Algorithms in behavioral systems theory

Cotroneo, Tommaso

Chapter 3

Gröbner Bases

As pointed out in the previous chapter, the relationship between modules over $\mathbb{R}[\xi_1, \ldots, \xi_n]$ and linear differential systems allows us to use tools from constructive algebra in order to manipulate and analyze systems. The present chapter is dedicated to introducing the main such tool, namely Gröbner bases.

The literature on the subject is, at this point, vast. We cite [19] as a very accessible introductory text; more advanced topics are covered in [20], [1], [4], the first two being recommended for a discussion of Gröbner bases over modules, the last for theoretical issues regarding orderings. Finally we wish to mention the very recently published monograph [40] providing a fresh treatment of the matter. Implementations of the algorithms we discuss can be found in packages such as Maple [50], Singular [31], CoCoA [11]. For the special case of polynomials in one variable, we sometimes refer to Matlab Polynomial Toolbox [44] for implementation of basic operations. The use of Gröbner bases in the study of nonlinear control systems is discussed in [25]; reference [72], discussing relationship between Gröbner bases and systems of partial differential equations, is very close to the spirit in which we are going to use this tool.

The present chapter contains a very basic introduction to the subject of Gröbner bases, enough to keep this work self-contained. In particular, we first review the concepts of *monomials* and *monomial orderings* in both the cases of scalar and vector polynomials. With these concepts at hand, we proceed to define a *division algorithm* for both cases and, finally, de-

fine Gröbner bases for ideals and modules over $\mathbb{R}[\xi_1, \ldots, \xi_n]$. Given the paramount importance for our purposes of polynomials in one variable, corresponding to systems of ordinary differential equations, we dedicate the last section of the chapter to showing how the theory of Gröbner bases ties up with the classical theory of canonical forms for polynomial matrices over $\mathbb{R}[\xi]$. This way we hope to convince the reader that Gröbner bases provide a unified computational framework for dealing with systems of linear differential equations.

3.1 Ordering monomials

The final goal of this chapter is defining and discussing Gröbner bases for submodules of $\mathbb{R}^m[\xi_1, \ldots, \xi_n]$; in order to do so, two concepts need to be introduced: the ordering of monomials and the division algorithm in $\mathbb{R}^m[\xi_1, \ldots, \xi_n]$. This section is dedicated to the former, while the next deals with the latter; in both cases we try to show how concepts which are common knowledge for polynomials in one indeterminate can be extended to polynomials in more indeterminates and to vectors of such polynomials.

3.1.1 Monomials in $\mathbb{R}[\xi]$

Monomials in $\mathbb{R}[\xi]$ are expressions of the form ξ^m, $m \in \mathbb{N}$; m is also called the *degree* of the monomial. Looking at degrees, the natural *ordering* of $\mathbb{N}$ induces an ordering on monomials by setting

$$\xi^j > \xi^k \Leftrightarrow j > k$$

A polynomial $f \in \mathbb{R}[\xi]$ has the structure $f = \sum_{i=0}^{N} c_i \xi^i$, $0 \neq c_i \in \mathbb{R}$ and can therefore be seen as a linear combination of monomials. The ordering on monomials introduced above allows us then to call ξ^N the *leading monomial* of f (abbreviated ξ^N=LM(f)), N the *degree* of f (abbreviated N=deg(f)), c_N the *leading coefficient* of f (abbreviated c_N=LC(f)) and $c_N \xi^N$ the *leading term* of f (abbreviated $c_N \xi^N$=LT(f)) .

We now give the words we had in italics (monomial, ordering, degree...) appropriate meaning for polynomials in $\mathbb{R}[\xi_1, \ldots, \xi_n]$.

3.1.2 Monomials in $\mathbb{R}[\xi_1, \ldots, \xi_n]$

A *monomial* in $\mathbb{R}[\xi_1, \ldots, \xi_n]$ is an expression of the form $\xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n}$ for $\alpha_i \in \mathbb{N}$; we often use the shorthand multi-index notation ξ^α with $\alpha = (\alpha_1 \cdots \alpha_n) \in \mathbb{N}^n$ to indicate the above monomial. We also denote by $\mathbb{M}_n$ the set of all monomials in $\mathbb{R}[\xi_1, \ldots, \xi_n]$. There is a natural bijection between $\mathbb{M}_n$ and $\mathbb{N}^n$, in the same way as monomials in one variable are associated to natural numbers. Extending the nomenclature used for $\mathbb{R}[\xi]$ we therefore call $\alpha \in \mathbb{N}^n$ the *degree* of ξ^α.

We are now going to show how the set $\mathbb{M}_n$ can be ordered, in other words how, given $\alpha,\ \beta \in \mathbb{N}^n$, we can in some appropriate sense say that $\xi^\alpha > \xi^\beta$. Because of the one to one correspondence between monomials and $\mathbb{N}^n$ the above question is equivalent to the possibility of establishing $\alpha > \beta$ for $\alpha,\ \beta \in \mathbb{N}^n$ and then setting $\xi^\alpha > \xi^\beta \Leftrightarrow \alpha > \beta$. In the following we therefore talk interchangeably of ordering of $\mathbb{N}^n$ or $\mathbb{M}_n$.

The "appropriate sense" we have mentioned above is that of a *monomial ordering* which we know define

Definition 36 : A *monomial ordering* on $\mathbb{N}^n$ is a binary relation $>$ on $\mathbb{N}^n$ satisfying:

i) $>$ is a *total* ordering, equivalently given any two $\alpha,\ \beta\ \in \mathbb{N}^n$ exactly one of three statements

$$\alpha > \beta, \quad \alpha = \beta, \quad \alpha < \beta$$

is true

ii) $\alpha > \beta\ \Rightarrow \alpha + \gamma > \beta + \gamma\ \forall \gamma \in \mathbb{N}^n$

iii) $(0 \cdots 0) < \alpha,\ \forall \alpha \in \mathbb{N}^n,\ \alpha \neq (0 \cdots 0)$

Because $\alpha + \gamma$ is the degree of $x^\alpha x^\gamma$, property *ii)* states that $\xi^\alpha > \xi^\beta\ \Rightarrow \xi^\alpha \xi^\gamma > \xi^\beta \xi^\gamma$; in other words order is preserved when multiplying two monomials by a same monomial.

Property iii), can be shown to be equivalent to requiring that our ordering of $\mathbb{N}^n$ is a *well ordering*. This means that every non empty subset of $\mathbb{N}^n$ has a smallest element under the given ordering or, equivalently, that there

is no infinite strictly decreasing sequence $\alpha_1 > \cdots > \alpha_k > \cdots$. In terms of elements of $\mathbb{M}_n$ it asks for the monomial 1 to be smaller than any other element of $\mathbb{M}_n$.

Having defined what are the properties we require of an ordering of monomials, we also wish to show that such properties can actually be achieved by giving two examples of relations that satisfy definition 36.

Example 37 : The *lexicographic ordering* on $\mathbb{N}^n$ is defined by

$$\alpha >_{lex} \beta \Leftrightarrow (\alpha - \beta) \text{ has leftmost non zero entry } > 0$$

For example, we have $(1,2) >_{lex} (0,3)$, $(2,0) >_{lex} (1,2)$ and $(1,2) >_{lex} (1,1)$.

Applying this order to $\mathbb{M}_n$ means that ξ^α is greater than ξ^β if for the smallest i such that $\xi_i^{\alpha_i} \neq \xi_i^{\beta_i}$ we have $\alpha_i > \beta_i$. For example, in the case of monomials in two variables, we have

$$1 < \xi_2 < \cdots < \xi_2^k < \cdots < \xi_1 < \xi_1\xi_2 < \xi_1\xi_2^k < \cdots < \xi_1^2 < \cdots$$

Example 38 : The *graded lexicographic ordering* on $\mathbb{N}^n$ is defined by

$$\alpha >_{glex} \beta \Leftrightarrow \sum_{i=1}^{n} \alpha_i > \sum_{i=1}^{n} \beta_i \text{ or } \sum_{i=1}^{n} \alpha_i = \sum_{i=1}^{n} \beta_i \text{ and } \alpha >_{lex} \beta$$

In this case we have $(1,2) >_{glex} (0,3)$, $(1,2) >_{glex} (2,0)$ and $(1,2) >_{glex} (1,1)$

Applying this order to monomials means we first compare the total degree and then break tie using lexicographic ordering as defined above. Again for the case of $\mathbb{M}_2$ we have

$$1 < \xi_2 < \xi_1 < \xi_2^2 < \xi_1\xi_2 < \xi_1^2 < \cdots$$

Trivial as this may be, notice that if we consider the special case of $\mathbb{N}$, the natural ordering $0 < 1 < 2 < \cdots$ is the only monomial ordering which can be established.

Any polynomial $f \in \mathbb{R}[\xi_1, \ldots, \xi_n]$ can be written as a linear combination of elements of $\mathbb{M}_n$. In particular, given any monomial ordering $>$ on $\mathbb{N}^n$,

$f = \sum_{k=1}^{N} c_k \xi^{\alpha_k}$ for some finite N, $0 \neq c_k \in \mathbb{R}$, $\xi^{\alpha_k} \in \mathbb{M}_n$ and $\alpha_i < \alpha_{i+1}$, $i = 1, \ldots, N-1$.

We then call ξ^{α_N} the *leading monomial* of f with respect to the given ordering; the notation ξ^{α_N}=LM(f) is also used, the dependence on the specific term ordering being silently implied. In the same way, α_N is called the *degree* of f (abbreviated α_N=deg(f)), c_N the *leading coefficient* of f (abbreviated c_N=LC(f)) and $c_N \xi^{\alpha_N}$ the *leading term* of f (abbreviated $c_N \xi^{\alpha_N}$=LT(f)). If $c_N = 1$, f is called a *monic* polynomial.

Notice how the concepts of degree and leading monomial are now well defined also for polynomials in more than one variable. Contrary to what happens in $\mathbb{R}[\xi]$, however, they are not unique, as a consequence of the fact that only one monomial order exists for $\mathbb{N}$ while more exist for the general case $\mathbb{N}^m$.

Example 39 : Consider $f = 2\xi_1^2 + 3\xi_1\xi_2^2$. If we use lexicographic ordering on $\mathbb{N}^2$ we have $\deg(f) = (2,0)$, $\mathrm{LM}(f) = \xi_1^2$ and $\mathrm{LC}(f) = 2$; if we use graded lexicographic ordering, instead, we have $\deg(f) = (1,2)$, $\mathrm{LM}(f) = \xi_1\xi_2^2$ and $\mathrm{LC}(f) = 3$.

3.1.3 Monomials in $\mathbb{R}^m[\xi_1, \ldots, \xi_n]$

Let $v_1, \ldots, v_m \in \mathbb{R}^m$ be a basis for $\mathbb{R}^m$; a *monomial* in $\mathbb{R}^m[\xi_1, \ldots, \xi_n]$ is an expression of the form $\xi^\alpha v_i$, with $\xi^\alpha \in \mathbb{M}_n$. The set of all monomials in $\mathbb{R}^m[\xi_1, \ldots, \xi_n]$ is denoted by $\mathbb{M}_n^m$; elements of $\mathbb{M}_n^m$ are thus $m-$ dimensional polynomial vectors obtained taking the product of an element of $\mathbb{M}_n$ by one of the given basis vectors.

Example 40 : Consider $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ as a basis for $\mathbb{R}^2$; then $\begin{pmatrix} \xi \\ \xi \end{pmatrix}$ and $\begin{pmatrix} 3\xi^2 \\ -3\xi^2 \end{pmatrix}$ are both monomials in $R^2[\xi]$, while $\begin{pmatrix} \xi \\ 0 \end{pmatrix}$ is not.

Often the basis for $\mathbb{R}^m$ is chosen to be $v_i = e_i$, $i = 1, \ldots, m$ with e_i the $i-$th unit vector in $\mathbb{R}^m$. If this is the case, $\mathbb{M}_n^m$ becomes the set of $m-$ dimensional

polynomial vectors having only one non zero component, that component being an element of $\mathbb{M}_n$.

If we define the index set $I = \{1, \ldots, m\}$ we can establish a bijection between $\mathbb{M}_n^m$ and $\mathbb{N}^n \times I$, the element $\xi^\alpha v_i \in \mathbb{M}_n^m$ being mapped to $(\alpha, i) \in \mathbb{N}^n \times I$. Extending the nomenclature used for $\mathbb{M}_n$ we call $(\alpha, i) \in \mathbb{N}^n \times I$ the *degree* of $\xi^\alpha v_i$.

As for $\mathbb{M}_n$, we can now show how elements of $\mathbb{M}_n^m$ can be ordered in some appropriate sense. Such a question is equivalent to the possibility of ordering $\mathbb{N}^n \times I$ and then setting $\xi^\alpha v_i > \xi^\beta v_j \Leftrightarrow (\alpha, i) > (\beta, j)$. Whenever convenient, we thus talk interchangeably of ordering of $\mathbb{M}_n^m$ or of $\mathbb{N}^n \times I$; most of the times it is, in fact, easier to define properties referring to $\mathbb{N}^n \times I$ but the relevant applications involve $\mathbb{M}_n^m$.

We first define the properties of the ordering we are after.

Definition 41 : Let $I = \{1, \ldots, m\}$; a *monomial ordering* on $\mathbb{N}^n \times I$ is a relation $>$ on $\mathbb{N}^n \times I$ satisfying:

i) $>$ is a *total* ordering, equivalently given any two (α, i) (β, j) $\in \mathbb{N}^n \times I$ exactly one of three statements

$$(\alpha, i) > (\beta, j), \quad (\alpha, i) = (\beta, j), \quad (\alpha, i) < (\beta, j)$$

is true

ii) $\forall i, j \in \mathbb{I},\ (\alpha, i) > (\beta, j) \ \Rightarrow (\alpha + \gamma, i) > (\beta + \gamma, j)\ \forall \gamma \in \mathbb{N}^n$

iii) $(0, i) < (\alpha, i),\ \forall i \in I,\ \forall \alpha \in \mathbb{N}^n,\ \alpha \neq (0 \cdots 0)$

Notice how the conditions required by the above definition are analogous to those required in definition 36 and actually correspond in the special case $I = 1$.

In particular, because $(\alpha + \gamma, i)$ is the degree of the monomial $x^\alpha x^\gamma v_i$, property ii) can be interpreted by saying that $\xi^\alpha v_i > \xi^\beta v_j \ \Rightarrow \xi^\alpha \xi^\gamma v_i > \xi^\beta + \xi^\gamma v_j$, in other word multiplying two monomials in $\mathbb{M}_n^m$ by the same monomial in $\mathbb{M}_n$ does not alter their ordering.

Property iii), instead, can be shown to be equivalent to requiring that our ordering of $\mathbb{N}^n \times I$ is a *well ordering* in the same sense described when commenting definition 36.

One can of course think of the set $\mathbb{N}^n \times I$ as a bunch of replicas of $\mathbb{N}^n$ and regard an element $(\alpha, i) \in \mathbb{N}^n \times I$ as a term $\alpha \in \mathbb{N}^n$ appearing in position i. It should not come as a surprise that, as the following examples show, "natural" ways of ordering $\mathbb{N}^n \times I$, exploit orders of $\mathbb{N}^n$ as starting point.

Example 42 : Assume an order of $\mathbb{N}^n$ is given. The corresponding *term over position* ordering of $\mathbb{N}^n \times I$ is defined by

$$(\alpha, i) >_{TOP} (\beta, j) \Leftrightarrow \alpha > \beta \text{ or } \alpha = \beta \text{ and } i < j$$

with $\alpha > \beta$ taken with respect to the given ordering of $\mathbb{N}^n$. In other words we first compare the terms α and β using an order for $\mathbb{N}^n$ and, in case equal, we try and break the tie by looking at the "position" in which they appear.

For example, in the case of $\mathbb{M}_2^2$ with basis vectors $v_i = e_i$ and lexicographic ordering for $\mathbb{M}_2$ we have $\begin{pmatrix} \xi_1\xi_2 \\ 0 \end{pmatrix} >_{TOP} \begin{pmatrix} \xi_1 \\ 0 \end{pmatrix} >_{TOP} \begin{pmatrix} 0 \\ \xi_1 \end{pmatrix} >_{TOP} \begin{pmatrix} \xi_2 \\ 0 \end{pmatrix}$

Example 43 : Assume an ordering of $\mathbb{N}^n$ is given. The corresponding *position over term* ordering of $\mathbb{N}^n \times I$ is defined by

$$(\alpha, i) >_{POT} (\beta, j) \Leftrightarrow i < j \text{ or } i = j \text{ and } \alpha > \beta$$

with $\alpha > \beta$ taken with respect to the given ordering of $\mathbb{N}^n$. This time, therefore, we first look at the position in which the terms α and β appear, and, in case of equality, try breaking tie comparing the terms using an order of $\mathbb{N}^n$. Again for $\mathbb{M}_2^2$ basis vectors $v_i = e_i$ and lexicographic ordering for $\mathbb{M}_2$ we have $\begin{pmatrix} \xi_1\xi_2 \\ 0 \end{pmatrix} >_{POT} \begin{pmatrix} \xi_1 \\ 0 \end{pmatrix} >_{POT} \begin{pmatrix} \xi_2 \\ 0 \end{pmatrix} >_{POT} \begin{pmatrix} 0 \\ \xi_1 \end{pmatrix}$

Example 44 : It turns out that the position over term orderings defined in example 43 are a special case of a wider class of orderings, namely *product orderings* which we now introduce.

Let $m = m_1 + m_2$, $I_1 = \{1, \ldots, m_1\}$, $I_2 = \{1, \ldots, m_2\}$ $I = \{1, \ldots, m\}$. Assume that monomial orderings on $\mathbb{N}^n \times I_1$ and $\mathbb{N}^n \times I_2$ are given. We then

define the *product ordering* on $\mathbb{N}^n \times I$ induced by the given orderings as

$$(\alpha, i) >_{PROD} (\beta, j) \Leftrightarrow \left[\begin{array}{c} \text{a) } i \in \{1, \ldots, m_1\} \text{ and } j \in \{m_1+1, \ldots, m\} \\ \text{or} \\ \text{b) } i, j \in \{1, \ldots, m_1\} \text{ and } (\alpha, i) > (\beta, j) \text{ in } \mathbb{N}^n \times I_1 \\ \text{or} \\ \text{c) } i, j \in \{m_1+1, \ldots, m\} \text{ and } (\alpha, i-m_1) > (\beta, j-m_1) \\ \text{in } \mathbb{N}^n \times I_2 \end{array}\right.$$

In terms of $\mathbb{M}_n^m$ what we are doing can be interpreted in the following way. We fix a basis $a_1, \ldots, a_{m_1}$ for $\mathbb{R}^{m_1}$ and a monomial ordering on $\mathbb{M}_n^{m_1}$; similarly we choose a basis $b_1, \ldots, b_{m_2}$ for $\mathbb{R}^{m_2}$ and a monomial ordering on $\mathbb{M}_n^{m_2}$. As a basis for $\mathbb{R}^m$ we then choose $v_1, \ldots, v_m$ such that $v_i = \begin{pmatrix} a_i \\ 0 \end{pmatrix}$, $i = 1, \ldots m_1$ and $v_i = \begin{pmatrix} 0 \\ b_{i-m_1} \end{pmatrix}$, $i = m_1+1, \ldots, m$. All monomials of the form $\xi^\alpha v_i$ $i = 1, \ldots, m_1$ are then greater than all monomials $\xi^\beta v_j$ $j = m_1+1, \ldots, m$. If $i, j \in \{1, \ldots, m_1\}$, instead $\xi^\alpha v_i > \xi^\beta v_j \Leftrightarrow \xi^\alpha a_i > \xi^\beta a_j$ according to the given ordering of $\mathbb{M}_n^{m_1}$. Similarly if $i, j \in \{m_1+1, \ldots, m_2\}$ $\xi^\alpha v_i > \xi^\beta v_j \Leftrightarrow \xi^\alpha b_{i-m_1} > \xi^\beta b_{j-m_1}$ according to the given ordering of $\mathbb{M}_n^{m_2}$. In the following we often talk of the product ordering on $\mathbb{M}_n^m = \mathbb{M}_{n_1+n_2}^m$ induced by the given orderings on $\mathbb{M}_{m_1}^n$ and $\mathbb{M}_{m_2}^n$.

The above construction can, of course, be extended to the case in which more indexing sets $I_k = \{1, \ldots, m_i\}$, $k = 1, \ldots, t$ are considered with $\sum_{i=1}^k m_k = m$. The POT ordering ordering from example 43 is then obtained as a special case of this construction.

Any vector polynomial $f \in \mathbb{R}^m[\xi_1, \ldots, \xi_n]$ can be written as a linear combination of elements of $\mathbb{M}_n^m$. In particular, given any monomial order $>$ on $\mathbb{M}_n^m$, $f = \sum_{k=1}^N c_k f_k$ for some finite N, $c_k \in \mathbb{R}$, $f_k \in \mathbb{M}_n^m$ and $f_i < f_{i+1}$, $i = 1, \ldots, N-1$. We then call f_N the *leading monomial* of f with respect to the given ordering (abbreviated f_N=LM(f). If $f_N = \xi^{\alpha_N} v_{i_N}$ we call $(\alpha_N, i_N) \in \mathbb{N}^N \times I$ the *degree* of f (abbreviated α_N=deg(f)), c_N the *leading coefficient* of f (abbreviated c_N=LC(f)) and $c_N f_N$ the *leading term* of f (abbreviated $c_N f_N$=LT(f)). If $c_N = 1$, f is called a *monic* polynomial.

3.2 Division algorithm

With ordering of monomials at hand, we are now ready to introduce a division algorithm for polynomials in more that one variable. Using the same strategy as for orderings in the previous section, we first briefly recap the situation for $\mathbb{R}[\xi]$ and then build extensions to $\mathbb{R}[\xi_1, \ldots, \xi_n]$ and $\mathbb{R}^m[\xi_1, \ldots, \xi_n]$.

3.2.1 Division in $\mathbb{R}[\xi]$

The following theorem holds for $\mathbb{R}[\xi]$

Theorem 45 : *Given any two polynomials* $f,\ t \in \mathbb{R}[\xi]$ *there exist uniquely defined* $q,\ r \in \mathbb{R}[\xi]$ *such that*

$$f = qt + r$$

and either $r = 0$ *or* $deg(r) < deg(q)$. *The polynomial* q *is called the* quotient *of the division of* f *by* t *and* r *the* remainder.

The existence part of the proof of the above theorem is given by showing an algorithm, commonly known as "long division", that actually performs the required computation. We briefly recall it using MATLAB pseudocode notation:

Algorithm 46 :

$[q, r]$=Division(f, t);

$q = 0$;
$r = f$;

while $((r \neq 0)$ and deg$(q) \leq$deg$(r))$
    $q = q + \frac{\mathrm{LT}(r)}{\mathrm{LT}(t)}$;
    $r = r - \frac{\mathrm{LT}(r)}{\mathrm{LT}(t)} t$;
endwhile

3.2.2 Division in $\mathbb{R}[\xi_1, \ldots, \xi_n]$

If one stops a while to think about the above theorem and algorithm, one realizes that all we needed in order to define a quotient and remainder of the division of two polynomials were the concept of degree of a polynomial, and the fact that degrees can be ordered. A great part of last section, however, was devoted exactly to extending this concept to the case of polynomials in n variables; it should not seem unreasonable, therefore, to also try and extend the idea of division to polynomials in $\mathbb{R}[\xi_1, \ldots, \xi_n]$.

When looking at monomials in n variables the concept of divisibility is obviously defined by saying that ξ^α *divides* ξ^β if there exists a third monomial ξ^γ such that $\xi^\beta = \xi^\alpha \xi^\gamma$. Given a monomial order on $\mathbb{N}^n$, a polynomial $r \in \mathbb{R}[\xi_1, \ldots, \xi_n]$ is then called *reduced* with respect to a set of polynomials $t_1, \ldots, t_s$ if no monomial in r is divisible by $\mathrm{LM}(t_i)$, $i = 1, \ldots, s$. This definition of reducedeness depends heavily on the chosen term order. Take for example $r = \xi_1 + \xi_2^4$, $t = \xi_1\xi_2 + \xi_2^3$; then r is reduced with respect to t if we use lexicographic order, but not reduced if we use graded lexicographic order. Also notice that, in case we are looking at $\mathbb{R}[\xi]$, r is reduced with respect to g if and only if $\deg(r) < \deg(g)$.

Using the above concept of reducedeness we can now write the following theorem for $\mathbb{R}[\xi_1, \ldots, \xi_n]$

Theorem 47 : *Given polynomials* $f,\ t_1, \ldots, t_s \in \mathbb{R}[\xi_1, \ldots, \xi_n]$ *and a monomial order on* $\mathbb{N}^n$, *there exist effectively constructible* $r,\ q_1, \ldots, q_s \in \mathbb{R}[\xi_1, \ldots, \xi_n]$ *such that*

$$f = q_1 t_1 + \cdots q_s t_s + r$$

and

i) $LM(f) = \max\{\max_i\{LM(q_i)LM(t_i)\}, LM(r)\}$

ii) $r = 0$ *or it is reduced with respect to* $t_1, \ldots, t_s$.

r *is called a* remainder *of the division of* f *by* $t_1, \ldots, t_s$.

Again the proof of the above theorem is given by simply providing an algorithm that performs the required computation. We now proceed to show it:

Algorithm 48 :

$[q_1, \ldots, q_s, r]$=Division$(f, t_1, \ldots, t_s)$;

for $i = 1 : s$ $q_i = 0$ endfor;
$h = f$;
$r = 0$;

while $(h \neq 0)$

if ($\exists i$ such that LM(t_i) divides LM(h)) then

j =min$(i$ such that LM(t_i) divides LM$(h))$;
$q_j = q_j + \frac{\mathrm{LT}(h)}{\mathrm{LT}(t_j)}$;
$h = h - \frac{\mathrm{LT}(h)}{\mathrm{LT}(t_j)} t_j$;

else

$r = r$+LT(h);
$h = h$−LT(h);

endif

endwhile

We leave it to the reader to verify that algorithm 48 actually produces an output satisfying the conditions of theorem 47; proofs can be found in the references indicated at the beginning of this chapter.

A few comments are, however, in order. In the first place it is not difficult to see that algorithm 48 coincides with 46 in case it receives as input $f,\ t \in \mathbb{R}[\xi]$. In the second place, one should remark that, contrary to theorem 45, theorem 47 only contains an existence statement about quotients and remainders, and no uniqueness. This is because quotients q_i and remainders r are indeed not unique in the case of n variable polynomials; they depend on the chosen monomial order (as might be expected) but also, for a given monomial order, on the order in which the polynomials $t_1, \ldots, t_s$ are taken. We now give two examples to illustrate these crucial issues.

Example 49 : Consider $f = \xi_1,\ t = \xi_1 + \xi_2^2$. If lexicographic order is chosen, then $[1, -\xi_2^2]$ =Division(f, t). On the other hand, if degree lexico-

graphic order is used $[0, \xi_1]$ =`Division`(f, t)

Example 50 : Consider $f = \xi_1\xi_2^2 - \xi_1$, $t_1 = \xi_1\xi_2 + 1$, $t_2 = \xi_2^2 - 1$ and assume lexicographic order has been chosen. Then $[\xi_2, 0, -\xi_1 - \xi_2]$ =`Division`(f, t_1, t_2), but $[\xi_1, 0, 0]$ =`Division`(f, t_2, t_1).

3.2.3 Division in $\mathbb{R}^m[\xi_1, \ldots, \xi_n]$

Looking at the division algorithm in $\mathbb{R}[\xi_1, \ldots, \xi_n]$ we realize it is entirely based on the possibility of ordering and dividing monomials. Because the issue of ordering $\mathbb{M}_n^m$ has already been cleared in the previous section, we may hope to be able to define a division for polynomials in $\mathbb{R}^m[\xi_1, \ldots, \xi_n]$ provided we can give a meaning to division of monomials in $\mathbb{M}_n^m$. This can be actually done by defining $\xi^\alpha v_i$ to *divide* $\xi^\beta v_j$ if and only if $i = j$ and there exists $\xi^\gamma \in \mathbb{N}^n$ such that $\xi^\beta = \xi^\alpha\xi^\gamma$; we then set $\xi^\beta v_i / \xi^\alpha v_j = \xi^\gamma$. Notice therefore how divisibility of monomials in $\mathbb{R}^m[\xi_1, \ldots, \xi_n]$ entails two conditions: the basis vectors have to be the same ($v_i = v_j$) and the monomials have to divide in the sense of $\mathbb{R}[\xi_1, \ldots, \xi_n]$. Also notice how, in case possible, the result of dividing two vector monomials is a monomial $\xi^\gamma \in \mathbb{R}[\xi_1, \ldots, \xi_n]$. For example $\begin{pmatrix} \xi \\ 0 \end{pmatrix}$ divides $\begin{pmatrix} \xi^2 \\ 0 \end{pmatrix}$ and the result of their division is ξ, but $\begin{pmatrix} \xi \\ 0 \end{pmatrix}$ does not divide $\begin{pmatrix} 0 \\ \xi^2 \end{pmatrix}$.

With division of monomials defined, the concept of reducedeness follows from what has been said for the case of $\mathbb{R}[\xi_1, \ldots, \xi_n]$. Given a term order on $\mathbb{N}^n \times I$, in fact, we call a polynomial $r \in \mathbb{R}^m[\xi_1, \ldots, \xi_n]$ *reduced* with respect to a set of polynomials $t_1, \ldots, t_s$ if no monomial in r is divisible by $\mathrm{LM}(t_i)$, $i = 1, \ldots, s$. Notice how this definition reduces to the one given for $\mathbb{R}[\xi_1, \ldots, \xi_n]$ in the special case $I = 1$; all considerations on reducedeness done in that situation extend to the vector case we are now dealing with.

We can then write

Theorem 51 : *Given vector polynomials $f,\ t_1, \ldots, t_s \in \mathbb{R}^m[\xi_1, \ldots, \xi_n]$ and a monomial order on $\mathbb{N}^n \times I$, there exist $r \in \mathbb{R}^m[\xi_1, \ldots, \xi_n]$ and $q_1, \ldots, q_s \in \mathbb{R}[\xi_1, \ldots, \xi_n]$ such that*

$$f = q_1 t_1 + \cdots + q_s t_s + r$$

and

i) $LM(f) = \max\{\max_i\{LM(q_i)LM(t_i)\},\ LM(r)\}$

ii) $r = 0$ *or it is reduced with respect to* $t_1, \ldots, t_s$.

r *is called a* remainder *of the division of* f *by* $t_1, \ldots, t_s$.

Notice that the statement of theorem 51 is the same as that of theorem 47, once reducedeness is suitably interpreted. Also as for theorem 47, the proof is given by actually establishing an algorithm $[q_1, \ldots q_s, r]=$ `Division`$(f, t_1, \ldots, t_s)$ that performs the required computations. We will not do so because the algorithm reads the same as 48 once division of monomials is interpreted in the vector sense just defined. We often use the matrix notation $[Q, r]=$`Division`$[f, T]$ with $Q = [q_1, \ldots, q_s] \in \mathbb{R}^s[\xi_1, \ldots, \xi_n]$ and $T = [t_1, \ldots, t_s] \in \mathbb{R}^{p\times s}[\xi_1, \ldots, \xi_n]$, so that $f = TQ + r$, r being a remainder of the division of f by $t_1, \ldots, t_s$.

3.3 Gröbner bases

We now come to the main issue of this chapter, namely establishing a theory of Gröbner bases for submodules of $\mathbb{R}^m[\xi_1, \ldots, \xi_n]$. The reason for this interest lies in the relationship, established in the previous chapter, between systems described by linear partial differential equations and submodules of $\mathbb{R}^m[\xi_1, \ldots, \xi_n]$.

Although, as just said, our main interest is in modules, we choose to follow the same strategy already used for orderings and division algorithm, namely start with a detailed account of the scalar case (i.e. ideals) and then discuss the extension to vectors (i.e. modules).

3.3.1 Gröbner bases for ideals

The last example from section 3.2.2 allows us to make a couple of remarks which are instrumental in understanding the motivation behind the definition of Gröbner basis which we are going to give. In the first place notice

that $f = \xi_1\xi_2^2 - \xi_1 \in < t_1, t_2 >$; we would therefore expect that dividing f by t_1, t_2 would result in $r = 0$ irrespective of the order in which the division is carried out; this is not the case as shown by performing $\texttt{Division}(f, t_1, t_2)$, which outputs $r = -\xi_1 - \xi_2$. In this respect the situation in $\mathbb{R}[\xi_1, \ldots, \xi_n]$ is crucially different from the one in $\mathbb{R}[\xi]$. In that case, in fact, if $f \in < t >$ (and because $\mathbb{R}[\xi]$ is a PID there is no loss of generality in only considering principal ideals) then $\texttt{Division}(f, t)$ will always result in a remainder $r = 0$. In order to obtain analogous properties when using division for polynomials in n variables we will need to resort to Gröbner bases.

Also notice that the reason why we get two different values for r in the above example is that there exists a polynomial $g = -\xi_1 - \xi_2 \in < t_1, t_2 >$ such that, in the given ordering, $\mathrm{LM}(g)$ is not divisible by either $\mathrm{LM}(t_1)$ or $\mathrm{LM}(t_2)$. As we shall see, it is exactly by ensuring this property that Gröbner Bases are defined. We have in fact:

Definition 52 : Given a monomial order and an ideal $\mathfrak{J} \subseteq \mathbb{R}[\xi_1, \ldots, \xi_n]$, a set $\mathcal{G} = \{g_1, \ldots, g_t\} \subseteq \mathfrak{J}$ is called a *Gröbner basis* for $\mathfrak{J}$ with respect to the given order if $\forall\ f \in \mathfrak{J},\ \exists i$ such that $\mathrm{LM}(g_i)$ divides $\mathrm{LM}(f)$.

It can be shown that for any given monomial order and for any given ideal $\mathfrak{J}$ a Gröbner basis for $\mathfrak{J}$ with respect to the given order actually exists. It can also be shown that if a set $\mathcal{G} = \{g_1, \ldots, g_t\} \subseteq \mathfrak{J}$ is a Gröbner basis for $\mathfrak{J}$ with respect to any monomial order, then $\mathfrak{J} =< g_1, \ldots, g_t >$. In other words, surprising as this may seem, by just requiring a property on leading monomials, we also ensure that the given set is a generating set for the whole ideal.

Because of the paramount importance they have in all of this work, we now stop to describe some properties of Gröbner bases.

Given an ideal $\mathfrak{J} \subseteq \mathbb{R}[\xi_1, \ldots, \xi_n]$ and a monomial order, one can build the set $\mathrm{LM}(\mathfrak{J})$ of all leading monomials of elements in the ideal $\mathfrak{J}$ and then look at the ideal $< \mathrm{LM}(\mathfrak{J}) >$ it generates. Because $\mathbb{R}[\xi_1, \ldots, \xi_n]$ is Noetherian, we know that, without loss of generality, we can consider $\mathfrak{J} =< t_1, \ldots, t_s >$ for some $t_i \in \mathbb{R}[\xi_1, \ldots, \xi_n]$. It is then trivial to see that $<\mathrm{LM}(t_1), \ldots, \mathrm{LM}(t_s) > \subseteq < \mathrm{LM}(\mathfrak{J}) >$, and one might expect that in general $< \mathrm{LM}(\mathfrak{J}) > = <\mathrm{LM}(t_1), \ldots, \mathrm{LM}(t_s) >$. Example 50 shows, however, that this is not true (in that case, in fact, for $\mathfrak{J} =< t_1, t_2 >$ we have $\xi_1 \in <\mathrm{LM}(\mathfrak{J}) >$ and $\xi_1 \notin < \mathrm{LM}(t_1), \mathrm{LM}(t_2) >$). This property, however, becomes true in

case the generating set for $\mathfrak{J}$ we are considering happens to be a Gröbner basis for $\mathfrak{J}$ with respect to the given order. We have in fact

Proposition 53 : *Given a monomial order, a set* $\mathcal{G} = \{g_1, \ldots, g_t\} \subseteq \mathfrak{J}$ *is a* Gröbner basis *for the ideal* $\mathfrak{J}$ *with respect to the given order if and only if* $< LM(\mathfrak{J}) >=< LM(g_1), \ldots, LM(g_t) >$

It can also be shown that using Gröbner basis in the division algorithm makes the remainder independent of the order in which the divisors are taken. In fact

Proposition 54 : *Given a monomial order, a set* $\mathcal{G} = \{g_1, \ldots, g_t\} \subseteq \mathfrak{J}$ *is a* Gröbner basis *for* $\mathfrak{J}$ *with respect to the given order if and only if for any* $f \in \mathbb{R}[\xi_1, \ldots, \xi_n]$ *the remainder* r *of division of* f *by* $g_1, \ldots, g_t$ *using* `Division` *is independent of the order in which* $g_1, \ldots, g_t$ *are taken.*

As a consequence of the above proposition we obtain that when dividing by a Gröbner basis for $\mathfrak{J}$, then $f \in \mathfrak{J} \Leftrightarrow$ `Division`$[f, g_1, \ldots, g_t]$ outputs $r = 0$.

We now want to discuss a final property of Gröbner bases which also allows us to sketch an algorithm to construct them. Before doing so, we define the concepts of least common multiple of two monomials, and of $S-$polynomial of two polynomials.

Given two monomials $\xi^\alpha,\ \xi^\beta \in \mathbb{R}[\xi_1, \ldots, \xi_n]$, let $\alpha = (\alpha_1, \ldots, \alpha_n)$, $\beta = (\beta_1, \ldots, \beta_n)$ and $\gamma = (\gamma_1, \ldots, \gamma_n) \in \mathbb{N}^n$ with $\gamma_i = \max(\alpha_i, \beta_i)$. Then ξ^γ is called the *least common multiple* of $\xi^\alpha,\ \xi^\beta$ (abbreviated $\xi^\gamma = \text{LCM}\ (\xi^\alpha,\ \xi^\beta)$). It is obvious that ξ^γ is a multiple of both ξ^α and ξ^β; using property ii) from definition 36 it is not difficult to see that ξ^γ is also the least such multiple under any term order; the terminology least common multiple is, therefore, jusitified.

Given two polynomials $f,\ g \in \mathbb{R}[\xi_1, \ldots, \xi_n]$ the *S-polynomial of* f *and* g is defined as

$$S(f,g) = \frac{\text{LCM}(\text{LM}(f), \text{LM}(g))}{\text{LT}(f)} f - \frac{\text{LCM}(\text{LM}(f), \text{LM}(g))}{\text{LT}(g)} g$$

and is, therefore, a combination of f and g in which the leading monomials of both polynomials cancel. Take, for example, $f = 2\xi_1^2 + 2\xi_1\xi_2$

and $g = \xi_1\xi_2^2 + \xi_2$; using lexicographic order $\mathrm{LM}(f) = \xi_1^2$, $\mathrm{LM}(g) = \xi_1\xi_2^2$, $\mathrm{LCM}(\mathrm{LM}(f), \mathrm{LM}(g)) = \xi_1^2\xi_2^2$ and $S(f,g) = \xi_1\xi_2^3 - \xi_1\xi_2$.

Assume we are given $\mathfrak{J} =< t_1, \ldots, t_s >$. From what has been said in this section we know that $< t_1, \ldots, t_s >$ will not be a Gröbner basis for $\mathfrak{J}$ if we can find a combination of t_i whose leading monomial is not in the ideal generated by the $\mathrm{LM}(t_i)$'s. Because, as just seen, $S(t_i, t_j)$ is a polynomial whose leading monomial is different from those of both t_i and t_j, it is not unreasonable to expect a relationship to exist between $S-$polynomials and Gröbner bases. This relationship is, in fact, very deep, as stated in the following:

Proposition 55 : *Given a monomial order, a set* $\mathcal{G} = \{g_1, \ldots, g_t\} \subseteq \mathfrak{J}$ *is a* Gröbner basis *for* $\mathfrak{J}$ *with respect to the given order if and only if* $\forall i \neq j$ `Division`$(S(g_i, g_j), g_1, \ldots, g_t)$ *yields remainder* $r = 0$.

The above property is of crucial importance because it allows to construct an algorithm that builds a Gröbner basis $G = \{g_1, \ldots, g_t\}$ for an ideal $\mathfrak{J}$, given any monomial ordering and an arbitrary set of generators $T = \{t_1, \ldots, t_s\}$ of $\mathfrak{J}$. This algorithm is the famous Buchberger algorithm, nowadays implemented in most computer algebra packages (e.g. [31], [11], [50]). For the sake of completeness we now sketch how this algorithm works in principle. Standard implemented versions are more refined and improved than the version we present, but still exploit the same basic idea.

Algorithm 56 :

$[g_1, \ldots, g_t]$=`Buchberger`$(t_1, \ldots, t_s)$;

$t = s$;
`for` $i = 1 : t$ $g_i = t_i$ `endfor`;
$C = \{(g_i, g_j) \mid i \neq j\}$;
`repeat`

`Choose` $(g_k, g_l) \in C$; $C = C - \{(g_k, g_l)\}$;
$[q_1, \ldots q_s, r] =$ `Division`$(S(g_k, g_l), g_1, \ldots, g_s)$;

`if` $r \neq 0$ `then`

`for` $k = 1 : t$ $C = C \cup \{(r, g_k)\}$ `endfor`;

$t = t + 1;\ g_t = r;$
endif;

until $(C == 0)$

The algorithm starts by setting the Gröbner basis equal to the set of generators for the ideal received as input ($g_i = t_i,\ i = 1, \ldots, t$) and by building all possible pairs thereof ($C = \{(g_i, g_j) \mid i \neq j\}$). At each step of the main loop, a pair of polynomials from the set C is chosen, the corresponding $S-$polynomial is built and divided by the polynomials that make up the Gröbner basis. If the remainder r is not 0, then it is added to the set of polynomials that constitute the basis ($t = t+1;\ g_t = r;$) and all the resulting new pairs of elements are taken into account ($C = C \cup \{(r, g_k)\},\ k = 1, \ldots, t\};$). We start with a set of generators for the ideal $\mathfrak{J}$, and at each step add to this set a polynomial $r \in \mathfrak{J}$. Thus the elements g_i always are a generating set for $\mathfrak{J}$. Because at each step we compute the remainder upon division of an $S-$polynomial, it is clear that if the algorithm terminates, then the condition from proposition 55 is satisfied and the resulting set $g_1, \ldots, g_t$ is a Gröbner basis for $\mathfrak{J}$. The fact that algorithm 56 actually terminates is guaranteed by the fact that, if $r \neq 0$, then $< \mathrm{LM}(g_1), \ldots, \mathrm{LM}(g_t) > \subset$ $< \mathrm{LM}(g_1), \ldots, \mathrm{LM}(g_t), \mathrm{LM}(r) >$ with the inclusion being strict. This implies that after a finite number of steps all the computed remainders must be $r = 0$, otherwise an infinite chain of ideals would be generated, which is not possible because $\mathbb{R}[\xi_1, \ldots, \xi_n]$ is a Noetherian ring.

It is clear that the Gröbner basis for an ideal $\mathfrak{J}$ with respect to a given order is by no means unique (to see this, just add to any basis $\mathcal{G}$ a linear combination of elements of $\mathcal{G}$). There is, however, one such basis which is unique and which we now define

Definition 57 : Given a term order, a set $\mathcal{G} = \{g_1, \ldots, g_t\} \subseteq \mathfrak{J}$ is a *reduced Gröbner basis* for the ideal $\mathfrak{J}$ with respect to the given order if it is a Gröbner basis and moreover

i) g_i is monic, $i = 1, \ldots, t$

ii) g_i is reduced with respect to g_k, $k \neq i$.

For the proof of uniqueness of the reduced Gröbner basis we refer to [19]. We do show, however, how such a basis can be built starting from an arbitrary basis $\mathcal{G} = \{h_1, \ldots, h_g\}$.

Algorithm 58 :

$[g_1, \ldots, g_t]$ =Redgro$(h_1, \ldots, h_g)$;

$t = 0$;
for $j = 1 : g$

$[q_1, \ldots, q_s, r] =$ Division $(h_j, g_1, \ldots g_t, h_{j+1}, \ldots, h_t)$;

if $r \neq 0$ then $t = t + 1$; $h_j = r$; endif

endfor;

for $i = 1 : t$ $g_i = \frac{1}{\mathrm{LC}(g_i)} g_i$ endfor

Combining algorithms 56 and 58 we then easily obtain a procedure $[g_1, \ldots, g_t]$=Gröbner$(t_1, \ldots, t_s)$ that computes a reduced Gröbner basis for $\mathfrak{J} =< t_1, \ldots, t_s >$

Algorithm 59 :

$[g_1, \ldots, g_t]$ =Gröbner$(t_1, \ldots, t_s)$;

$[h_1, \ldots, h_g]$=Buchberger$(t_1, \ldots, t_s)$;
$[g_1, \ldots, g_t]$ =Redgro$(h_1, \ldots, h_g)$;

Notice that most standard computer algebra packages actually compute the reduced Gröbner basis for a given ideal.

3.3.2 Gröbner bases for modules

If one reads carefully through the previous section, one realizes that all we needed in order to define Gröbner bases for ideals were the concept of ordering for monomials, therefore of leading monomial of a polynomial, and the availability of a division algorithm. Because both have also been suitably defined for vectors, it is reasonable to set

Definition 60 : Given a submodule $\mathfrak{M} \subseteq \mathbb{R}^m[\xi_1, \ldots, \xi_n]$ and a monomial order on $\mathbb{N}^n \times I$, a set $\mathcal{G} = \{g_1, \ldots, g_t\} \subseteq \mathfrak{M}$ is called a *Gröbner basis* for $\mathfrak{M}$ with respect to the given order if $\forall\ f \in \mathfrak{M}$, $\exists j$ such that $\mathrm{LM}(g_j)$ divides $LM(f)$.

As for ideals, it can be shown that for any given term order and any given submodule $\mathfrak{M}$, a Gröbner basis for $\mathfrak{M}$ with respect to the given order actually exists and that if $\mathcal{G} = \{g_1, \ldots, g_t\} \subseteq \mathfrak{M}$ is a Gröbner basis for $\mathfrak{M}$ with respect to any order, then $\mathfrak{M} =< g_1, \ldots, g_t >$.

All characterizations of Gröbner bases given by propositions 53, 54, 55 for ideals extend to the case of modules. For the sake of completeness we now explicitly restate them.

Given a submodule $\mathfrak{M} \subseteq \mathbb{R}^m[\xi_1, \ldots, \xi_n]$ and a term order on $\mathbb{N}^n \times I$, we build the set $\mathrm{LM}(\mathfrak{M})$ of all leading monomials of elements in $\mathfrak{M}$ and then look at the submodule $< \mathrm{LM}(\mathfrak{M}) >$ it generates. We then have

Proposition 61 : *Given a monomial order on $\mathbb{N}^n \times I$, a set $\mathcal{G} = \{g_1, \ldots g_t\}$ $\subseteq \mathfrak{M}$ is a* Gröbner basis *for $\mathfrak{M}$ with respect to the given order if and only if $< LM(\mathfrak{M}) >=< LM(g_1), \ldots, LM(g_t) >$*

We also have

Proposition 62 : *Given a monomial order on $\mathbb{N}^n \times I$, a set $G = \{g_1, \ldots, g_t\}$ $\subseteq \mathfrak{M}$ is a* Gröbner basis *for $\mathfrak{M}$ with respect to the given order if and only if for any $f \in \mathbb{R}^m[\xi_1, \ldots, \xi_n]$ the remainder r of division of f by $g_1, \ldots, g_t$ using* `Divide` *is independent of the order in which $g_1, \ldots, g_t$ are taken.*

Most important of all, the characterization of Gröbner bases in terms of $S-$polynomials also holds in this case.

Given two monomials $\xi^\alpha v_i,\ \xi^\beta v_j \in \mathbb{M}_n^m$, we define their *least common multiple* as

$$\mathrm{LCM}(\xi^\alpha v_i,\ \xi^\beta v_j) = \begin{cases} 0 \text{ if } i \neq j \\ \mathrm{LCM}(\xi^\alpha,\ \xi^\beta) v_i \text{ if } i = j \end{cases}$$

with $\mathrm{LCM}(\xi^\alpha,\ \xi^\beta)$ defined for monomials in $\mathbb{R}[\xi_1, \ldots, \xi_n]$ as in the previous section. Notice that in the special case $\xi^\alpha v_i,\ \xi^\beta v_j \in \mathbb{M}^m$ (i.e. $\alpha,\ \beta \in \mathbb{N}$) we

have

$$\mathrm{LCM}(\xi^\alpha v_i,\ \xi^\beta v_j) = \begin{cases} 0 \text{ if } i \neq j \\ \xi^\alpha v_i \text{ if } i = j \end{cases}$$

where, without loss of generality, we have taken $\alpha > \beta$.

Given two vector polynomials $f,\ g \in \mathbb{R}^m[\xi_1, \ldots, \xi_n]$ the *S-polynomial* of f and g is then defined as

$$S(f,g) = \frac{\mathrm{LCM}(\mathrm{LM}(f), \mathrm{LM}(g))}{\mathrm{LT}(f)} f - \frac{\mathrm{LCM}(\mathrm{LM}(f), \mathrm{LM}(g))}{\mathrm{LT}(g)} g \tag{3.1}$$

and is also in the vector case a combination of f and g in which the leading monomials of both polynomials cancel. In the special case $f,\ g \in \mathbb{R}^m[\xi]$ assume, again without loss of generality, $\mathrm{LT}(f) = \xi^\alpha v_i$, $\mathrm{LT}(g) = \xi^\beta v_j$ with $\alpha > \beta$. From what we have just discussed about least common multiples in $\mathbb{M}^m$, it follows that $S(f,g)$ is either zero or has the special structure

$$S(f,g) = \frac{1}{\mathrm{LC}(f)} f - \frac{\xi^{\alpha-\beta}}{\mathrm{LC}(g)} g \tag{3.2}$$

This remark turns out to be crucial in next section when analyzing special properties of Gröbner basis for submodules of $\mathbb{R}^m[\xi]$.

We can now show the characterization of Gröbner basis in terms of $S-$polynomials that we had announced

Proposition 63 : *Given a term order on $\mathbb{N}^n \times I$, a set $G = \{g_1, \ldots, g_t\} \subseteq \mathfrak{M}$ is a* Gröbner basis *for the submodule $\mathfrak{M}$ with respect to the given order if and only if $\forall i \neq j$ `Division`$(S(g_i, g_j), g_1, \ldots, g_t)$ yields remainder $r = 0$.*

The importance of the above theorem is, of course, that it enables one to construct the Buchberger algorithm also in the vector case. We will not explicitly state such an algorithm because it is literally the same as the one given for ideals. Also the definition and construction of a reduced Gröbner basis for a module follow from what we described in the case of ideals, and are not repeated for the modules. In the following, however, we assume available a procedure $[g_1, \ldots, g_t]$ =`Gröbner`$(t_1, \ldots, t_s)$ that builds the reduced Gröbner basis $\{g_1, \ldots, g_t\}$ for a submodule $\mathfrak{M}$ given an ordering and arbitrary set of generators $\{t_1, \ldots, t_s\}$ of $\mathfrak{M}$. We also often use the matrix notation G=`Gröbner`(T) to indicate that the columns of G are a reduced Gröbner basis for the module generated by the columns of T. Notice

that, as a consequence of $< G >=< T >$, there exists a polynomial matrix H such that $G = TH$. H can be very easily computed by keeping track of the reduction steps in Buchberger's algorithm; we chose not to show this feature for sake of brevity and refer, for example, to [40] for a detailed explanation. In the following we do, however, sometimes use the notation (G, H)=`Gröbner`(T) to indicate availability of a procedure that computes both the reduced Gröbner basis G and the transformation matrix H.

No discussion of algorithms is complete without at least touching the issue of complexity. The literature on the subject is rich, we refer to [49], [38], [45], [47], alongside the general references given in the introduction. Here we simply report the essential property that the complexity of computing the Gröbner basis of $< t_1, \ldots, t_s > \mathbb{R}^m[\xi, \ldots, \xi_n]$ is, in general, $(smd)^{O(2^n)}$ with d the maximal total degree of the elements t_i. Better bounds can be obtained in special cases (e.g. polynomials in 2 variables, zero dimensional ideals, etc.), but we do not enter into the details and, again, refer the interested reader to the literature.

The complexity of Buchberger's algorithm is itself, in general, $(smd)^{O(2^n)}$ and is, thus, polynomial in the number(s), dimension(m), and degree (d) of the given generators, but exponential in the number of variables of the polynomial ring we work in. This feature is, of course, very unpleasant in applications where n is actually a variable of the problem (see [40], [1] for use of Gröbner basis in integer programming or statistics). In our case, however, n is a given number, corresponding to the number of independent variables describing our dynamical systems (typically $n \leq 4$ corresponding to space and time). For n fixed, we then have an algorithm that is of polynomial complexity.

3.4 Gröbner basis for modules over $\mathbb{R}[\xi]$

As discussed in chapter 2, polynomials in one variable (and matrices thereof) play a special role in our presentation, being associated to systems of ordinary differential equations. We thus find it fitting to spend a section to discuss how the general theory of Gröbner bases specializes to the case of modules over $\mathbb{R}[\xi]$. In particular, we first establish a connection between Gröbner bases and well known canonical forms for polynomial matrices in one variable. After that, we proceed to show how algorithm 56 can be implemented in the special case of modules over $\mathbb{R}[\xi]$.

We start by recalling that, given a polynomial matrix $T \in \mathbb{R}^{m\times q}[\xi]$, its highest column coefficient matrix T^{hc} is defined as the real matrix whose i-th column equals the coefficient vector of the highest power of ξ that occurs in the i-th column of T.

Example 64 : If $T = \begin{bmatrix} 3\xi^2+\xi+1 & 2 \\ 2\xi & \xi+1 \end{bmatrix}$ then $T^{hc} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$

A polynomial matrix T is defined to be *column proper* (or column reduced) if the columns of T_{hc} are linearly independent. If T is not column proper, a *column proper form* of T is defined to be any column proper polynomial matrix T_{cp} such that the modules $\langle T\rangle$ and $\langle T_{cp}\rangle$ are the same, and such that T_{cp} is column proper. It is easily seen that any matrix T actually admits infinitely many column proper forms T_{cp}. There is one however which is unique and which we now define (see [39])

Definition 65 : A polynomial matrix $T \in \mathbb{R}^{p\times q}[\xi]$ is said to be in *column echelon* form if:

i) It is column proper, and has column degrees

$$\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_q$$

ii) For each column $t_i,\ i = 1,\ldots,q$ there exists a row index j_i such that

   a) $t_{j_i i}$ is monic and of degree α_i
   b) $\deg(t_{ki}) < \alpha_i$ if $k < j_i$
   c) $\deg(t_{j_i k}) < \alpha_i$ if $k \neq i$
   d) If $\alpha_i = \alpha_k$ and $i < j$ then $j_i < j_k$

Given an arbitrary polynomial matrix T, its *column echelon form* is defined as the unique column echelon polynomial matrix T_{ce} such that $\langle T\rangle = \langle T_{ce}\rangle$.

Classical algorithms for finding one column proper form T_{cp} and the column echelon form T_{ce} starting from an arbitrary polynomial matrix T are described in the literature (e.g. in [39], [24]). Possible implementations thereof are the functions `pcolred` and `echelon` from Matlab Polynomial Toolbox; the latter, however, only works for non-singular T.

The following proposition establishes a relationship between Gröbner bases and column proper matrices

Proposition 66 : *Let $T \in \mathbb{R}^{m\times q}[\xi]$ be a column proper matrix with leading coefficient matrix $T^{hc} = [t_1^{hc} \ \cdots t_q^{hc}] \in \mathbb{R}^{m\times q}$. Choose a set $v_i,\ i = 1,\ldots,m$ of generators for $\mathbb{R}^m$ such that $v_i = t_i^{hc},\ i = 1,\ldots,q$. Then T is a Gröbner basis for $< T >$ with respect to the TOP ordering on $\mathbb{M}^m$.*

The next proposition, instead, relates column echelon forms and Gröbner bases

Proposition 67 : *Let $T \in \mathbb{R}^{m\times q}[\xi]$ be a column echelon matrix. Then T is the reduced Gröbner basis for $< T >$ with respect to the TOP ordering on $\mathbb{M}^m$ if the standard basis $v_i = e_i$ is chosen for $\mathbb{R}^m$.*

Given proposition 67, one might expect also the reduced Gröbner basis with respect to the POT ordering to have a significant interpretation in terms of canonical forms for one variable polynomial matrices. In order to investigate this aspect we first recall the definition of *row Hermite* form of a polynomial matrix

Definition 68 : A polynomial matrix $T \in \mathbb{R}^{m\times q}[\xi]$ is said to be in *row Hermite form* if for each column $t_i,\ i = 1,\ldots,q$ there exists row indices $j_1 < j_2 < \cdots < j_q$ such that

i) $t_{ki} = 0$ for $k < j_i$ (lower staircase structure)

ii) $t_{j_i i}$ is monic and of degree α_i

iii) If $\alpha_i > 0$ then $\deg(t_{j_i k}) < \alpha_i$ for $k < i$

iv) If $\alpha_i = 0$ then $(t_{j_i k}) = 0$ for $k < i$

Given an arbitrary polynomial matrix T, its *row Hermite form* is defined as the unique row Hermite polynomial matrix T_{rh} such that $\langle T\rangle = \langle T_{rh}\rangle$.

Algorithms for finding the row Hermite form T_{rh} starting from an arbitrary polynomial matrix T are described in the literature (e.g. in [39], [24], [34]). A possible implementation thereof is the function `hermite` from Matlab Polynomial Toolbox.

The following proposition should now not be a surprise (see also [51])

Proposition 69 : *Let $T \in \mathbb{R}^{m\times q}[\xi]$ be a matrix in row Hermite form. Then T is the the reduced Gröbner base for $< T >$ with respect to the POT ordering on $\mathbb{M}^m$ if the standard basis $v_i = e_i$ is chosen for $\mathbb{R}^m$.*

We can now show how the algorithm to compute a reduced Gröbner basis simplifies in case we are dealing with polynomials in one variable; the key property behind such simplification is, as we shall see, the fact that now S-polynomials have the special structure $S(g_i, g_j) = g_i - \xi^\gamma g_j$ for some $\gamma \in \mathbb{N}$.

Algorithm 70 :

$[g_1, \ldots, g_t]$ =`Gröbner`$(t_1, \ldots, t_s)$;

$t = s$;
`for` $i = 1 : t$ $g_i = t_i$ `endfor;`
`nored=true;`

`repeat`
  `nored=true;`
  `while` $i \leq t$
    $[q_1, \ldots, q_{t-1}, r]$ =`Division`$(g_i, g_1, g_2, \ldots, g_{i-1}, g_{i+1}, \ldots g_t)$;
    `if` $(\exists q_j \neq 0)$
      `nored=false;`
      `if` $(r \neq 0)$
        $g_i = r$;
      `else`
        `for` $k = i + 1 : t$ $g_{k-1} = g_k$ `endfor`

$t = t - 1$;

endif

endif

$i = i + 1$

endwhile

until (nored==true)

for $i = 1 : t$ $g_i = \frac{1}{\mathrm{LC}_{g_i}} g_i$ **endfor**

The following comments, hopefully, clarify how algorithm 70 works

i) If $[q_1, \ldots q_{t-1}, r]$ =**Division** $(g_i, g_1, g_2, \ldots, g_{i-1}, g_{i+1}, \ldots, g_t)$ outputs at least one quotient $q_j \neq 0$ it means that there exist at least one g_j, $j \neq i$ such that $LM(g_i) = \xi^\gamma \mathrm{LM}(g_j)$ for $\gamma \in \mathbb{N}$. If this is the case then the remainder r is the same as would be obtained by performing $[q_1, \ldots, q_t, r]$ =**Division**$(S(g_i, g_j), g_1, g_2, \ldots, g_{i-1}, g_i, g_{i+1}, \ldots, g_t)$.
Computing $[q_1, \ldots q_{t-1}, r]$ =**Division** $(g_i, g_1, g_2, \ldots, g_{i-1}, g_{i+1}, \ldots, g_t)$ is thus equivalent to performing a basic step in Buchberger's algorithm, namely computing the $S-$polynomial of two elements and taking its remainder after division.

ii) Because $g_i - r \in < g_1, g_2, \ldots, g_{i-1}, g_{i+1}, \ldots g_t >$ it follows that we can simply replace g_i by r and still have a generating set for the given module, in other words $< g_1, \ldots, g_i, \ldots, g_t > = < g_1, \ldots, r, \ldots g_t >$
(**if** $(r \neq 0)$ $g_i = r$).

Replacing g_i with r is, of course, not necessary if $r = 0$, in which case we just continue with the remaining elements **(for** $k = i + 1 : t$ $g_{k-1} = g_k$**)**.

Moreover $< \mathrm{LT}(g_1), \ldots, \mathrm{LT}(g_i), \ldots, \mathrm{LT}(g_t) > \subset < \mathrm{LT}(g_1), \ldots, \mathrm{LT}(r), \ldots$ $\mathrm{LT}(g_t) >$ with the inclusion being strict if $r \neq 0$. Using the ascending chain condition this guarantees that, after a finite number of steps, $[q_1, \ldots, q_{t-1}, r] =$
Division $(g_i, g_1, g_2, \ldots, g_{i-1}, g_{i+1}, \ldots, g_t)$ always outputs $r = g_i$, thus the main **while** loop thus terminates with **(nored==true)** guaranteeing the end of the algorithm.

The fact that $[q_1, \ldots, q_{t-1}, r]$ =**Divide** $(g_i, g_1, g_2, \ldots, g_{i-1}, g_{i+1}, \ldots, g_t)$ always outputs $r = g_i$ is equivalent to all the $S-$polynomials being

zero; proposition 63 is thus verified, showing that the output is a Gröbner basis for the given module. The final `for` cycle is simply meant to make all the elements monic; It is then easily verified that the resulting set is, in fact, the reduced Gröbner basis.

Notice that, as a consequence of all $S-$ polynomials being zero as discussed in remark *ii*) it follows that the reduced Gröbner basis is composed of independent elements. This might have been expected, given that all $\mathbb{R}[\xi]-$modules are free; it is, however, not true in the general case of polynomials in $n > 1$ variables, as the following example shows

Example 71 : Let $t_1 = \begin{pmatrix} \xi_2 \\ 1 \end{pmatrix}$, $t_2 = \begin{pmatrix} \xi_1 \\ 0 \end{pmatrix}$; then $< t_1, t_2 >$ is obviously a free module. Its reduced Gröbner basis with respect to POT lexicographic ordering, however, is $g_1 = t_1$, $g_2 = t_2$, $g_3 = \begin{pmatrix} 0 \\ \xi_1 \end{pmatrix}$ which is not an independent set.

Finally let us remark that as consequence of what we said at the end of the previous section, algorithm 70 is of polynomial complexity in m, s and d with d the highest degree of the given generators $t_1, \ldots, t_s$. A complete comparison with classical algorithms (see e.g. [39], [24], [34]) to bring matrices in row Hermite or row echelon form still needs to be performed. It is clear, however, that algorithm 70 being parameterized by the possible orderings on $\mathbb{R}^m[\xi]$ has an advantage of flexibility and generality with respect to algorithms that only compute a particular transformation of a given input matrix.

3.5 Proofs

Proof of Proposition 66
With TOP ordering and the chosen basis for $\mathbb{R}^m$ we have $\mathrm{LM}(t_i) = t_i^{hc}\xi_i^{\alpha}$. Moreover $t \in < T >$ if and only if $t = \sum_{i=1}^{q} p_i t_i,\ p_i \in \mathbb{R}[\xi]$. Let $\beta_i \in \mathbb{N}$ be the degree of p_i, then because the t_i^{hc} are linearly independent it follows that $\mathrm{LM}(t) = \xi^{\alpha_k+\beta_k} = t_k^{hc}$ with $k = \min\{j \mid \alpha_j + \beta_j \text{ is maximum}\}$ and is thus a multiple of $\mathrm{LM}(t_i)$. The claim follows using definition 60.

Proof of Proposition 67
From ii.a and ii.b of definition 65 follows that $\mathrm{LM}(t_i) = \xi_i^{\alpha} e_{j_i}$. The same kind of argument used in the proof of proposition 66 shows that $t \in < T >$ implies that $\mathrm{LM}(t) = \xi^{\alpha_k+\beta_k} e_{j_k}$ for some $k \in \{1, \ldots, q\}$ and some $\beta_k \in \mathbb{N}$. The columns of T are, thus, a Gröbner basis for $< T >$. From ii.a of definition 65 it follows that the vectors t_i are monic, while ii.c assures that t_i is reduced with respect to t_k, $k \neq i$ and thus that the basis is reduced.
Proof of Proposition 69
From i and ii of definition 68 follows that, with respect to POT ordering, $\mathrm{LM}(t_i) = \xi_i^{\alpha} e_{j_i}$. From i and the fact that the row indices are strictly decreasing follows that $t \in < T >$ implies that $\mathrm{LM}(t) = \xi^{\alpha_k+\beta_k} e_{j_k}$ for some $k \in \{1, \ldots, q\}$ and some $\beta_k \in \mathbb{N}$. The columns of T are, thus, a Gröbner basis for $< T >$. From ii follows that the t_i are monic vectors and from i, iii and iv that t_i is reduced with respect to t_k, $k \neq i$ and thus that the basis is reduced.

3.6 Conclusion and further research

The main aim of this chapter has been to show how Gröbner bases provide a unified framework for computations involving submodules over the polynomial ring $\mathbb{R}[\xi_1, \ldots, \xi_n]$. Together with the connections between modules and behaviors established in the previous chapter, this opens the way for the application of Gröbner bases techniques to the analysis of linear systems which are presented in the following chapters.

In order to define Gröbner bases, total orderings are needed. For the applications which we have in mind, this is not always a good feature. It amounts, in fact, to establishing an order on system variables which we would, instead, often like to regard on equal footing; in order to do so, partial orderings must be used instead of total ones. It would therefore be important to try and develop a theory similar to that of Gröbner basis presented in this chapter, this time relying on partial rather than total ordering of monomials. In particular, the theory of H-basis ([48]) might provide some insight in this direction.

As already pointed out in sections 3.3.2 and 3.4, we only briefly touch the issue of computational complexity and efficiency of Gröbner basis. Given the special role played in systems theory by ordinary differential equations and associated polynomials in one variable, we think it would be especially

interesting to understand how the algorithm presented in section 3.4 compares to those existing in the literature for computing canonical forms of one variable polynomial matrices.