product ordering on } M_n^{l+v};
\]

\[
G = \text{Gröbner}(\hat{R});
\]

\[
R = \text{Select columns of } G \text{ with first } l \text{ components equal to zero};
\]

\[
R = R^T;
\]

The difference between matrix \(R\) as output by algorithm 82 and algorithm 86 is that in the latter case the columns of \(R^T\) are a Gröbner basis with respect to the ordering on \(M_n^{v}\) inducing the given product ordering on \(M_n^{l+v}\).
Example 87: Calling Eliminationord\((N, M)\) with \(N\) and \(M\) as in example 75 with POT ordering on \(M_3^e\), for \(v_1 = e_4, v_2 = e_3, v_3 = e_2, v_4 = e_1\) and lexicographic ordering on \(M_3^e\) one obtains output

\[ n = \begin{bmatrix}
\rho_1 \xi_2^2 - (\lambda_1 + 2\rho_1) \xi_1 \xi_2 - \rho_1 \xi_2^2 \\
(\rho_1 + \rho_2) \xi_2^2 - (\lambda_1 + \lambda_2 + 2(\rho_1 + \rho_2)) \xi_1 \xi_2 \\
-\rho_1 \xi_2^2 - (\lambda_1 + 2\rho_1) \xi_1 \xi_2 - \rho_1 \xi_2^2 \\
(\rho_1 + \rho_2) \xi_2^2 - (\lambda_1 + \lambda_2 + 2(\rho_1 + \rho_2)) \xi_1 \xi_2 \\
-1 & 0 & 1 \\
0 & 0 & 0
\end{bmatrix} \]

corresponding to equations 1.9.

4.4 Proofs

Proof of Theorem 78

\(\Rightarrow\) If \(w \in \mathcal{B}\), then by definition of manifest and full behavior, there exists a distribution \(\ell\) such that \(N (\frac{d}{dt}) w - M (\frac{d}{dt}) \ell = 0\). Given any \(h \in \mathbb{R}^p [\xi]\) this of course implies \(h^T (\frac{d}{dt}) (N (\frac{d}{dt}) w - M (\frac{d}{dt}) \ell) = 0\). Therefore \(n \in SYZ(M^T) \Rightarrow n^T (\frac{d}{dt}) N (\frac{d}{dt}) w = 0\).

\(\Leftarrow\) Let \(U = \begin{bmatrix}
U_1 \\
U_2
\end{bmatrix}\) be an unimodular matrix such that \(UM = \begin{bmatrix}
U_1 M \\
U_2 M
\end{bmatrix}\)

\[
\begin{bmatrix}
0 \\
M_2
\end{bmatrix}
\]

with \(M_2 \in \mathbb{R}^{p+1} [\xi]\) a full row rank polynomial matrix; such a \(U\) always exists. We then know that \(\mathcal{B}_{\text{full}} = \ker([N (\frac{d}{dt}) - M (\frac{d}{dt})]) = \ker([U (\frac{d}{dt}) N (\frac{d}{dt}) - U (\frac{d}{dt}) M (\frac{d}{dt})])\). Since \(M_2\) is of full row rank, it can be shown that \(M_2 (\frac{d}{dt})\) is a surjective map from \(\mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^1)\) to \(\mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^p)\), thus the equations \(M_2 (\frac{d}{dt}) \ell = N_2 (\frac{d}{dt}) w\) admit a solution \(\ell\) for any \(w\). The only constraint on \(w\) is hence given by \(N_1 (\frac{d}{dt}) w = 0\), thus showing that \(\mathcal{B} = \ker(G_1 (\frac{d}{dt})).\)

The columns of \(U_1^T\) span \(SYZ(M^T)\). The hypothesis \(n^T (\frac{d}{dt}) N (\frac{d}{dt}) w = 0\) for any \(n \in SYZ(M^T)\) implies that \(U_1 (\frac{d}{dt}) N (\frac{d}{dt}) w = N_1 (\frac{d}{dt}) w = 0\), equivalently \(w \in \mathcal{B}\) as claimed.

Proof of Lemma 79

Of course \(c \subseteq \mathcal{S}\); we are now going to show that \(\forall q \in \mathcal{S}, \exists c \in \mathcal{C}\) such that \(LM(c)\) divides \(LM(q)\), with leading monomials being computed with respect to the given ordering on \(M_3^n\).
Assume $q \in S$. Because \( \begin{pmatrix} 0 \\ q \end{pmatrix} \in < H > \) and $G$ be a Gröbner basis for $< H >$ it follows that there exists $g \in G$ such that $LM(g)$ divides the leading monomial of \( \begin{pmatrix} 0 \\ q \end{pmatrix} \). This means, using notation from example 44, that $LM(g) = \xi^\alpha v_i$, $i \in \{s + 1, \ldots, s + p\}$ and thus $g \in C$.

### 4.5 Conclusions and further research

In this chapter we have presented algorithms to solve the elimination problem, in other words to find equations describing the behavior of given manifest variables starting form equations that also include latent variables. The relevance of this problem in the context of modeling systems has been evidenced by illustrating it by means of a number of examples which had arisen in the first chapter of this work.

As discussed when commenting on example 85, the most pressing future development in this direction is extending the theory and algorithms for elimination to nonlinear systems.