



ARISTOTLE UNIVERSITY OF THESSALONIKI  
DEPARTMENT OF MATHEMATICS  
THEORETICAL INFORMATICS AND SYSTEMS & CONTROL THEORY

“Notions of equivalences of multivariate polynomial matrices.”

**M.Sc.THESIS**

**EVAGELIA LAMPIRI**

**Supervisor:** Nicholas Karampetakis

Associate Professor, Aristotle University Of Thessaloniki

Thessaloniki, June 2014





ARISTOTLE UNIVERSITY OF THESSALONIKI  
DEPARTMENT OF MATHEMATICS  
THEORETICAL INFORMATICS AND SYSTEMS & CONTROL THEORY

“Notions of equivalences of multivariate polynomial matrices.”

**M.Sc.THESIS**

**EVAGELIA LAMPIRI**

**Supervisor:** Nicholas Karampetakis

Associate Professor, Aristotle University Of Thessaloniki

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή

.....  
N. Καραμπετάκης  
Αν. Καθηγητής Α.Π.Θ.

.....  
Ε. Αντωνίου  
Επικ. Καθηγητής Μηχ.  
Πληροφορικής  
Τμήμα Α.Τ.Ε.Ι. Θεσ.

.....  
Γ. Ραχώνης  
Αν. Καθηγητής Α.Π.Θ.

Thessaloniki, June 2014

.....

Λαμπίρη Ευαγγελία

Πτυχιούχος Μαθηματικός Α.Π.Θ.

Copyright © Λαμπίρη Ευαγγελία, 2014.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευτεί ότι εκφράζουν τις επίσημες θέσεις του Α.Π.Θ.

## ΠΕΡΙΛΗΨΗ

Στην παρούσα διπλωματική εργασία γίνεται μελέτη των βάσεων Gröbner και του αλγόριθμου εύρεσης τους (Buchberger αλγόριθμος). Επίσης πραγματεύεται σχέσεις ισοδυναμίας και συγκεκριμένα ισοδυναμίες 1-D και 2-D πολυωνυμικών πινάκων. Η εργασία είναι χωρισμένη σε δύο μέρη.

Το πρώτο μέρος της εργασίας, ασχολείται με τις βάσεις Gröbner, καθώς και τον αλγόριθμο του Buchberger για την εύρεση αυτών. Επίσης, δίνεται ο τρόπος εύρεσης τους μέσα από το πρόγραμμα Mathematica.

Το δεύτερο μέρος της εργασίας, περιλαμβάνει έννοιες των 1-D και 2-D πολυωνυμικών πινάκων και τα σημεία στα οποία αυτές διαφέρουν. Επίσης, γίνεται εκτενής ανάλυση της αντιστρέψιμης ισοδυναμίας (unimodular equivalence) που αφορά πολυωνυμικούς πίνακες ιδίων διαστάσεων και της γενικευμένης αντιστρέψιμης ισοδυναμίας (extended unimodular equivalence) που αφορά πολυωνυμικούς πίνακες διαφορετικών διαστάσεων. Τέλος, δίνονται οι zero coprime και factor coprime ισοδυναμίες και τα αναλλοίωτα στοιχεία αυτών.

## ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ

Μονωνυμική διάταξη, αλγόριθμος διαίρεσης, βάσεις Gröbner, Buchberger αλγόριθμος, ισοδυναμίες, Smith μορφή, πολυωνυμικοί πίνακες, αντιστρέψιμη ισοδυναμία, γενικευμένη αντιστρέψιμη ισοδυναμία, zero coprime ισοδυναμία, factor coprime ισοδυναμία

## ABSTRACT

In the present Thesis is studied the Gröbner bases and their algorithm (Buchberger algorithm). Moreover, it concerns about equivalence relations and specifically equivalences of 1-D and 2-D polynomial matrices. This paper consists of two parts.

The first part of this thesis deals with Gröbner bases, as well as Buchberger algorithm for finding these. Furthermore, it is given the way to find them through the program of Mathematica.

The second part of this thesis includes notions of 1-D and 2-D polynomial matrices and the points on which they differ. There is also an extensive analysis of the unimodular equivalence which is about polynomial matrices of the same dimension and of the extended unimodular equivalence that is about polynomial matrices of different dimensions. Lastly, given the zero coprime equivalence and factor coprime equivalence and the invariants of them.

## ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ

Monomial ordering, division algorithm, Gröbner bases, Buchberger algorithm, equivalences, Smith form, polynomial matrices, unimodular equivalence, extended unimodular equivalence, zero coprime equivalence, factor coprime equivalence

## ΠΡΟΛΟΓΟΣ

Θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες στον επιβλέποντα καθηγητή μου κ.Νικόλαο Καραμπετάκη για τη βοήθεια του, την επιστημονική του καθοδήγηση και την συνολική επίβλεψη της εργασίας αυτής. Όπως επίσης και για την ευκαιρία που μου έδωσε να συνεχίσω τις σπουδές μου.

Οφείλω ακόμη να ευχαριστήσω τα υπόλοιπα μέλη της τριμελούς επιτροπής κ. Αντωνίου Ευστάθιο και τον κ. Ραχώνη Γεώργιο για το χρόνο που αφιέρωσαν στην μελέτη και αξιολόγηση της εργασίας.

Τέλος θα ήθελα να ευχαριστήσω την οικογένεια μου για τη ψυχολογική και οικονομική υποστήριξη καθ' όλη την διάρκεια των σπουδών μου. Ιδιαίτερα θα ήθελα να ευχαριστήσω τον σύζυγο μου, ο οποίος συνεχίζει να με στηρίζει στην ακαδημαϊκή μου πορεία.

## Table of Contents

Glossary of Notation .....	9
Chapter 1 .....	10
Gröbner bases.....	10
1.1 Primary notions and definitions .....	11
1.2 Monomial Ordering .....	16
1.3 Division algorithm.....	25
1.4 Gröbner bases .....	30
1.5 Buchberger’s algorithm .....	34
1.6 Gröbner bases in Mathematica.....	39
Chapter 2 .....	42
Equivalence of Polynomial Matrices .....	42
2.1 1-D Polynomial Matrices.....	43
2.1.1 Important Notions and Definitions .....	43
2.1.2 Coprimeness of 1-D polynomial matrices .....	50
2.2 Equivalences in 1-D polynomial matrices .....	56
2.2.1 Equivalence relation.....	56
2.2.2 Unimodular Equivalence .....	57
2.2.3 Equivalence of Binomials .....	61
2.2.4 Extended Unimodular Equivalence .....	63
2.3 2-D Polynomial Matrices.....	68
2.3.1 Notions of Coprimeness .....	70
2.3.2 Invariant Polynomials and Zeros.....	79
2.4 Equivalences in 2-D polynomial matrices .....	81
References .....	96



## Glossary of Notation

$F$	$\mathbb{R}$ or $\mathbb{C}$ ;
$\mathbb{R}$	the field of real numbers;
$R[x]$	the ring of polynomials in the single indeterminate $x$ with coefficients in the field $F$ ;
$R[x_1, x_2, \dots, x_n]$	the ring of polynomials in the $n$ indeterminates $x_1, x_2, \dots, x_n$ with coefficients in the field $F$ ;
$I$	an ideal;
$\mathbb{N}$	the natural numbers;
$A_{j_1, j_2, \dots, j_k}^{i_1, i_2, \dots, i_k}$	the $k^{th}$ order minor of the matrix $A$ using rows $i_1, i_2, \dots, i_k$ and columns $j_1, j_2, \dots, j_k$ ;
$ A , \det(A)$	the determinant of the matrix $A$ ;
$adj(A)$	the adjoint of the matrix $A$ ;
$rank(A)$	the rank of the matrix $A$ ;
$A^{-1}$	the inverse of the matrix $A$ ;
$A^+$	the generalized inverse of the matrix $A$ ;
$A^T$	the transpose of the matrix $A$ ;
$I_p$	the identity matrix with dimension $p \times p$ ;

# Chapter 1

---

## Gröbner bases

## 1.1 Primary notions and definitions

---

Initially, we will mention some primary notions and definitions, which will help us to understand better the aftermath. The ring is one such a notion, which is an algebraic structure with two binary operations.

The theory of rings sprang through the study of two specific rings classes, the ring of polynomials in  $n$  variables over the real or complex numbers and the “integers” of a body of algebraic numbers. First it was *David Hilbert* (1862-1943) who introduced the term *ring*, but it needed to get to the second decade of 20<sup>th</sup> century to see a completely abstract definition. The theory of computing rings was founded by *Emmy Noether* (1882-1935) in her monumental work “*Theory of Ideals in Rings*”, which appeared in 1921.

### Definition 1.1 [1]

A *group*, denoted  $\langle G, * \rangle$ , is a set  $G$ , together with a binary operation  $*$  in  $G$ , such that the following axioms are satisfied:

- G1. The binary operation  $*$  is associative.
- G2. There exists an element  $e$  in  $G$ , such that  $e * x = x * e = x$  for every  $x \in G$ .  
(This element  $e$  is called identity element for  $*$  in  $G$ )
- G3. For every  $a$  in  $G$ , there exists an element  $a'$  in  $G$  with the property  $a' * a = a * a' = e$ . (The element  $a'$  is called inverse of  $a$  as to the operation  $*$ )

*Note:* A group  $G$  is called abelian (in honour of Niels Abel) if the binary of the operation  $*$  is commutative.

### Definition 1.2 [1]

A *ring*, denoted  $\langle G, +, * \rangle$ , is a set with two binary operations  $+$  and  $*$ , referred to as addition and multiplication, which satisfies the following axioms:

- R1.  $\langle G, + \rangle$  is an abelian group.
- R2. Multiplication is associative.
- R3. For all  $a, b, c \in R[x]$  the left distributivity law is valid,  $a(b + c) = ab + ac$ , and the right distributivity law,  $(a + b)c = ac + bc$ .

If, in addition, multiplication is commutative too, i.e.  $a \cdot b = b \cdot a$  for all

$a, b \in \langle G, +, * \rangle$ , then  $\langle G, +, * \rangle$  is called a commutative ring.  $\langle G, +, * \rangle$  is called a ring with 1, or ring with unity if it contains a distinguished element 1 with  $1 \neq 0$  and  $1 \cdot a = a$  for all  $a \in \langle G, +, * \rangle$ .

**Definition 1.3** [9]

The set of polynomials  $a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ ,  $a_i \in \mathbb{R}$ ,  $n$  finite, is denoted  $R[x]$ . On  $R[x]$  we can define with the natural way the addition and multiplication. Then it can be easily proved that  $R[x]$  is a ring.

Since  $R[x]$  is a ring, it makes sense to consider the set  $(R[x])[y]$ , i.e., polynomials in  $y$  with coefficients in  $R[x]$ . We see that

$$\begin{aligned} (R[x])[y] &= \{f_0 + f_1y + f_2y^2 + \dots + f_ny^n, f_i \in R[x]\} \\ &= \{r_0 + r_1x + r_2y + r_3x^2 + r_4xy + r_5y^2 + \dots\} \\ &= \{g_0 + g_1x + g_2x^2 + \dots + g_mx^m | g_i \in R[y]\} \\ &= (R[y])[x]. \end{aligned}$$

So  $(R[x])[y] = (R[y])[x]$ , and can therefore be denoted by  $R[x, y]$  without any risk for confusion. Likewise, we can define the polynomial ring over  $\mathbb{R}$  in  $n$  variables  $x_1, x_2, \dots, x_n$ . This polynomial ring is denoted by  $R[x_1, x_2, \dots, x_n]$ .

**Example 1.1**

We know that the axioms  $R1 - R3$  of the ring are valid in every subset of complex numbers which is group with the addition and closed as to the multiplication. For example,  $\langle \mathbb{Z}, +, \cdot \rangle$ ,  $\langle \mathbb{Q}, +, \cdot \rangle$ ,  $\langle \mathbb{R}, +, \cdot \rangle$  and  $\langle \mathbb{C}, +, \cdot \rangle$  are all rings.

In ring theory, a special subset of a ring is an ideal. Ideals generalize certain subsets of the integers. *Ernest Eduard Kummer* (1810-1893) was the person who introduced the notion of “*ideal of complex number*” in 1847, in his effort to retain the notion of unambiguously analysis in some rings of algebraic integers.

**Definition 1.4** [4]

Let  $R[x]$  and  $S[x]$  be rings and  $f: R[x] \rightarrow S[x]$  a mapping. Then  $f$  is called a *homomorphism* of rings if the following hold:

- (i).  $f(a + b) = f(a) + f(b)$  for all  $a, b \in R[x]$
- (ii).  $f(ab) = f(a)f(b)$  for all  $a, b \in R[x]$

If in addition  $R[x]$  and  $S[x]$  are rings with unities  $1_R$  and  $1_S$ , respectively, then we require that

- (iii).  $f(1_R) = f(1_S)$

We will often drop this distinction and just write “1” and “0” even when more than one ring is involved. A homomorphism  $f$  is called an *embedding* if  $f$  is injective, and an isomorphism if  $f$  is *bijective*. A homomorphism from a ring  $R[x]$  to itself is called an *endomorphism*, and an *isomorphism* from  $R$  to itself is called an *automorphism*.

*Note:* It follows from the definition that if  $f: R[x] \rightarrow S[x]$  is a homomorphism, then the image of the zero element in  $R[x]$  is the zero element in  $S[x]$ , i.e.  $f(0_R) = 0_S$ , because for any  $r$  we have  $f(r) = f(r + 0_R) = f(r) + f(0_R)$  and hence  $f(0_R) = 0_S$ .

**Definition 1.5** [4]

Let  $f: R[x] \rightarrow S[x]$  be a homomorphism of rings. We define the *kernel* of  $f$  by setting

$$\ker(f) = \{a \in R[x] | f(a) = 0\}$$

**Lemma 1.1** [4]

Let  $f: R[x] \rightarrow S[x]$  be a homomorphism of rings. The  $\text{Ker}(f)$  is a proper ideal of  $R[x]$ .

**Proof.**  $\text{Ker}(f) \neq \emptyset$  since  $0 \in \text{Ker}(f)$ . If  $a, b \in \text{Ker}(f)$ , then  $f(a) = f(b) = 0$ , hence  $f(a + b) = f(a) + f(b) = 0 + 0 = 0$ , and thus  $a + b \in \text{Ker}(f)$ . If  $a \in \text{Ker}(f)$  and  $r \in R$ , then  $f(a) = 0$ , hence

$$f(ar) = f(a)f(r) = 0f(r) = 0$$

and thus  $ar \in \text{Ker}(f)$ . The ideal  $\text{Ker}(f)$  is proper since  $f(1_R) = 1_S \neq 0$  and thus  $1_R \notin \text{Ker}(f)$ .  $\square$

**Definition 1.6** [2]

A non-empty subset  $I$  of a ring  $R[x]$  is called *ideal* of  $R[x]$ , if the following are valid:

- (i).  $a - b \in I$ , for all  $a, b \in I$ , and
- (ii).  $ra \in I$ , and  $ar \in I$ , for all  $r \in R$ , and  $a \in I$

If instead of relations  $ra \in I$ , and  $ar \in I$ , is valid only  $ra \in I$ , for all  $r \in R[x]$ , and  $a \in I$ , then  $I$  is called left ideal of  $R[x]$ . Respectively, if is valid only the relation  $ar \in I$ , for all  $r \in R[x]$ , and  $a \in I$ , then  $I$  is called right ideal of  $R[x]$ .

If  $I = \{0\}$  is called *trivial* and if  $I \neq R[x]$  is called *proper*.

Before the notion of what a basis is, we will need some other notions first. The most important notion is that of *linear independence*.

**Definition 1.7** [3]

Let  $v_1, v_2, \dots, v_n$  elements of  $k$ -vector space  $V$ . Linear relation of  $v_1, v_2, \dots, v_n$  is a relation

$$k_1 v_1 + k_2 v_2 + \dots + k_n v_n = 0 \tag{1.1}$$

where  $k_1, k_2, \dots, k_n \in F$ .

The  $v_1, v_2, \dots, v_n$  are called linearly independent if there exists no relation between them, except the trivial, where all coefficients  $k_i$  are zero. Equivalents  $v_1, v_2, \dots, v_n$  are linearly independent if from a linear relation (1.1) we conclude that  $k_i = 0$ , for  $i = 1, 2, \dots, n$

$$k_1 v_1 + k_2 v_2 + \dots + k_n v_n = 0 \Leftrightarrow k_1 = k_2 = \dots = k_n = 0.$$

**Example 1.2**

- A. Elements  $(1,0)$  and  $(0,1)$  of  $\mathbb{R}$ -vector space  $\mathbb{R}^2$  is linearly independent, since if

$$k_1(1,0) + k_2(0,1) = (0,0) \Rightarrow (k_1, k_2) = (0,0) \Rightarrow k_1 = k_2 = 0$$

- B. In  $k$ -vector space  $F^n$  denote

$$e_1 = (1,0,0, \dots, 0), e_2 = (0,1,0, \dots, 0), \dots, e_n = (0,0, \dots, 0,1)$$

The elements  $e_1, e_2, \dots, e_n$  are linearly independent since if

$$k_1 e_1 + k_2 e_2 + \dots + k_n e_n = (0, 0, \dots, 0) \Rightarrow (k_1, k_2, \dots, k_n) = (0, 0, \dots, 0)$$

$$\Rightarrow k_1 = k_2 = \dots = 0$$

We also observe that, for the random element  $(a_1, a_2, \dots, a_n) \in F^n$ , is valid that

$$(a_1, a_2, \dots, a_n) = a_1 e_1 + a_2 e_2 + \dots + a_n e_n$$

Consequently,  $F^n = S(\{e_1, e_2, \dots, e_n\})$ , where  $S(X)$ : the space spanned by the set of  $X$ .

**Definition 1.8** [3]

*Basis* of a  $k$ -vector space  $V$  is called a linearly independent generating set of  $V$  and it is denoted by  $B$ , namely

$$B \text{ basis of } V \Leftrightarrow V = S(B) \text{ and } B \text{ is linearly independent}$$

**Example 1.3**

A. The set  $B = \{e_1, e_2\}$  where  $e_1 = (1, 0)$  and  $e_2 = (0, 1)$  is a spanning set of  $\mathbb{R}^2$ . It is also linearly independent for the only solution of the vector equation  $c_1 e_1 + c_2 e_2 = 0$  is the trivial solution. Therefore,  $B$  is a basis for  $\mathbb{R}^2$ . It is called the standard basis for  $\mathbb{R}^2$ .

B. The elements  $x, x^2, \dots, x^n, \dots$  of  $R[x]$  are a basis of  $k$ -vector space  $R[x]$ . The set  $\{x, x^2, \dots, x^n, \dots\}$  of  $R[x]$  is linearly independent (since  $k_1 x^{i_1} + k_2 x^{i_2} + \dots + k_s x^{i_s} = 0$  then  $k_j = 0, 1 \leq j \leq s$  namely every finite subset of  $\{x, x^2, \dots, x^n, \dots\}$  is linearly independent) and every polynomial  $P(x) = a_0 + a_1 x + \dots + a_n x^n$  belongs to  $S(x)$ .

## 1.2 Monomial Ordering

---

Before we introduce the notion of monomial ordering we should know the form of a polynomial with more than one variable. Below is given first the definition of polynomials of one and then of  $n$  variables.

### Definition 1.9 [1]

Let  $R[x]$  a ring. A *polynomial*  $f(x)$  with coefficients in  $R[x]$  is an infinity typical sum

$$\sum_{i=0}^{\infty} a_i x^i = a_0 + a_1 x^1 + \dots + a_n x^n + \dots \quad (1.2)$$

where  $a_i \in R[x]$  and  $a_i = 0$  except for finite range of values of  $i$ . The  $a_i$  are coefficients of  $f(x)$ . If for some  $i > 0$  is valid  $a_i \neq 0$ , the largest of them is called *degree* of  $f(x)$ . If there exists no such a value, then we say that  $f(x)$  has zero degree.

### Definition 1.10 [10]

A *monomial* is a product of powers of variables with nonnegative integer exponents, or, in other words, a product of variables, possibly with repetitions.

The constant 1 is a monomial, being equal to the empty product and  $x^0$  for any variable  $x$  is considered, this means that a monomial is either 1 or a power  $x^n$  of  $x$ , with  $n$  a positive integer. If several variables are considered, say,  $x, y, z$  then each can be given an exponent, so that any monomial is of the form  $x^a y^b z^c$  with  $a, b, c$  non-negative integers.

Every monomial has a representation of the form  $a x_1^{\beta_1} x_2^{\beta_2} \dots x_n^{\beta_n}$  ( $a \in F, \beta \in \mathbb{N}^n$ ).

### Definition 1.11 [5]

We consider polynomials  $f(x_1, x_2, \dots, x_n)$  in  $n$  variables with coefficients in  $F$ . Such polynomials are finite sums of the form  $a x_1^{\beta_1} x_2^{\beta_2} \dots x_n^{\beta_n}$ , where  $a \in F$ , and  $\beta_i \in \mathbb{N}$ ,  $i = 1, \dots, n$ . We call  $x_1^{\beta_1} x_2^{\beta_2} \dots x_n^{\beta_n}$  a *power product*.



**Example 1.4**

The  $f = x_1^3 + x_2^2 + 1$  and  $g = x_1^2 - 4x_2 + 8x_1x_3$  are polynomials in  $n$  variables.

The basis of  $R[x_1, x_2, \dots, x_n]$ , which is a  $k$ -vector space, is the set,  $\mathbb{T}^n$ , of all power products,

$$\mathbb{T}^n = \{x_1^{\beta_1} x_2^{\beta_2} \dots x_n^{\beta_n} \mid \beta_i \in \mathbb{N}, i = 1, \dots, n\} \quad (1.3)$$

Sometimes we will denote  $x_1^{\beta_1} x_2^{\beta_2} \dots x_n^{\beta_n}$  by  $x^\beta$ , where  $\beta = (\beta_1, \beta_2, \dots, \beta_n) \in \mathbb{N}^n$ .

**Definition 1.12** [9]

The *degree* of an element  $m = x_1^{\beta_1} x_2^{\beta_2} \dots x_n^{\beta_n}$  in a polynomial ring  $R[x_1, x_2, \dots, x_n]$  is

$$\deg(m) = \beta_1 + \beta_2 + \dots + \beta_n \quad (1.4)$$

The *degree* of a non-zero polynomial  $f(x_1, x_2, \dots, x_n) = \sum a_{\beta_1, \beta_2, \dots, \beta_n} x_1^{\beta_1} x_2^{\beta_2} \dots x_n^{\beta_n}$  equals

$$\deg(f) = \max \{ \deg(x_1^{\beta_1} x_2^{\beta_2} \dots x_n^{\beta_n}) \mid a_{\beta_1, \beta_2, \dots, \beta_n} \neq 0 \} \quad (1.5)$$

To introduce the notion of *Gröbner bases* and the algorithm which helps us to find them, we should be able to compare any two power products.

The orders in the linear and one variable cases are used to define a division (or reduction) algorithm.

**Definition 1.13** [5]

By a *term order* on  $\mathbb{T}^n$  we mean a total order  $<$  on  $\mathbb{T}^n$  satisfying the following two conditions:

- (i).  $1 < x^\beta$  for all  $x^\beta \in \mathbb{T}^n, x^\beta \neq 1$
- (ii). If  $x^a < x^\beta$ , then  $x^a x^\gamma \leq x^\beta x^\gamma$ , for all  $x^\gamma \in \mathbb{T}^n$ .

**Proposition 1.1** [5]

For  $x^a, x^\beta \in \mathbb{T}^n$ , if  $x^a$  divided  $x^\beta$  then  $x^a \leq x^\beta$ .

**Proof.** By assumption there is an  $x^\gamma \in \mathbb{T}^n$  such that  $x^\beta = x^a x^\gamma$ . By condition (i) in Definition 1.11 we have  $x^\gamma \geq 1$  and so by condition (ii) we have  $x^\beta = x^a x^\gamma \geq x^a$ , as desired.  $\square$

To compare any two power products, the order must be a total order, that is, given any  $x^a, x^\beta \in \mathbb{T}^n$ , exactly one of the following relations must hold

$$x^a < x^\beta \text{ or } x^a > x^\beta \text{ or } x^a = x^\beta \quad (1.6)$$

Also, for any  $x^a, x^\beta, x^\gamma \in \mathbb{T}^n$  we have:

$$\text{if } x^a < x^\beta \text{ and } x^\beta < x^\gamma \text{ then } x^a < x^\gamma \quad (1.7)$$

If we want the reduction to be finite, we need that order be a well-ordering, that is, there is no infinite descending chain  $x^{\beta_1} > x^{\beta_2} > \dots$  in  $\mathbb{T}^n$ .

The most frequently used descriptions of ordering have at most two defining conditions: a degree and a (normal or reverse) lexicographical comparison. The most famous are:

- a) *Lexicographic order*
- b) *Reverse lexicographic order*
- c) *Degree lexicographic order*
- d) *Degree reverse lexicographic order*

- Let  $a = (a_1, a_2, \dots, a_n)$  and  $\beta = (\beta_1, \beta_2, \dots, \beta_n)$  be vectors belonging to  $\mathbb{Z}_+^n$ . We define the total order  $<_{lex}$  on  $\mathbb{T}^n$  by setting  $x^a <_{lex} x^\beta$  if either (i)  $\sum_{i=1}^n a_i < \sum_{i=1}^n \beta_i$ , or (ii)  $\sum_{i=1}^n a_i = \sum_{i=1}^n \beta_i$  and the left most non-zero component of the vector  $a - \beta$  is negative. It follows that  $<_{lex}$  is a monomial order on  $R[x_1, x_2, \dots, x_n]$  which is called the **lexicographic order** on  $R[x_1, x_2, \dots, x_n]$  included by the ordering  $x_1 > x_2 > \dots > x_n$ .
- Let  $a = (a_1, a_2, \dots, a_n)$  and  $\beta = (\beta_1, \beta_2, \dots, \beta_n)$  be vectors belonging to  $\mathbb{Z}_+^n$ . We define the total order  $<_{rev}$  on  $\mathbb{T}^n$  by setting  $x^a <_{rev} x^\beta$  if either (i)  $\sum_{i=1}^n a_i < \sum_{i=1}^n \beta_i$ , or (ii)  $\sum_{i=1}^n a_i = \sum_{i=1}^n \beta_i$  and the right most non-zero component of the vector  $a - \beta$  is negative. It follows that  $<_{rev}$  is a monomial order on  $R[x_1, x_2, \dots, x_n]$  which is called the **reverse lexicographic order** on  $R[x_1, x_2, \dots, x_n]$  included by the ordering  $x_1 > x_2 > \dots > x_n$ .
- Let  $a = (a_1, a_2, \dots, a_n)$  and  $\beta = (\beta_1, \beta_2, \dots, \beta_n)$  be vectors belonging to  $\mathbb{Z}_+^n$ . We define the total order  $<_{deglex}$  on  $\mathbb{T}^n$  by setting  $x^a <_{deglex} x^\beta$  if either (i)  $\sum_{i=1}^n a_i < \sum_{i=1}^n \beta_i$ , or (ii)  $\sum_{i=1}^n a_i = \sum_{i=1}^n \beta_i$  and  $x^a < x^\beta$  with respect to

lexicographic order with  $x_1 > x_2 > \dots > x_n$ . It follows that  $<_{deglex}$  is a monomial order on  $R[x_1, x_2, \dots, x_n]$  which is called the **degree reverse lexicographic order** on  $R[x_1, x_2, \dots, x_n]$  included by the ordering  $x_1 > x_2 > \dots > x_n$ .

- Let  $a = (a_1, a_2, \dots, a_n)$  and  $\beta = (\beta_1, \beta_2, \dots, \beta_n)$  be vectors belonging to  $\mathbb{Z}_+^n$ . We define the total order  $<_{degrevlex}$  on  $\mathbb{T}^n$  by setting  $x^a <_{degrevlex} x^\beta$  if either (i)  $\sum_{i=1}^n a_i < \sum_{i=1}^n \beta_i$ , or (ii)  $\sum_{i=1}^n a_i = \sum_{i=1}^n \beta_i$  and  $x^a < x^\beta$  the right most non-zero component of the vector  $a - \beta$  is positive. It follows that  $<_{deglex}$  is a monomial order on  $R[x_1, x_2, \dots, x_n]$  which is called the **degree reverse lexicographic order** on  $R[x_1, x_2, \dots, x_n]$  included by the ordering  $x_1 > x_2 > \dots > x_n$ .

### Example 1.5

A. Consider the polynomial

$$p(x_1, x_2, x_3) = 2x_1^2x_2^3 + 5x_1x_2x_3 - 7x_1^2x_3^3$$

We assume that  $x_1 > x_2 > x_3$ .

Let be  $\alpha=(2,3,0)$ ,  $\beta=(1,1,1)$  and  $\gamma=(2,0,3)$ , so  $\alpha-\beta=(1,2,-1)$ ,  $\beta-\gamma=(-1,1,-2)$ ,  $\alpha-\gamma=(0,3,-3)$  and  $|\alpha|=5, |\beta|=3$  and  $|\gamma|=5$ .

- With respect to **Lexicographic order**, p is written in decreasing order as :

$$p(x_1, x_2, x_3) = 2x_1^2x_2^3 - 7x_1^2x_3^3 + 5x_1x_2x_3$$

Because the left-most non-zero entries of  $\alpha-\beta$  and  $\alpha-\gamma$  are positive, therefore  $x_1^2x_2^3 >_{lex} x_1x_2x_3$  and  $x_1^2x_2^3 >_{lex} x_1^2x_3^3$  respectively. But the left-most non-zero entry of  $\beta-\gamma$  is negative, therefore  $x_1^2x_3^3 >_{lex} x_1x_2x_3$ .

- With respect to **Reverse lexicographic order**, p is written in decreasing order as :

$$p(x_1, x_2, x_3) = -7x_1^2x_3^3 + 5x_1x_2x_3 + 2x_1^2x_2^3$$

Because the right-most non-zero entries of  $\alpha-\beta$ ,  $\alpha-\gamma$  and  $\beta-\gamma$  are negative, therefore  $x_1x_2x_3 >_{revlex} x_1^2x_2^3$ ,  $x_1^2x_3^3 >_{revlex} x_1^2x_2^3$  and  $x_1^2x_3^3 >_{revlex} x_1x_2x_3$  respectively.

- With respect to **Degree lexicographic order**, p is written in decreasing order as :

$$p(x_1, x_2, x_3) = 2x_1^2x_2^3 - 7x_1^2x_3^3 + 5x_1x_2x_3$$

Because  $x_1^2x_2^3 >_{lex} x_1x_2x_3$ ,  $x_1^2x_2^3 >_{lex} x_1^2x_3^3$ ,  $x_1^2x_3^3 >_{lex} x_1x_2x_3$  and  $|\alpha| > |\beta|$ ,  $|\alpha| = |\gamma|$ ,  $|\beta| < |\gamma|$ , therefore  $x_1^2x_2^3 >_{deglex} x_1x_2x_3$ ,  $x_1^2x_3^3 >_{deglex} x_1^2x_3^3$  and  $x_1^2x_3^3 >_{deglex} x_1x_2x_3$  respectively.

- With respect to **Degree Reverse lexicographic ordering**,  $p$  is written in decreasing order as :

$$p(x_1, x_2, x_3) = 2x_1^2x_2^3 - 7x_1^2x_3^3 + 5x_1x_2x_3$$

Because the right-most non-zero entries of  $\alpha - \beta$ ,  $\alpha - \gamma$ ,  $\beta - \gamma$  are negative and  $|\alpha| > |\beta|$ ,  $|\alpha| = |\gamma|$ ,  $|\beta| < |\gamma|$ , therefore  $x_1^2x_2^3 >_{degrevlex} x_1x_2x_3$ ,  $x_1^2x_2^3 >_{degrevlex} x_1^2x_3^3$  and  $x_1^2x_3^3 >_{degrevlex} x_1x_2x_3$  respectively.

B. Consider the polynomial

$$p(x_1, x_2, x_3, x_4) = x_1x_4^2 - 3x_1x_2x_3 + 8x_2x_4^3$$

We solve the exercise with  $x_1 > x_2 > x_3 > x_4$ .

Let be  $a = (1, 0, 0, 2)$ ,  $\beta = (1, 1, 1, 0)$  and  $\gamma = (0, 1, 0, 3)$ , so  $a - \beta = (0, -1, -1, 2)$ ,  $a - \gamma = (1, -1, 0, -1)$  and  $\beta - \gamma = (1, 0, 1, -3)$ , and  $|a| = 3 = |\beta|$ ,  $|\gamma| = 4$ .

- With **Lexicographic order**,  $p$  is written in decreasing order as:

$$p(x_1, x_2, x_3, x_4) = -3x_1x_2x_3 + x_1x_4^2 + 8x_2x_4^3$$

Because  $|a| = |\beta|$ ,  $|a| < |\gamma|$  and  $|\beta| < |\gamma|$ , and the left-most non-zero component of the vector  $a - \beta$  is negative and  $a - \gamma$ ,  $\beta - \gamma$  is positive, therefore  $x_1x_2x_3 >_{lex} x_1x_4^2$ ,  $x_1x_4^2 >_{lex} x_2x_4^3$  and  $x_1x_2x_3 >_{lex} x_2x_4^3$  respectively.

- With **Reverse lexicographic order**,  $p$  is written in decreasing order as:

$$p(x_1, x_2, x_3, x_4) = -3x_1x_2x_3 + x_1x_4^2 + 8x_2x_4^3$$

Because  $|a| = |\beta|$ ,  $|a| < |\gamma|$  and  $|\beta| < |\gamma|$ , and the right-most non-zero component of the vector  $a - \beta$  is positive and  $a - \gamma$ ,  $\beta - \gamma$  is negative, therefore  $x_1x_2x_3 >_{lex} x_1x_4^2$ ,  $x_1x_4^2 >_{lex} x_2x_4^3$  and  $x_1x_2x_3 >_{lex} x_2x_4^3$  respectively.

- With **Degree lexicographic order**,  $p$  is written in decreasing order as:

$$p(x_1, x_2, x_3, x_4) = -3x_1x_2x_3 + x_1x_4^2 + 8x_2x_4^3$$

Because  $|a| = |\beta|$ ,  $|a| < |\gamma|$  and  $|\beta| < |\gamma|$ , and  $x_1x_2x_3 >_{lex} x_1x_4^2$ ,  $x_1x_4^2 >_{lex} x_2x_4^3$ ,  $x_1x_2x_3 >_{lex} x_2x_4^3$ , therefore  $x_1x_2x_3 >_{deglex} x_1x_4^2$ ,  $x_1x_4^2 >_{deglex} x_2x_4^3$  and  $x_1x_2x_3 >_{deglex} x_2x_4^3$  respectively.

- With **Degree Reverse lexicographic order**,  $p$  is written in decreasing order as:

$$p(x_1, x_2, x_3, x_4) = -3x_1x_2x_3 + x_1x_4^2 + 8x_2x_4^3$$

Because  $|a| = |\beta|$ ,  $|a| < |\gamma|$  and  $|\beta| < |\gamma|$ , and the right-most non-zero component of the vector  $a - \beta$  is positive and  $\alpha - \gamma$ ,  $\beta - \gamma$  is negative, therefore  $x_1x_2x_3 >_{degrevlex} x_1x_4^2$ ,  $x_1x_4^2 >_{degrevlex} x_2x_4^3$  and  $x_1x_2x_3 >_{degrevlex} x_2x_4^3$  respectively.

**Definition 1.14** [9]

We call an ordering a *degree ordering* if the most important criterion for comparison is the degree of the monomials. In the sequel we assume that a monomial order for  $\mathbb{T}^n$  has been fixed.

**Definition 1.15** [6]

Let be a non-zero polynomial in  $R[x_1, x_2, \dots, x_n]$ .

One can write:

$$f = \sum_{i=1}^r a_i m_i \tag{1.8}$$

with  $a_i \in F - \{0\}$ ,  $m_i \in \mathbb{T}^n$  and  $m_1 > \dots > m_r$ .

- the *leading term*  $m_1$  of  $f$  is denoted by  $lt(f)$
- the *leading coefficient*  $a_1$  of  $f$  is denoted by  $lc(f)$  and
- the *leading monomial*  $a_1 m_1$  is denoted by  $lm(f)$ .

**Example 1.6**

A. Consider the polynomial which is used in Example 1.5A:

$$p(x_1, x_2, x_3) = 2x_1^2x_2^3 + 5x_1x_2x_3 - 7x_1^2x_3^3$$

we prove that :

- with respect to **lexicographic ordering**, p is written in decreasing order as:

$$p(x_1, x_2, x_3) = 2x_1^2x_2^3 - 7x_1^2x_3^3 + 5x_1x_2x_3$$

and therefore it can be seen from the above definition that :

- i.  $lt(p) = x_1^2x_2^3$
- ii.  $lc(p) = 2$
- iii.  $lm(p) = 2x_1^2x_2^3$

- with respect to **Reverse lexicographic ordering**, p is written in decreasing order as:

$$p(x_1, x_2, x_3) = -7x_1^2x_3^3 + 5x_1x_2x_3 + 2x_1^2x_2^3$$

- i.  $lt(p) = x_1^2x_3^3$
- ii.  $lc(p) = -7$
- iii.  $lm(p) = -7x_1^2x_3^3$

- with respect to **Degree lexicographic ordering**, p is written in decreasing order as:

$$p(x_1, x_2, x_3) = 2x_1^2x_2^3 - 7x_1^2x_3^3 + 5x_1x_2x_3$$

- i.  $lt(p) = x_1^2x_2^3$
- ii.  $lc(p) = 2$
- iii.  $lm(p) = 2x_1^2x_2^3$

- with respect to **Degree Reverse lexicographic ordering**, p is written in decreasing order as:

$$p(x_1, x_2, x_3) = 2x_1^2x_2^3 - 7x_1^2x_3^3 + 5x_1x_2x_3$$

- i.  $lt(p) = x_1^2x_2^3$
- ii.  $lc(p) = 2$

iii.  $lm(p) = 2x_1^2x_2^3$

B. Consider the polynomial which is used in Example 1.5B:

$$p(x_1, x_2, x_3, x_4) = x_1x_4^2 - 3x_1x_2x_3 + 8x_2x_4^3$$

we prove that :

- with respect to **Lexicographic ordering**, p is written in decreasing order as:

$$p(x_1, x_2, x_3, x_4) = -3x_1x_2x_3 + x_1x_4^2 + 8x_2x_4^3$$

and therefore it can be seen from the above definition that :

- i.  $lt(p) = x_1x_2x_3$
- ii.  $lc(p) = -3$
- iii.  $lm(p) = -3x_1x_2x_3$

- with respect to **Reverse lexicographic ordering**, p is written in decreasing order as:

$$p(x_1, x_2, x_3, x_4) = -3x_1x_2x_3 + x_1x_4^2 + 8x_2x_4^3$$

- i.  $lt(p) = x_1x_2x_3$
- ii.  $lc(p) = -3$
- iii.  $lm(p) = -3x_1x_2x_3$

- with respect to **Degree lexicographic ordering**, p is written in decreasing order as:

$$p(x_1, x_2, x_3, x_4) = -3x_1x_2x_3 + x_1x_4^2 + 8x_2x_4^3$$

- i.  $lt(p) = x_1x_2x_3$
- ii.  $lc(p) = -3$
- iii.  $lm(p) = -3x_1x_2x_3$

- with respect to **Degree Reverse lexicographic ordering**, p is written in decreasing order as:

$$p(x_1, x_2, x_3, x_4) = -3x_1x_2x_3 + x_1x_4^2 + 8x_2x_4^3$$

- i.  $lt(p) = x_1x_2x_3$
- ii.  $lc(p) = -3$
- iii.  $lm(p) = -3x_1x_2x_3$

**Definition 1.16** [7]

Let  $I$  be an ideal of  $R[x_1, x_2, \dots, x_n]$ . The initial ideal of  $I$ , denoted by  $in(I)$ , is given by

$$in(I) = (\{lt(f) \mid 0 \neq f \in I\}). \quad (1.9)$$



## 1.3 Division algorithm

---

Before we define *Gröbner bases*, we have to deal with the division algorithm for polynomials in more than one variables.

Firstly, we will see *Dickson's lemma*, which is important because it will justify our choice of  $\mathbb{T}^n$  to be a finite set, in our examples.

### Lemma 1.1 [9]

Every monomial ideal in  $R[x_1, x_2, \dots, x_n]$  is finitely generated.

**Proof.** Let  $S = \mathbb{T}^n$ . We proceed by induction on  $n$ . Since in  $R[x]$  is principal, the lemma is true for  $n = 1$ . Suppose it is true for  $n - 1$  variables, and let  $a$  be a monomial ideal in  $R[x_1, x_2, \dots, x_n]$ . Let

$$b_j = (a = \langle x_n^j \rangle) \cap R[x_1, x_2, \dots, x_{n-1}] = \langle S_j \rangle.$$

Since  $b_j$  is an ideal in  $R[x_1, x_2, \dots, x_{n-1}]$ , we can choose  $S_j$  to be finite. We have  $b_0 \subseteq b_1 \subseteq \dots$ . It follows easily that  $\cup b_j$  is an ideal  $b$  in  $R[x_1, x_2, \dots, x_{n-1}]$  and hence finitely generated  $b = \langle S \rangle$ . If  $m \in a$  is a monomial then  $m = m' x_n^k$  for some monomial  $m \in R[x_1, x_2, \dots, x_{n-1}]$  and some  $k$ .

Since  $m' x_n^k \in a$  we get  $m' \in a : \langle x_n^k \rangle$ , so  $m \in \langle x_n^k S_k \rangle$ . Thus  $S' = S_0 \cup x_n S_1 \cup x_n^2 S_2 \cup \dots$  is a generating set for  $a$ . But for some  $r$  we have that  $S_k = S_r$  if  $k \geq r$ , so  $S_0 \cup x_n S_1 \cup \dots \cup x_n^r S_r$  is a (finite) generating set for  $a$ .

### Definition 1.17 [5]

Given  $f, g, h$  in  $R[x_1, x_2, \dots, x_n]$  with  $g \neq 0$ , we say that  $f$  reduces to  $h$  modulo  $g$  in one step, written  $f \xrightarrow{g} h$ , if and only if  $lt(g)$  divides a non-zero term  $X$  that appears in  $f$  and

$$h = f - \frac{X}{lt(g)} \tag{1.10}$$

### Example 1.7

Let  $f = x^2y + 4xy - 3y^2$ ,  $g = 2x + 4y + 1$ . Also, let the order be lexicographic with  $x > y$ . We have  $lt(f) = x^2y$  and  $lt(g) = x$ .

- $lt(g)/lt(f) \Rightarrow f = \frac{xy}{2}g + p_1$ , where  $p_1 = -2xy^2 + \frac{7}{2}xy - 3y^2$
- $lt(g)/lt(p_1) \Rightarrow f = \frac{xy}{2}g - y^2g + p_2$ , where  $p_2 = \frac{7}{2}xy + 4y^3 - 2y^2$
- $lt(g)/lt(p_2) \Rightarrow f = \frac{xy}{2}g - y^2g + \frac{7}{4}g + p_3$ , where  $p_3 = 4y^3 - 9y^2 - \frac{7}{4}y$
- $lt(g)$  doesn't divide  $lt(p_3)$

So, we have

$$f \xrightarrow{g} -2xy^2 + \frac{7}{2}xy - 3y^2 \xrightarrow{\frac{7}{2}g} \frac{7}{2}xy + 4y^3 - 2y^2 \xrightarrow{4y^3 - 9y^2 - \frac{7}{4}y}$$

In the multivariable case we may have to divide by more than one polynomial at a time, and so we extend the process of reduction defined above to include this more general setting.

**Definition 1.18** [5]

Let  $f, h$  and  $f_1, \dots, f_s$  be polynomials in  $R[x_1, x_2, \dots, x_n]$ , with  $f_i \neq 0$  ( $1 \leq i \leq s$ ), and let  $F = \{f_1, \dots, f_s\}$ . We say that  $f$  reduces to  $h$  modulo  $F$ , denoted

$$f \xrightarrow{F} h,$$

if and only if there exist a sequence of indices  $i_1, \dots, i_t \in \{1, \dots, s\}$  and a sequence of polynomial  $h_1, \dots, h_t \in R[x_1, x_2, \dots, x_n]$  such that

$$f \xrightarrow{f_{i_1}} h_1 \xrightarrow{f_{i_2}} h_2 \xrightarrow{f_{i_3}} \dots \xrightarrow{f_{i_{t-1}}} h_{t-1} \xrightarrow{f_{i_t}} h$$

**Example 1.8**

Let  $f = xy^2 + 2x$ ,  $f_1 = xy - y$ ,  $f_2 = y^2 - 1$ . Also, let the order be lexicographic with  $x > y$ . We have  $lt(f) = xy^2, lt(f_1) = xy$  and  $lt(f_2) = y^2$ .

- $lt(f_1)/lt(f) \Rightarrow f = yf_1 + p_1$ , where  $p_1 = y^2 + 2x$
- $lt(f_1)$  doesn't divide  $lt(p_1)$ , while
- $lt(f_2)/lt(p_1) \Rightarrow f = yf_1 + f_2 + p_2$ , where  $p_2 = 2x + 1$
- None of  $lt(f_i)$  doesn't divide  $lt(p_2)$ , so  $p_2 = 2x + 1$ .

Eventually,

$$f = yf_1 + f_2 + 2x + 1$$

The above process according to Definition 1.18 can be described as follows:

$$f \xrightarrow{f_1} xy^2 + 2x \xrightarrow{f_2} 2x + 1$$

It's important to say that quotients and remainders of division depending as much on the monomial ordering as the order of polynomials.

Now we will present the *Generalized Division Algorithm*:

**Theorem 1.1** [7]

Let  $S = R[x_1, x_2, \dots, x_n]$  denote the polynomial ring in  $n$  variables over a field  $F$  and fix a monomial order  $<$  on  $S$ . Let  $g_1, \dots, g_s$  be non-zero polynomial of  $S$ . Then, given a polynomial  $0 \neq f \in S$ , there exist polynomials  $f_1, \dots, f_s$  and  $f'$  of  $S$  with

$$f = f_1g_1 + f_2g_2 + \dots + f_s g_s + f' \quad (1.11)$$

such that the following conditions are satisfied:

- (i). if  $f' \neq 0$  and if  $u \in \text{supp}(f')$ , where  $\text{supp}(f') = \{u \in \mathbb{T}^n : a_u \neq 0\}$ , then none of the divides  $u$ , i.e. no monomial  $u \in \text{supp}(f')$  belongs to  $(\text{lt}(g_1), \dots, \text{lt}(g_s))$ ;
- (ii). if  $f_i \neq 0$ , then

$$\text{lt}(f) \geq \text{lt}(f_i g_i)$$

The right-hand side of equation (1.11) is said to be a standard expression for  $f$  with respect to  $g_1, \dots, g_s$  and the polynomial  $f'$  is said to be a remainder of  $f$  with respect to  $g_1, \dots, g_s$ .

To prove the Theorem 1.1, we need the following lemma:

**Lemma 1.2** [7]

Let be a monomial order on  $S = R[x_1, x_2, \dots, x_n]$ . Then, for any monomial  $u$  of  $S$ , there is no infinite descending sequence of the form

$$\dots < u_2 < u_1 < u_0 = u \quad (1.12)$$

**Proof.** Suppose, on the contrary, that one has an infinite descending sequence (1.12) and write  $M$  for the set of monomials  $\{u_0, u_1, \dots\}$ . It follows from Dickson’s lemma (Lemma 1.1) that  $M^{\min}$  is a finite set, say  $M^{\min} = \{u_{i_1}, \dots, u_{i_s}\}$  with  $i_1 < i_2 < \dots < i_s$ . Then the monomial  $u_{i_{s+1}}$  is divided by  $u_{i_j}$  for some  $1 \leq j \leq s$ . Thus  $u_{i_j} < u_{i_{s+1}}$ , which contradicts  $i_j < i_{s+1}$ .  $\square$

Now, we are ready to prove *Theorem 1.1*:

**Proof.** Let  $I = (\text{lt}(g_1), \dots, \text{lt}(g_s))$ . If none of the monomials  $u \in \text{supp}(f)$  belongs to  $I$  and write respect to  $<$  among the monomials  $u \in \text{supp}(f)$  belonging to  $I$ . Let, say,  $\text{lt}(g_{i_0})$  divide  $u_0$  and  $w_0 = u_0/\text{lt}(g_{i_0})$ .

We rewrite

$$f = c'_0 c_{i_0}^{-1} w_0 g_{i_0} + h_1,$$

where  $c'_0$  is the coefficient of  $u_0$  in  $f$  and  $c_{i_0}$  is that of  $\text{lt}(g_{i_0})$  in  $g_{i_0}$ . One has

$$lt(w_0 g_{i_0}) = w_0 lt(g_{i_0}) = u_0 \leq lt(f)$$

If either  $h_1 = 0$  or, in case of  $h_1 \neq 0$ , none of the monomials  $u \in \text{supp}(h_1)$  belongs to  $I$ , then  $f = c'_0 c_{i_0}^{-1} w_0 g_{i_0} + h_1$  is a standard expression of  $f$  with respect to  $g_1, \dots, g_s$  and  $h_1$  is a remainder of  $f$ .

If a monomial of  $\text{supp}(h_1)$  belongs to  $I$  and if  $u_1$  is the monomial which is biggest with respect to  $<$  among the monomials  $u \in \text{supp}(h_1)$  belonging to  $I$ , then one has

$$u_0 > u_1$$

In fact, if a monomial  $u$  with  $u > u_0 (= \text{lm}(w_0 g_{i_0}))$  belongs to  $\text{supp}(h_1)$ , then  $u$  must belong to  $\text{supp}(f)$ . This is impossible. Moreover,  $u_0$  itself cannot belong to  $\text{supp}(h_1)$ .

Let, say,  $lt(g_{i_1})$  divide  $u_1$  and  $w_1 = u_1 / lt(g_{i_1})$ . Again, we rewrite

$$f = c'_0 c_{i_0}^{-1} w_0 g_{i_0} + c'_1 c_{i_1}^{-1} w_1 g_{i_1} + h_2$$

where  $c'_1$  is the coefficient of  $u_1$  in  $h_1$  and  $c_{i_1}$  is that of  $lt(g_{i_1})$  in  $g_{i_1}$ . One has

$$lt(w_1 g_{i_1}) < lt(w_0 g_{i_0}) \leq lt(f).$$

Continuing these procedures yields the descending sequence

$$u_0 > u_1 > u_2 > \dots$$

*Lemma 1.2* thus guarantees that these procedures will stop after a finite number of steps, say  $N$  steps, and we obtain an expression

$$f = \sum_{q=0}^{N-1} c'_q c_{i_q}^{-1} w_q g_{i_q} + h_N,$$

where either  $h_N = 0$  or, in case  $h_N \neq 0$ , none of the monomials  $u \in \text{supp}(h_N)$  belongs to  $I$ , and where

$$lt(w_q g_{i_q}) < \dots < lt(w_0 g_{i_0}) \leq lt(f).$$

Thus, by letting  $\sum_{i=1}^s f_i g_i = \sum_{q=0}^{N-1} c'_q c_{i_q}^{-1} w_q g_{i_q}$  and  $f' = h_N$ , we obtain an expression  $f = \sum_{i=1}^s f_i g_i + f'$  satisfying the conditions (i) and (ii), as desired.  $\square$

**Example 1.9**

Let  $f = x^2 + y^2 - 1$ ,  $f_1 = x + y$  and  $f_2 = y + 1$ . Als, let the order be lexicographic with  $x > y$ . We have  $lt(f) = x^2$ ,  $lt(f_1) = x$  and  $lt(f_2) = y$ .

- $lt(f_1)/lt(f) \Rightarrow f = xf_1 + p_1$ , where  $p_1 = -xy + y^2 - 1$
- $lt(f_1) / lt(p_1) \Rightarrow f = xf_1 - yf_1 + p_2$ , where  $p_2 = 2y^2 - 1$
- $lt(f_1)$  doesn't divide  $lt(p_2)$ , while
- $lt(f_2) / lt(p_2) \Rightarrow f = xf_1 - yf_1 + 2yf_2 + p_3$ , where  $p_3 = -2y - 1$
- $lt(f_2) / lt(p_3) \Rightarrow f = xf_1 - yf_1 + 2yf_2 - 2f_2 + p_4$ , where  $p_4 = 1$

Eventually,

$$f = xf_1 - yf_1 + 2yf_2 - 2f_2 + 1,$$

or

$$f \xrightarrow{f_1} -xy + y^2 - 1 \xrightarrow{f_1} 2y^2 - 1 \xrightarrow{f_2} -2y - 1 \xrightarrow{f_2} 1$$

## 1.4 Gröbner bases

---

Gröbner bases were introduced in 1965, together with an algorithm to compute them (Buchberger; algorithm), by *Bruno Buchberger* in his PhD thesis. He named them after his advisor *Wolfgang Gröbner*. In 2007, Buchberger received the Association for Computing Machinery’s Paris Kanellakis Theory and Practice Award for his work. An analogous concept for local rings was developed independently by Heisuke Hirokawa in 1964 who named them standard bases.

### Definition 1.19 [6]

Let  $I \neq (0)$  be an ideal of  $R[x_1, x_2, \dots, x_n]$  and let  $G = \{g_1, g_2, \dots, g_s\}$  be a subset of  $I$ . The set  $G$  is called a *Gröbner basis* of  $I$  if

$$\text{in}(I) = (\text{lt}(g_1), \text{lt}(g_2), \dots, \text{lt}(g_s)) \quad (1.13)$$

where  $g_1, g_2, \dots, g_s$  are ordered with respect to a common ordering.

### Theorem 1.2 [5]

Let  $I$  be a non-zero ideal of  $R[x_1, x_2, \dots, x_n]$ . The following statements are equivalent for a set of non-zero polynomials  $G = \{g_1, g_2, \dots, g_s\} \subseteq I$

- (i).  $G$  is a Gröbner basis for  $I$
- (ii).  $f \in I$  if and only if  $f \xrightarrow{G}_+ 0$
- (iii).  $f \in I$  if and only if  $f = \sum_{i=1}^s h_i g_i$  with  $\text{lt}(f) = \max_{1 \leq i \leq s} (\text{lt}(h_i) \text{lt}(g_i))$
- (iv).  $Lm(G) = Lm(I)$ , where  $Lm(S) = \{lm(s) | s \in S\}$ .

**Proof.** (i)  $\implies$  (ii) Let  $f \in R[x_1, x_2, \dots, x_n]$ . Then by *Theorem 1.1*, there exists  $f' \in R[x_1, x_2, \dots, x_n]$ , reduced with respect to  $G$ , such that  $f \xrightarrow{G}_+ f'$ . Thus  $f - f' \in I$  and so  $f \in I$ , if and only if  $f' \in I$ . Clearly, if  $f' = 0$  (that is  $f \xrightarrow{G}_+ 0$ ), then  $f \in I$ . Conversely, if  $f \in I$  and  $f' \neq 0$  then  $f' \in I$  and by (i), there exists  $i \in \{1, \dots, s\}$  such that  $\text{lt}(g_i)$  divides  $\text{lt}(f')$ . This is a contradiction to the fact that  $f'$  is reduced with respect to  $G$ . Thus  $f' = 0$  and  $f \xrightarrow{G}_+ 0$ .

(ii)  $\implies$  (iii) For  $f \in I$ , we know by hypothesis that  $f \xrightarrow{G}_+ 0$  and since the process of reduction is exactly the same as the Division Algorithm, we see that (iii) follows from *Theorem 1.1*.

(iii)  $\implies$  (iv) Clearly,  $Lm(G) \subseteq Lm(I)$ . For the reverse inclusion it suffices to show that for all  $f \in I$ ,  $lm(f) \in Lm(G)$ , since the  $lm(f)$ ' generate  $Lm(I)$ . Writing  $f$  as in the hypothesis, it immediately follows that

$$lm(f) = \sum_i lm(h_i)lm(g_i),$$

where the sum is over all  $i$  such that  $lt(f) = lt(h_i)lt(g_i)$ . The result follows immediately.

(iv)  $\Rightarrow$  (i) Let  $f \in I$ . Then  $lm(f)$  is in  $Lm(G)$ , and hence

$$lm(f) = \sum_{i=1}^t h_i lm(g_i), \quad (1.14)$$

for some  $h_i \in R[x_1, x_2, \dots, x_n]$ . If we expand the right-hand side of Equation we see that each term is divisible by some  $lt(g_i)$ . Thus  $lm(f)$ , the only term in the left-hand side, is also divisible by some  $lt(g_i)$ , as desired.  $\square$

### Example 1.10

Let  $f = x^3y^2 + 2xy + 2x + 2y$ ,  $f_1 = x^2y^2 + 2x$  and  $f_2 = xy + 1$ . Also, let the order be lexicographic with  $x > y$ . We have  $lt(f) = x^3y^2$ ,  $lt(f_1) = x^2y^2$  and  $lm(f_2) = xy$ .

- $lt(f_1)/lt(f) \Rightarrow f = xf_1 + p_1$ , where  $p_1 = -2x^2 + 2xy + 2x + 2y$
- $lt(f_1), lt(f_2)$  don't divide  $lt(p_1)$ , so  $p_1 = -2x^2 + 2xy + 2x + 2y$  is a remainder of  $f$ .

Now we will solve it slightly different:

- $lt(f_2)/lt(f) \Rightarrow f = x^2yf_2 + p_1$ , where  $p_1 = -x^2y + 2xy + 2x + 2y$
- $lt(f_2)/lt(p_1) \Rightarrow f = x^2yf_2 - xf_2 + p_1$ , where  $p_2 = 2xy + 3x + 2y$
- $lt(f_2)/lt(p_2) \Rightarrow f = x^2yf_2 - xf_2 + 2f_2 + p_3$ , where  $p_3 = 3x + 2y - 2$
- $lt(f_2), lt(f_1)$  don't divide  $lt(p_3)$ , so  $p_3 = 3x + 2y - 2$  is another remainder of  $f$ .

We notice that in the division algorithm a remainder of  $f$  is, in general, not unique. However,

### Lemma 1.3 [7]

If  $G = \{g_1, g_2, \dots, g_s\}$  is a Gröbner basis of  $I = (g_1, g_2, \dots, g_s)$ , then for any non-zero polynomial  $f$  of  $R[x_1, x_2, \dots, x_n]$ , there is a *unique remainder* of  $f$  with respect to  $g_1, g_2, \dots, g_s$ .

**Proof.** Suppose there exist remainders  $f'$  and  $f''$  with respect to  $g_1, g_2, \dots, g_s$  with  $f' \neq f''$ . Since  $0 \neq f' - f'' \in I$ , the initial monomial  $w = lt(f' - f'')$  must belong to  $lm(I)$ . However, since  $w \in \text{supp}(f') \cup \text{supp}(f'')$ , it follows that none of the

monomials  $lt(g_1), lt(g_2), \dots, lt(g_t)$  divides  $w$ . Hence  $in(I) \neq (lt(g_1), lt(g_2), \dots, lt(g_t))$  a contradiction.

**Corollary 1.1** [7]

If  $G = \{g_1, g_2, \dots, g_s\}$  is a Gröbner basis of  $I = (g_1, g_2, \dots, g_s)$ , then a non-zero polynomial  $f$  of  $R[x_1, x_2, \dots, x_n]$  belongs to  $I$  if and only if the *unique remainder* of  $f$  with respect to  $g_1, g_2, \dots, g_s$  is 0.

**Proof.** First, in general, if a remainder of a non-zero polynomial  $f$  of  $R[x_1, x_2, \dots, x_n]$  with respect to  $g_1, g_2, \dots, g_s$  is 0, then  $f$  belongs to  $I = (g_1, g_2, \dots, g_s)$ .

Second, suppose that a non-zero polynomial  $f$  belongs to  $I$  and  $f = f_1g_1 + f_2g_2 + \dots + f_sg_s + f'$  is a standard expression of  $f$  with respect to  $g_1, g_2, \dots, g_s$ . Since  $f \in I$ , one has  $f' \in I$ . If  $f' \neq 0$ , then  $lt(f') \in in(I)$ . Since  $G$  is a Gröbner basis of  $I$ , it follows that  $in(I) = (lt(g_1), lt(g_2), \dots, lt(g_s))$ . However, since  $f'$  is a remainder, none of the monomials  $u \in supp(f')$  can belong to  $(lt(g_1), lt(g_2), \dots, lt(g_s))$ .  $\square$

**Definition 1.20** [6]

A Gröbner basis  $G = \{g_1, g_2, \dots, g_s\}$  of an ideal  $I$  is called a *Reduced Gröbner basis* for  $I$  if

- (i).  $lc(g_i) = 1 \forall i$  and
- (ii). none of the terms occurring in  $f_i$  belongs to  $in(G - \{g_i\}) \forall i$

*Reduced Gröbner basis* is very important due to its uniqueness. As a result of its uniqueness, we have the next theorem:

**Theorem 1.3** [6]

A reduced Gröbner basis exists and is *uniquely determined*.

**Proof.** (*Existence*) Let  $I$  be a non-zero ideal of  $R[x_1, x_2, \dots, x_n]$  and  $\{u_1, u_2, \dots, u_s\}$  the unique monomial system of monomial generators of  $in(I)$ . Thus, for  $i \neq j$ , the monomial  $u_i$  cannot be divided by  $u_j$  for each  $1 \leq i \leq s$ , we choose a polynomial  $g_i \in I$  with  $lt(g_i) = u_i$ .

Let  $g_1 = f_2g_2 + \dots + f_sg_s + h_1$  be a standard expression of  $g_1$  with respect to  $g_2, \dots, g_s$ , where  $h_1$  a remainder. It follows from the property (ii) required in the division algorithm that  $lt(g_1)$  coincides with one of the monomials



$lt(f_2), lt(g_2), \dots, lt(f_s), lt(g_s), lt(h_1)$ . Since  $lt(g_1) = u_1$  can be divided by none of the monomials  $lt(g_2), \dots, lt(g_s)$ , one has  $lt(h_1) = lt(g_1)$ . Hence  $\{h_1, g_2, \dots, g_s\}$  is a Gröbner basis of  $I$ . Since the monomial  $h_1$  is a remainder of a standard expression of  $g_1$  with respect to  $g_2, \dots, g_s$ , each monomial of  $supp(h_1)$  is divided by none of the monomials  $lt(g_2), \dots, lt(g_s)$ .

Similarly, if  $h_2$  is a remainder of a standard expression of  $g_2$  with respect to  $h_1, g_3, g_4, \dots, g_s$ , the one has  $lt(h_2) = lt(g_2)$  and each monomial of  $supp(h_2)$  is divided by none of the monomials  $lt(h_1), lt(g_3), \dots, lt(g_s)$ . Moreover,  $\{h_1, h_2, g_3, \dots, g_s\}$  is a Gröbner basis of  $I$ . Since  $lt(h_2) = lt(g_2)$ , each monomial of  $supp(h_1)$  is divided by none of the monomials  $lt(h_2), lt(g_3), \dots, lt(g_s)$ .

Continuing these procedures yields the polynomials  $\{h_3, h_4, \dots, h_s\}$  which satisfies condition (ii). Dividing  $h_i$  by the coefficient of  $lt(h_i)$  for all  $i$ , we obtain a reduced basis of  $I$ .

(Uniqueness) Let  $\{g_1, g_2, \dots, g_s\}$  and  $\{h_1, h_2, \dots, h_t\}$  be reduced Gröbner bases of  $I$ . Since  $\{lt(g_1), lt(g_2), \dots, lt(g_s)\}$  and  $\{lt(h_1), lt(h_2), \dots, lt(h_t)\}$  are the minimal system of monomial generators of the initial ideal  $in(I)$  of  $I$ , we may assume that  $s = t$  and  $lt(g_i) = lt(h_i)$  for all  $1 \leq i \leq s (= t)$ . If  $g_i \neq h_i$ , then  $0 \neq g_i - h_i \in I$  and  $lt(g_i - h_i) < lt(g_i)$ . In particular  $lt(g_i)$  cannot divide  $lt(g_i - h_i)$ . Since the monomial  $lt(g_i - h_i)$  must appear in either  $supp(g_i)$  or  $supp(h_i)$ , it follows that  $lt(g_i - h_i)$  cannot be divided by  $lt(g_j)$  with  $j \neq i$ . Hence,  $lt(g_i - h_i) \notin in(I)$ . This contradicts  $g_i - h_i \in I$ .  $\square$

## 1.5 Buchberger’s algorithm

---

In this chapter, we will present the Buchberger’s algorithm. The Buchberger’s algorithm helps us to find a Gröbner basis, which got its name by Bruno Buchberger. A useful computational definition of Gröbner bases is in terms of *S-polynomials*. Let it be  $L = lcm(lt(f), lt(g))$ , where *lcm* is a least common multiple.

**Definition 1.21** [5]

Let  $0 \neq f, g \in R[x_1, x_2, \dots, x_n]$ . The polynomial

$$S(f, g) = \frac{L}{lm(f)}f - \frac{L}{lm(g)}g \quad (1.15)$$

is called the *S – polynomial* of  $f$  and  $g$ .

**Example 1.11**

A. Let  $G = \{f, g\}$ , where  $f = 2xy^2z - xyz^2$  and  $g = x^2yz - z^2$  with respect to lexicographic order and  $x > y > z$ . We have  $lm(f) = 2xy^2z$ ,  $lt(f) = xy^2z$  and  $lm(g) = lt(g) = x^2yz$ , and therefore  $L = x^2y^2z$ .

$$S(f, g) = \frac{x^2y^2z}{2xy^2z}f - \frac{x^2y^2z}{x^2yz}g = \frac{x}{2}f - yg = \frac{-1}{2}x^2yz^2 + yz^2$$

B. Let  $G = \{f, g\}$ , where  $f = x^2y + 2x + 1$  and  $g = xy + 2y$  with respect to degree lexicographic order and  $x > y > z$ . We have  $lm(f) = x^2y = lt(f)$  and  $lm(g) = lt(g) = xy$ , and therefore  $L = x^2y$ .

$$S(f, g) = \frac{x^2y}{x^2y}f - \frac{x^2y}{xy}g = f - xg = -2xy + 2x + 1$$

**Theorem 1.4** [5]

Let  $G = \{g_1, g_2, \dots, g_s\}$  be a set of non-zero polynomials in  $S = R[x_1, x_2, \dots, x_n]$ . Then  $G$  is a Gröbner basis for the ideal  $I = \langle g_1, g_2, \dots, g_s \rangle$  if and only if for all  $i \neq j$ ,

$$S(g_i, g_j) \xrightarrow{G} 0.$$

In order to prove *Theorem 1.4*, we need the following *lemma* :

**Lemma 1.4** [5]

Let  $g_1, g_2, \dots, g_s \in R[x_1, x_2, \dots, x_n]$  be such that  $lt(g_i) = X \neq 0$  for all  $i = 1, \dots, s$ . Let  $g = \sum_{i=1}^s c_i g_i$  with  $c_i \in F$ ,  $i = 1, \dots, s$ . If  $lt(g) < X$ , then  $g$  is a linear combination, with coefficients in  $F$ , of  $S(g_i, g_j)$ ,  $1 \leq i, j \leq s$ .

**Proof.** Where  $g_i = a_i X + \text{lower term}$ ,  $a_i \in F$ . Then the hypothesis says that  $\sum_{i=1}^s c_i g_i = 0$ , since the  $c_i$ 's are in  $F$ . Now, by Definition,  $S(g_i, g_j) = \frac{1}{a_i} g_i - \frac{1}{a_j} g_j$ , since  $lt(g_i) = lt(g_j) = X$ . Thus

$$\begin{aligned} g &= c_1 g_1 + c_2 g_2 + \dots + c_s g_s \\ &= c_1 a_1 \left(\frac{1}{a_1} g_1\right) + c_2 a_2 \left(\frac{1}{a_2} g_2\right) + \dots + c_s a_s \left(\frac{1}{a_s} g_s\right) \\ &= c_1 a_1 \left(\frac{1}{a_1} g_1 - \frac{1}{a_2} g_2\right) + (c_1 a_1 + c_2 a_2) \left(\frac{1}{a_2} g_2 - \frac{1}{a_3} g_3\right) + \dots + \\ &\quad (c_1 a_1 + \dots + c_{s-1} a_{s-1}) \left(\frac{1}{a_{s-1}} g_{s-1} - \frac{1}{a_s} g_s\right) + (c_1 a_1 + \dots + c_s a_s) \frac{1}{a_s} g_s \\ &= c_1 a_1 S(g_1, g_2) + (c_1 a_1 + c_2 a_2) S(g_2, g_3) + \dots + \\ &\quad (c_1 a_1 + \dots + c_{s-1} a_{s-1}) S(g_{s-1}, g_s), \end{aligned}$$

since  $c_1 a_1 + \dots + c_s a_s = 0$ .  $\square$

Now we can prove *Theorem 1.4*:

**Proof.** If  $G = \{g_1, g_2, \dots, g_s\}$  is a Gröbner basis for  $I = \langle g_1, g_2, \dots, g_s \rangle$ , then

$S(g_i, g_j) \xrightarrow{G} + 0$  for all  $i \neq j$  by *Theorem 1.2*, since  $S(g_i, g_j) \in I$ .

Conversely, let us assume that  $S(g_i, g_j) \xrightarrow{G} + 0$  for all  $i \neq j$ . We will use *Theorem 1.2* (iii) to prove that  $G$  is a Gröbner basis for  $I$ . Let  $f \in I$ . Then  $f$  can be written in many ways as a linear combination of the  $g_i$ 's. We choose to write  $f = \sum_{i=1}^s h_i g_i$ , with

$$X = \max_{1 \leq i \leq s} (lt(h_i) lt(g_i))$$

least (here we use the well-ordering property of the term order). If  $X = lt(f)$ , we are done. Otherwise,  $lt(f) < X$ . We will find a representation of  $f$  with a smaller  $X$ , and this will be a contradiction. Let  $S = \{i \mid lt(h_i) lt(g_i) = X\}$ . For  $i \in S$ , write  $h_i = c_i X_i + \text{lower terms}$ . Set  $g = \sum_{i \in S} c_i X_i g_i$ . Then,  $lt(X_i g_i) = X$ , for all  $i \in S$ , but  $lt(g) < X$ . By *Lemma 1.4*, there exist  $a_{ij} \in F$  such that

$$g = \sum_{i, j \in S, i \neq j} d_{ij} S(X_i g_i, X_j g_j)$$

Now,  $X = lcm(lt(X_i g_i), lt(X_j g_j))$ , so

$$\begin{aligned} S(X_i g_i, X_j g_j) &= \frac{X}{\text{lm}(X_i g_i)} X_i g_i - \frac{X}{\text{lm}(X_j g_j)} X_j g_j \\ &= \frac{X}{\text{lm}(g_i)} g_i - \frac{X}{\text{lm}(g_j)} g_j = \frac{X}{X_{ij}} S(g_i, g_j), \end{aligned}$$

where  $X_{ij} = \text{lcm}(\text{lt}(g_i), \text{lt}(g_j))$ . By hypothesis,  $S(g_i, g_j) \xrightarrow{G} 0$ , and so we see from this last equation that  $S(X_i g_i, X_j g_j) \xrightarrow{G} 0$ . This gives a representation

$$S(X_i g_i, X_j g_j) = \sum_{v=1}^s h_{ijv} g_v,$$

where *Theorem 1.1*

$$\max_{1 \leq v \leq s} (\text{lt}(h_{ijv}) \text{lt}(g_v)) = \text{lt}(S(X_i g_i, X_j g_j)) < \max(\text{lt}(X_i g_i), \text{lt}(X_j g_j)) = X$$

Substituting these expression into  $g$  above, and  $g$  into  $f$ , we get  $f = \sum_{i=1}^s h'_i g_i$ , with

$$\max_{1 \leq i \leq s} (\text{lt}(h'_i), \text{lt}(g_i)) < X.$$

This is contradiction.  $\square$

### Example 1.12

A. According to *Example 1.11A*, we have  $f = 2xy^2z - xyz^2$ ,  $g = x^2yz - z^2$  and  $S(f, g) = \frac{-1}{2}x^2yz^2 + yz^2$ . We have  $\text{lt}(f) = xy^2z$ ,  $\text{lt}(g) = x^2yz$  and  $\text{lt}(S(f, g)) = x^2yz^2$ .

- $\text{lt}(f)$  doesn't divide  $\text{lt}(S(f, g))$ , while
- $\text{lt}(g)/\text{lt}(S(f, g)) \Rightarrow S(f, g) = \frac{-z}{2}g + p_1$ , where  $p_1 = yz^2 - \frac{z^3}{2}$
- $\text{lt}(g), \text{lt}(f)$  don't divide  $\text{lt}(p_1)$ , so  $p_1 = yz^2 - \frac{z^3}{2}$  is the remainder.

As a result,  $G = \{f, g\}$  isn't a Gröbner basis.

B. According to *Example 1.11B*, we have  $f = x^2y + 2x + 1$ ,  $g = xy + 2y$  and  $S(f, g) = -2xy + 2x + 1$ . We have  $\text{lt}(f) = x^2y$ ,  $\text{lt}(g) = xy$  and  $\text{lt}(S(f, g)) = xy$ .

- $lt(f)$  doesn't divide  $lt(S(f, g))$ , while
- $lt(g)/lt(S(f, g)) \Rightarrow S(f, g) = -2g + p_1$ , where  $p_1 = 2x + 4y + 1$
- $lt(g), lt(f)$  don't divide  $lt(p_1)$ , so  $p_1 = 2x + 4y + 1$  is the remainder.

As a result,  $G = \{f, g\}$  isn't a Gröbner basis.

Finally, below is given Buchberger's algorithm for computing Gröbner bases. [5,6,7,8]

$F = \{f_1, f_2, \dots, f_s\} \subseteq R[x_1, x_2, \dots, x_n]$  with  $g_i \neq 0$  ( $1 \leq i \leq s$ ), we find a set  $G = \{g_1, g_2, \dots, g_s\}$ , which is a Gröbner basis for  $I = \langle f_1, f_2, \dots, f_s \rangle$ .

Step1. Let  $G = F$  and  $\mathcal{G} = \{\{f_i, f_j\} \mid f_i \neq f_j \in G\}$

Step2. While  $\mathcal{G} \neq \emptyset$ , we choose any pair of  $\mathcal{G}$  and then update  $\mathcal{G} = \mathcal{G} - \{\{f_i, f_j\}\}$

Step3. Compute  $S(f_i, f_j) \xrightarrow{G}_+ h$ .

Step4. If  $h = 0$ : repeat Step 2-Step 4 until  $\mathcal{G} = \emptyset$ , otherwise update  $\mathcal{G}$  and  $G$  as follows:

$$\mathcal{G} = \mathcal{G} \cup \{\{u, h\} \mid \text{for all } u \in G\}$$

$$G = G \cup \{h\}$$

Step5. Repeat Step 2 – Step 4, until  $\mathcal{G} = \emptyset$ . □

### Example 1.13

A. Let  $I = \langle f_1 = xy - x, f_2 = x^2 - y \rangle$ . We use the lexicographic order with  $x > y$ .

- Let  $G = \{f_1, f_2\}$  and  $\mathcal{G} = \{\{f_1, f_2\}\}$ . We choose the pair  $\{f_1, f_2\}$ .
- Then  $\mathcal{G} = \emptyset$
- We compute  $S(f_1, f_2) = -x^2 + y^2$  and  $S(f_1, f_2) \xrightarrow{G}_+ y^2 - y = f_3$ , so  $G = \{f_1, f_2, f_3\}$  and  $\mathcal{G} = \{\{f_1, f_3\}, \{f_2, f_3\}\}$
- $S(f_1, f_3) = 0$  and  $S(f_2, f_3) = x^2y - y^3$ , we find  $S(f_2, f_3) \xrightarrow{G}_+ 0$ .
- $\mathcal{G} = \emptyset$ , so we stop here.

Eventually,  $G = \{f_1, f_2, f_3\}$  is a Gröbner basis for the ideal  $I$ .

Reduced Gröbner basis of  $G = \{f_1, f_2, f_3\}$  is the same because none of the  $lt(f_i)$  divide any of the others  $lt(f_j)$ , with  $f_i \neq f_j$ .

B. Let  $I = \langle f_1 = y^2 + yx + x^2, f_2 = y + x, f_3 = y \rangle$ . We use the lexicographic order with  $y > x$ .

- Let  $G = \{f_1, f_2, f_3\}$  and  $\mathcal{G} = \{\{f_1, f_2\}, \{f_1, f_3\}, \{f_2, f_3\}\}$
- Then  $\mathcal{G} = \{\{f_1, f_3\}, \{f_2, f_3\}\}$
- We compute  $S(f_1, f_2) = x^2 = f_4$ , so  $G = \{f_1, f_2, f_3, f_4\}$  and  $\mathcal{G} = \{\{f_1, f_3\}, \{f_2, f_3\}, \{f_1, f_4\}, \{f_2, f_4\}, \{f_3, f_4\}\}$
- $\mathcal{G} = \{\{f_2, f_3\}, \{f_1, f_4\}, \{f_2, f_4\}, \{f_3, f_4\}\}$  and  $S(f_1, f_3) \xrightarrow{G} 0$ , then
- $\mathcal{G} = \{\{f_1, f_4\}, \{f_2, f_4\}, \{f_3, f_4\}\}$  and  $S(f_2, f_3) \xrightarrow{G} x = f_5$  and  $G = \{f_1, f_2, f_3, f_4, f_5\}$
- $\mathcal{G} = \{\{f_1, f_4\}, \{f_2, f_4\}, \{f_3, f_4\}, \{f_1, f_5\}, \{f_2, f_5\}, \{f_3, f_5\}, \{f_4, f_5\}\}$  and  $S(f_1, f_4) \xrightarrow{G} 0$  and then,  $\mathcal{G} = \{\{f_2, f_4\}, \{f_3, f_4\}, \{f_1, f_5\}, \{f_2, f_5\}, \{f_3, f_5\}, \{f_4, f_5\}\}$
- All the other  $S(f_2, f_4), S(f_3, f_4), S(f_1, f_5), S(f_2, f_5), S(f_3, f_5), S(f_4, f_5) \xrightarrow{G} 0$  and  $\mathcal{G} = \emptyset$ , so we stop here.

Eventually,  $G = \{f_1, f_2, f_3, f_4, f_5\}$  is a Gröbner basis for the ideal  $I$ .

Reduced Gröbner basis of  $G = \{f_3, f_5\}$ .

## 1.6 Gröbner bases in Mathematica

Generally, computing Gröbner basis by hand is cumbersome, and may be impractical in many occasions. However, with the development of fast computers, Buchberger’s algorithm is now implemented in many Computer Algebra packages e.g Theorema, Maple, Mathematica, Singular and CoCoA. Theorema is developed by Buchberger and his team.

*The **Theorema language** allows to integrate computations (“programming”) along with mathematical theorem proving as another crucial ingredient of theory exploration. As an example, it is coded an operator-based algorithm for computing Green’s operators for linear boundary problems (both ODEs and simple PDEs) directly in a Theorema notebook. Their approach relies on a noncommutative Gröbner basis that describes the relations of the basic analysis operators appearing in boundary problems;*

In Mathematica, we have to use the command :

`GroebnerBasis[{...},{x,y,z,...}]` and then *shift + enter*.

In the first bracket, we put  $f_i$  and in the second bracket, we put the unknown variables.

### Example 1.14

A. Consider the ideal

$$I = \langle x^3y + x + y + 1, y^3 + x^2 + z, z^2y + z^2 \rangle$$

The reduced Gröbner basis for  $I$  with respect to the lex ordering with  $z < y < x$  is computed using Mathematica and is given by:

The screenshot shows a Mathematica notebook window titled "Untitled-4 \*". The input cell contains the command: `In[30]= GroebnerBasis[{y*x^3+x+y+1, z+x^2+y^3, y*z^2+z^2}, {x, y, z}]`. The output cell shows the result: `Out[30]= {-z^3+z^4, z^2+y z^2, 1+2 y+y^2+y^3-2 y^7+y^11+z-4 y^4 z-3 y^5 z-z^2+z^3, 1+y+y^3-y^4+y^5-y^6-y^7+y^8-y^9+y^10+z+x z-y z+y^2 z-2 y^3 z-2 y^4 z+2 y^5 z-2 y^6 z+3 y^7 z+5 z^2-z^3, x+x y-y^3+y^7-z+2 y^4 z-z^2, x^2+y^3+z}`. Below the output, there is a toolbar with buttons for "Groebner basis", "curl", "div", "polynomial reduce", and "more...".

This is the reduced Gröbner basis :

$$f_1 = -z^3 + z^4$$

$$f_2 = z^2 + yz^2$$

$$f_3 = 1 + 2y + y^2 + y^3 - 2y^7 + y^{11} + z - 4y^4z + 3y^8z - z^2 + z^3$$

$$f_4 = 1 + y + y^3 - y^4 + y^5 - y^6 - y^7 + y^8 - y^9 + y^{10} + z + xz - yz + y^2z - 2y^3z - 2y^4z + 2y^5z - 2y^6z + 3y^7z + 5z^2 - z^3$$

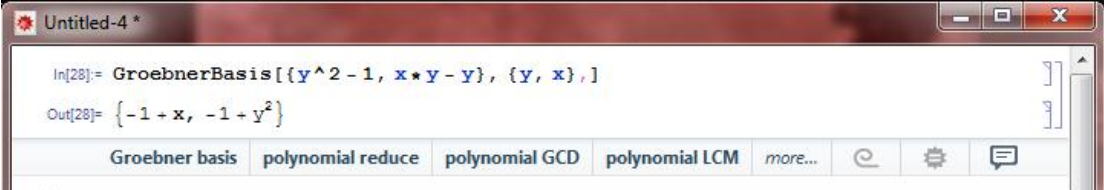
$$f_5 = x + xy - y^3 + y^7 - z + 2y^4z - z^2$$

$$f_6 = x^2 + y^3 + z$$

B. Consider the ideal

$$I = \langle xy - x, y^2 - 1, x - 1 \rangle$$

The reduced Gröbner basis for  $I$  with respect to the lex ordering with  $x > y$  is computed using Mathematica and is given by:



```

In[28]:= GroebnerBasis[{y^2 - 1, x*y - y}, {y, x}, ]
Out[28]:= {-1 + x, -1 + y^2}
    
```

This is the *reduced Gröbner basis* :

$$f = y^2 - 1 \text{ and } g = x - 1$$

C. Consider the ideal

$$I = \langle y^2 + yx + x^2, y + x, y \rangle.$$

The reduced Gröbner basis for  $I$  with respect to the lex ordering with  $y > x$  is computed using Mathematica and is given by:



```
Untitled-2 *  
  
In[4]:= GroebnerBasis[{y^2 + y*x + x^2, y + x, y}, {y, x}]  
  
Out[4]= {x, y}
```

reverse curl div array rules more... ↻ ⚙️ 💬

This is the *reduced Gröbner basis* :

$$f_3 = y \text{ and } f_5 = x$$

## **Chapter 2**

---

# **Equivalence of Polynomial Matrices**

## 2.1 1-D Polynomial Matrices

---

### 2.1.1 Important Notions and Definitions

**Definition 2.1** [3]

A polynomial  $n \times m$  – matrix over the body  $\mathcal{F}$  is a matrix with elements polynomials where they have coefficients of  $\mathcal{F}$ . Such a matrix is denoted

$$P(x) = \begin{pmatrix} p_{11}(x) & \cdots & p_{1m}(x) \\ p_{21}(x) & \cdots & p_{2m}(x) \\ \vdots & \ddots & \vdots \\ p_{n1}(x) & \cdots & p_{nm}(x) \end{pmatrix} \quad (2.1)$$

where  $p_{ij}(x) \in R[x]$ . For a matrix  $A = (a_{ij}) \in M_{n,m}(\mathcal{F})$  and for a natural number  $s$ , we define

$$x^s A = (x^s a_{ij}).$$

We may represent the polynomial matrix  $P(x)$  in the form of a matrix polynomial in  $x$ , i.e., in the form of a polynomial in  $x$  with matrix coefficients:

$$P(x) = P_0 x^l + P_1 x^{l-1} + \cdots + P_{l-1} x + P_l.$$

Now we will see some important properties of polynomial matrices without attempting to generalize these properties to include all matrices with elements in a commutative ring.

The theory of elementary divisors, invented by Sylvester, H.J.S Smith, and, more particularly, Weierstrass, and perfected in important respects by Kronecker, Frobenius, and others.

In particular, we now have:

**Definition 2.2** [12]

The following three elementary row (column) operations on the polynomial matrix  $P(x)$  with coefficients in  $R$  are defined

- (i) Interchange of rows (columns)  $i$  and  $j$ .
- (ii) Multiplication of row (column)  $i$  by a nonzero scalar in  $R$  ( $c \neq 0$ ).
- (iii) Replacement of row (column)  $i$  by itself plus any polynomial multiplied by any other row (column)  $j$  (for example the  $i$  –  $th$  of any other row, for example the  $j$  –  $th$  multiplied by any arbitrary polynomial  $b(x)$ ).

The operations (i), (ii) and (iii) are equivalent to a multiplication of the polynomial matrix  $P(x)$  on the left by the following square matrices of order, respectively [13]:

$$(i) \quad S' = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & & & & \vdots \\ \vdots & & 0 & 1 & & \vdots \\ \vdots & & 1 & 0 & & \vdots \\ \vdots & & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 1 \end{pmatrix}$$

$$(ii) \quad S'' = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & & & \vdots \\ \vdots & & c & & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 1 \end{pmatrix}$$

$$(iii) \quad S''' = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & & & & \vdots \\ \vdots & & 1 & b(x) & & \vdots \\ \vdots & & & \ddots & & \vdots \\ \vdots & & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 1 \end{pmatrix}$$

In the same way, we define the right elementary operations on a polynomial matrix, the matrices (of order  $m$ ) corresponding to them are:

$$(i) \quad T' = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & & & & \vdots \\ \vdots & & 0 & 1 & & \vdots \\ \vdots & & 1 & 0 & & \vdots \\ \vdots & & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 1 \end{pmatrix}$$

$$(ii) \quad T'' = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & & & \vdots \\ \vdots & & c & & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 1 \end{pmatrix}$$

$$(iii) \quad T''' = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & & & & \vdots \\ \vdots & & 1 & & & \vdots \\ \vdots & & b(x) & \ddots & & \vdots \\ \vdots & & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 1 \end{pmatrix}$$

**Definition 2.3** [12]

The *rank*  $\rho$  of a matrix  $P$ , denoted a  $\rho(P)$ , is equal to the maximum number of linearly independent columns (or rows) of  $P$  over the smallest field  $F$  which contains the elements of  $P$ .

**Definition 2.4** [12]

We introduce a concise notation for *determinants* formed from elements of the given matrix:

$$A \begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ k_1 & k_2 & \cdots & k_p \end{pmatrix} = \begin{vmatrix} a_{i_1 k_1} & a_{i_1 k_2} & \cdots & a_{i_1 k_p} \\ a_{i_2 k_1} & a_{i_2 k_2} & \cdots & a_{i_2 k_p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_p k_1} & a_{i_p k_2} & \cdots & a_{i_p k_p} \end{vmatrix} \quad (2.2)$$

**Definition 2.5** [14]

A  $n \times m$  polynomial matrix  $P(x)$  is a matrix with entries that are real coefficient polynomials in  $x$ .

A square ( $n = m$ ) polynomial matrix  $P(x)$  is called *nonsingular* if  $\det P(x)$  is a nonzero polynomial, and *unimodular* if  $\det P(x)$  is a nonzero real number.

Thus an alternative characterization is *nonsingular* if and only if  $\det P(x_0) \neq 0$  for all but a finite number of complex number  $x_0$ . And  $P(x)$  is *unimodular* if and only if  $\det P(x_0) \neq 0$  for all complex numbers  $x_0$ .

Between a matrix and its inverse we have that,  $P^{-1}(x)$  is unimodular if  $P(x)$  is unimodular. Conversely if  $P(x)$  and  $P^{-1}(x)$  both are polynomial matrices, then both are unimodular.

**Example 2.1**

A. We have the matrix

$$A = \begin{pmatrix} 2x & x^2 + x + 1 \\ 2 & x + 1 \end{pmatrix}$$

$\det(A) = -2 \neq 0$ . Since  $\det(A)$  is a nonzero real number, the matrix  $A$  is unimodular.

B. We have the matrix

$$B = \begin{pmatrix} x & x + 1 \\ x - 1 & x \end{pmatrix}$$

$\det(B) = 1 \neq 0$ . Since  $\det(A)$  is a nonzero real number, the matrix  $B$  is unimodular.

We study mostly systems which has rational transfer function. Let give the definition of a rational matrix:

**Definition 2.6**

Let  $R[x]$  be the ring of polynomials with coefficients of  $R$  and  $R(x)$  is the bode of *rational functions* over  $R[x]$

$$R(x) = \{g(x) | g(x) = \frac{n(x)}{d(x)}, n(x), d(x) \in R[x], d(x) \neq 0\} \quad (2.3)$$

$R(x)$  is called the body of rational functions.

*Note:* The elementary row and column operations where we can implement in a rational matrix are the same as in Definition 2.2.

So, we have that every rational transfer function can be expressed as  $g(x) = \frac{n(x)}{d(x)}$ , where  $n(x), d(x)$  are polynomials of  $x$ . The  $g(x)$  can be classified as follows [21]:

- $g(x)$  proper  $\Leftrightarrow \deg d(x) \geq \deg n(s) \Leftrightarrow g(\infty) = \text{zero or nonzero constant}$ .
- $g(x)$  strictly proper  $\Leftrightarrow \deg d(x) > \deg n(s) \Leftrightarrow g(\infty) = 0$ .
- $g(x)$  biproper  $\Leftrightarrow \deg d(x) = \deg n(s) \Leftrightarrow g(\infty) = \text{nonzero constant..}$
- $g(x)$  improper  $\Leftrightarrow \deg d(x) < \deg n(s) \Leftrightarrow |g(\infty)| = \infty$ .

*Note:* Improper rational transfer functions will amplify high-frequency noise, which often exists in the real world; therefore improper rational functions rarely arise in practice.

We introduce the concept of invariant polynomials of a matrix  $A(x)$ . Since in the theory of 2-D matrix we distinguish between two types of invariants; those associated with isolated points of  $\mathbb{C}^2$ .

In 1-D theory, we have the factors and are described by Smith form. In 1-D case the Smith form can be obtained by pre- and post- multiplication by unimodular matrices, but this is impossible in 2-D case. We will see 2-D case later. Now let's give the definition of invariant polynomials.

**Definition 2.7** [19]

The  $i$  – *th* *determinantal divisor*  $d_i(x)$ ,  $i = 1, \dots, r$  of the matrix  $A(x)$  is the greatest common divisor of the  $i$  – *th* order minors of  $A(x)$ . The zeros of  $d_i(x)$  are called the  $i$  – *th* *determinantal zeros* of  $A(x)$ .

*Note:* The set of  $i$  – *th* *order determinantal zeros* of a polynomial matrix  $A(x)$  is denoted  $\eta_i\{A(x)\}$ .

**Definition 2.8** [11]

We define  $d_0(x) \equiv 1$  and we have

$$i_1(x) = \frac{d_0(x)}{d_1(x)}, i_2(x) = \frac{d_1(x)}{d_2(x)}, \dots, i_r(x) = \frac{d_{r-1}(x)}{d_r(x)} \quad (2.4)$$

these are called *invariant polynomials* of the polynomial matrix  $A(x)$ .

*Note :* The characterization “*invariant*” is due to the fact that polynomials  $i_1(x), i_2(x), \dots, i_r(x)$  remain invariant beneath equivalence transformations of  $A(x)$ . This is also evident from (2.4)

**Theorem 2.1** [13]

Let  $A(x) \in R[x]^{n \times m}$  with  $\text{rank}A(x) = r, r = \min \{n, m\}$ . Then  $A(x)$  is equivalent with a diagonal matrix  $S_{A(x)}^C(x) \in R[x]^{n \times n}$  where it has the form:

$$S_{A(x)}^C(x) = \text{diag}[i_1(x), i_2(x), \dots, i_r(x), 0_{m-r, n-r}] \quad (2.5)$$

and it is called *Smith form* of  $A(x)$  in  $\mathbb{C}$  where the polynomials  $i_1(x), i_2(x), \dots, i_r(x)$  are not identically equal to zero and each of the polynomials  $i_2(x), \dots, i_r(x)$  is divisible by the preceding. Moreover, it is assumed that the highest coefficient of all the polynomials  $i_2(x), \dots, i_r(x)$  are equal to 1.

**Proof.** Among all the elements  $a_{ij}(x)$  of  $A(x)$  that are not identically equal to zero we choose one which has the least degree in  $x$  and by suitable permutations of the rows and columns we make this element into  $a_{11}(x)$ . Then we find the quotients and remainders of the polynomials  $a_{i1}(x)$ :

$$a_{i1}(x) = a_{11}(x)q_{i1}(x) + r_{i1}(x), \quad a_{ij}(x) = a_{11}(x)q_{1j}(x) + r_{1j}(x)$$

$$(i = 2, \dots, n, j = 2, \dots, m)$$

If at least one of the remainders  $r_{i1}(x), r_{1j}(x)$  ( $i = 2, \dots, n, j = 2, \dots, m$ ), for example  $r_{1j}(x)$ , is not identically equal to zero, then by subtracting from the  $j - th$  column the first column multiplied by  $q_{1j}(x)$ , we replace  $a_{1j}(x)$  by the remainder  $r_{1j}(x)$ , which is of smaller degree than  $a_{11}(x)$ . Then we can again reduce the degree of the element in the top left corner of the matrix by putting in its place an element of smaller degree in  $r$ .

But if all the remainders  $r_{21}(x), \dots, r_{n1}(x); r_{12}(x), \dots, r_{1m}(x)$  are identically equal to zero, then by subtracting from the  $i - th$  row the first multiplied by  $q_{1j}(x)$ , ( $j = 2, \dots, m$ ), we reduce our polynomial matrix to the form

$$\begin{pmatrix} a_{11}(x) & 0 & \dots & 0 \\ 0 & a_{22}(x) & \dots & a_{2m}(x) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}(x) & \dots & a_{nm}(x) \end{pmatrix} \quad (2.6)$$

If at least one of the elements  $a_{ij}(x)$  ( $i = 2, \dots, n, j = 2, \dots, m$ ) is not divisible without remainder by  $a_{11}(x)$ , then by adding to the first column that column which contains such an element we arrive at the preceding case and can therefore again replace the element  $a_{11}(x)$  by a polynomial of smaller degree.

Since the original element  $a_{11}(x)$  had a definite degree and since the process of reducing this degree cannot be continued indefinitely, we must, after a finite number of elementary operations, obtain a matrix of the form

$$\begin{pmatrix} i_1(x) & 0 & \dots & 0 \\ 0 & b_{22}(x) & \dots & b_{2m}(x) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & b_{n2}(x) & \dots & b_{nm}(x) \end{pmatrix} \quad (2.7)$$



in which all the elements  $b_{ij}(x)$  are divisible without remainder by  $i_1(x)$ . If among these elements  $b_{ij}(x)$  there is one not identically equal to zero, then continuing the same reduction process on the rows numbered 2, ...,  $n$  and the columns 2, ...,  $m$ , we reduce the matrix (2.7) to the form

$$\begin{pmatrix} i_1(x) & 0 & 0 & \dots & 0 \\ 0 & i_2(x) & 0 & \dots & 0 \\ 0 & 0 & c_{33}(x) & \dots & c_{3m}(x) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & c_{n3}(x) & \dots & c_{nm}(x) \end{pmatrix}$$

where  $i_2(x)$  is divisible without remainder by  $i_1(x)$  and all the polynomial  $c_{ij}(x)$  are divisible without remainder by  $i_2(x)$  continuing the process further, we finally arrive at a matrix of the form

$$\begin{pmatrix} i_1(x) & 0 & \dots & 0 & \dots & \dots & 0 \\ 0 & i_2(x) & & 0 & \ddots & \dots & 0 \\ \vdots & 0 & \ddots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & & i_r(x) & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}$$

where the polynomials  $i_1(x), i_2(x), \dots, i_r(x)$  are not identically equal to zero and each is divisible by the preceding one.

By multiplying the first  $r$  rows by suitable nonzero numerical factors, we can arrange that the highest coefficients of the polynomials  $i_1(x), i_2(x), \dots, i_r(x)$  are equal to 1.  $\square$

### Example 2.2

Consider the 1-D matrix

$$A(x) = \begin{pmatrix} x & 1 \\ x & 1 \end{pmatrix}$$

where  $d_0 = d_1 = 1, d_2 = 0$  and  $i_1 = 1, i_2 = 0$ . Hence the Smith form of  $A$  is given by :

$$S_{A(x)}^c(x) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

### **2.1.2 Coprimeness of 1-D polynomial matrices**

Now we will see the notion of Coprimeness, which is one “only” in one dimension. In more than one dimensions, this changes and the one become two or three different notions as we will see in the next chapters

#### **Definition 2.9** [21]

Two polynomials are said to be *coprime* if they have no common factor of degree at least 1.

#### **Definition 2.10** [22]

Let  $R(x), N(x), \widehat{N}(x) \in R[x]^{q \times q}$ ,  $D(x), \widehat{D}(x) \in R[x]^{q \times r}$  be such that

$$N(x) = R(x)\widehat{N}(x) \text{ and } D(x) = R(x)\widehat{D}(x)$$

$R(x)$  is said to be a *common left divisor* of  $N(x)$  and  $D(x)$ .  $R(x)$  is said to be a *greatest common left divisor* (gclid) of  $N(x)$  and  $D(x)$  if every other common left divisor  $L(x)$  of  $N(x)$  and  $D(x)$  is such that  $R(x) = L(x)M(x)$  for some  $M(x) \in R[x]^{q \times q}$ .

It is important to say that, if the gclid of  $N(x)$  and  $D(x)$  are unimodular, then  $N(x)$  and  $D(x)$  are said to be *left coprime*.

*Note:* Respectively, common right divisors and *greatest common right divisor* (gcrd) of  $N(x) \in R[x]^{q \times q}$  and  $D(x) \in R[x]^{r \times q}$  can be studied analogously. So, we have

$$N(x) = \widehat{N}(x)R(x) \text{ and } D(x) = \widehat{D}(x)R(x)$$

and  $R(x)$  is said to be a *common right divisor* of  $N(x)$  and  $D(x)$ . Clearly,  $R(x)$  is a gcrd if every other common right divisor  $L(x)$  of  $N(x)$  and  $D(x)$  is such that  $R(x) = M(x)L(x)$  for some  $(x) \in R[x]^{q \times q}$ .

Moreover,  $N(x)$  and  $D(x)$  are said to be *right coprime* if the gcrd of  $N(x)$  and  $D(x)$  are unimodular.

Computation of greatest common right divisors can be based on capabilities of elementary row operations on a polynomial matrix. To set up this approach we present a preliminary result.

**Theorem 2.2** [14]

Suppose  $P(x)$  is a  $p \times r$  polynomial matrix and  $Q(x)$  is a  $r \times r$  polynomial matrix. If a unimodular  $(p + r) \times (p + r)$  polynomial matrix  $U(x)$  and an  $r \times r$  polynomial matrix  $R(x)$  are such that

$$U(x) \begin{bmatrix} Q(x) \\ P(x) \end{bmatrix} = \begin{bmatrix} R(x) \\ 0 \end{bmatrix} \quad (2.8)$$

the  $R(x)$  is a gcd of  $P(x)$  and  $Q(x)$ .

**Proof.** Partition  $U(x)$  in the form

$$U(x) = \begin{bmatrix} U_{11}(x) & U_{12}(x) \\ U_{21}(x) & U_{22}(x) \end{bmatrix} \quad (2.9)$$

where  $U_{11}(x)$  is  $r \times r$  and  $U_{22}(x)$  is  $p \times p$ . Then the polynomial matrix  $U^{-1}(x)$  can be partitioned similarly as

$$U^{-1}(x) = \begin{bmatrix} U^{-1}_{11}(x) & U^{-1}_{12}(x) \\ U^{-1}_{21}(x) & U^{-1}_{22}(x) \end{bmatrix} \quad (2.10)$$

Using this notation to rewrite (2.12) gives

$$\begin{bmatrix} Q(x) \\ P(x) \end{bmatrix} = \begin{bmatrix} U^{-1}_{11}(x) & U^{-1}_{12}(x) \\ U^{-1}_{21}(x) & U^{-1}_{22}(x) \end{bmatrix} \begin{bmatrix} R(x) \\ 0 \end{bmatrix}$$

That is,

$$Q(x) = U^{-1}_{11}(x)R(x), \quad P(x) = U^{-1}_{21}(x)R(x)$$

Therefore  $R(x)$  is a common right divisor of  $P(x)$  and  $Q(x)$ . But, from (2.8), (2.9),

$$R(x) = U_{11}(x)Q(x) + U_{12}(x)P(x) \quad (2.11)$$

so that if  $R_a(x)$  is another common right divisor of  $P(x)$  and  $Q(x)$ , say

$$Q(x) = \hat{Q}_a(x)R_a(x), \quad P(x) = \hat{P}_a(x)R_a(x)$$

then (2.11) gives

$$R(x) = [U_{11}(x)\hat{Q}_a(x) + U_{12}(x)\hat{P}_a(x)]R_a(x)$$

This show  $R_a(x)$  also is a right divisor of  $R(x)$ , and thus  $R(x)$  is a gcd of  $P(x)$  and  $Q(x)$ .  $\square$

To calculate greatest common right divisors using Theorem 2.2, we consider the three types of elementary row operations which referred in Definition 2.2.

The below Theorem is the same as Theorem 2.2, but for greatest common left divisors (gclid) of  $P(x)$  and  $Q(x)$ .

**Theorem 2.3** [14]

Suppose  $P(x)$  is a  $q \times p$  polynomial matrix and  $Q(x)$  is a  $q \times q$  polynomial matrix. If a unimodular  $(q + p) \times (q + p)$  polynomial matrix  $U(x)$  and an  $q \times q$  polynomial matrix  $L(x)$  are such that

$$[Q(x) \ P(x)]U(x) = [L(x) \ 0] \quad (2.12)$$

the  $L(x)$  is a gclid of  $P(x)$  and  $Q(x)$ .

**Example 2.3**

For

$$Q(x) = \begin{pmatrix} x^2 + x + 1 & x + 1 \\ x^2 - 3 & 2x - 2 \end{pmatrix}$$

$$P(x) = (x + 2 \ 1)$$

calculation of a gclid via Theorem 2.2 is a sequence of elementary row operations.

$$\begin{aligned} M(x) = \begin{bmatrix} Q(x) \\ P(x) \end{bmatrix} &= \begin{bmatrix} x^2 + x + 1 & x + 1 \\ x^2 - 3 & 2x - 2 \\ x + 2 & 1 \end{bmatrix} \xrightarrow{\Gamma 1 \leftrightarrow \Gamma 3} \begin{bmatrix} x + 2 & 1 \\ x^2 - 3 & 2x - 2 \\ x^2 + x + 1 & x + 1 \end{bmatrix} \\ &= \begin{bmatrix} x + 2 & 1 \\ (x - 2)(x + 2) + 1 & 2x - 2 \\ (x - 2)(x + 2) + 3 & x + 1 \end{bmatrix} \\ &\xrightarrow{\begin{matrix} \Gamma 2 \rightarrow \Gamma 2 - (x - 2)\Gamma 1 \\ \Gamma 3 \rightarrow \Gamma 3 - (x - 2)\Gamma 1 \end{matrix}} \begin{bmatrix} x + 2 & 1 \\ 1 & x \\ 3 & 3 \end{bmatrix} \xrightarrow{\Gamma 1 \leftrightarrow \Gamma 2} \begin{bmatrix} 1 & x \\ x + 2 & 1 \\ 3 & 3 \end{bmatrix} \\ &\xrightarrow{\Gamma 2 \rightarrow \Gamma 2 - (x + 2)\Gamma 1} \begin{bmatrix} 1 & x \\ 0 & -(x + 2) \\ 3 & 3 \end{bmatrix} \xrightarrow{\Gamma 3 \rightarrow \frac{1}{3}\Gamma 3} \begin{bmatrix} 1 & x \\ 0 & -(x + 2) \\ 1 & 1 \end{bmatrix} \\ &\xrightarrow{\Gamma 1 \leftrightarrow \Gamma 3} \begin{bmatrix} 1 & 1 \\ 0 & -(x + 2) \\ 0 & x \end{bmatrix} \xrightarrow{\Gamma 2 \rightarrow \Gamma 2 + \Gamma 3} \begin{bmatrix} 1 & 1 \\ 0 & -2 \\ 0 & x \end{bmatrix} \end{aligned}$$

$$\overrightarrow{\Gamma 2 \rightarrow -\frac{1}{2}\Gamma 2} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & x \end{bmatrix} \overrightarrow{\Gamma 3 \rightarrow \Gamma 3 - x\Gamma 2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$\Gamma 1 = \Gamma 1 - \Gamma 2$$

This calculation shows that a gcd is the identity, and  $P(x)$  and  $Q(x)$  are right coprime.

There are two different characterizations of right coprimeness that are used in the sequel.

**Theorem 2.4** [14]

For a  $p \times r$  polynomial matrix  $P(x)$  and a nonsingular  $r \times r$  polynomial matrix  $Q(x)$ , the following statements are equivalent.

- (i) The polynomial matrices  $P(x)$  and  $Q(x)$  are right coprime.
- (ii) There exist an  $r \times p$  polynomial matrix  $X(x)$  and an  $r \times r$  polynomial matrix  $Y(x)$  satisfying the  $x_0$  called Bezout identity

$$X(x)P(x) + Y(x)Q(x) = I_r \tag{2.13}$$

- (iii) For every complex number  $x_0$ ,

$$\text{rank} \begin{bmatrix} Q(x_0) \\ P(x_0) \end{bmatrix} = r \tag{2.14}$$

**Proof.** Beginning a demonstration that each claim implies the next, first we show that (i) implies (ii). If  $P(x)$  and  $Q(x)$  are right coprime, then reduction to row Hermite form as in (2.8) yields polynomial matrices  $U_{11}(x)$  and  $U_{12}(x)$  such that

$$U_{11}(x)Q(x) + U_{12}(x)P(x) = I_r$$

and this has the form of (2.13).

To prove that (ii) implies (iii), write the condition (2.13) in the matrix form

$$\begin{bmatrix} Y(x) & X(x) \end{bmatrix} \begin{bmatrix} Q(x) \\ P(x) \end{bmatrix} = I_r$$

If  $x_0$  is a complex number for which

$$\text{rank} \begin{bmatrix} Q(x_0) \\ P(x_0) \end{bmatrix} < r$$

then we have a rank contradiction.

To show (iii) implies (i), suppose that (2.14) holds and  $R(x)$  is a common right divisor of  $P(x)$  and  $Q(x)$ . Then for some  $p \times r$  polynomial matrix  $\tilde{P}(x)$  and some  $r \times r$  polynomial matrix  $\tilde{Q}(x)$ .

$$\begin{bmatrix} Q(x) \\ P(x) \end{bmatrix} = \begin{bmatrix} \tilde{Q}(x) \\ \tilde{P}(x) \end{bmatrix} R(x) \quad (2.15)$$

If  $\det R(x)$  is a polynomial of degree at least one and  $x_0$  is a root of this polynomial, then  $R(x_0)$  is a complex matrix of less than full rank. Thus we obtain the contradiction

$$\text{rank} \begin{bmatrix} Q(x_0) \\ P(x_0) \end{bmatrix} \leq \text{rank} R(x_0) < r$$

Therefore  $\det R(x)$  is a nonzero constant, that is,  $R(x)$  is unimodular. This proves that  $P(x)$  and  $Q(x)$  are right coprime.  $\square$

Respectively, for the case of left coprimeness, we have the next theorem:

**Theorem 2.5** [14]

For a  $q \times p$  polynomial matrix  $P(x)$  and a nonsingular  $q \times q$  polynomial matrix  $Q(x)$ , the following statements are equivalent.

- (i) The polynomial matrices  $P(x)$  and  $Q(x)$  are left coprime.
- (ii) There exist an  $p \times q$  polynomial matrix  $X(x)$  and an  $q \times q$  polynomial matrix  $Y(x)$  such that

$$P(x)X(x) + Q(x)Y(x) = I_q \quad (2.16)$$

- (iii) For every complex number  $x_0$ ,

$$\text{rank} [Q(x_0) \quad P(x_0)] = q \quad (2.17)$$

**Definition 2.11** [12]

A pair  $\{ P(x), R(x) \}$  ( $\{ P(x), Q(x) \}$ ) of polynomial matrices which has the same number of columns (rows) is said to be relatively right prime if and only if their gcd (gclid) are unimodular matrices.

*Note:* If two polynomial matrices may be relatively right prime but not left prime and vice versa. Above is given an example in which we can see that:

**Example 2.4**

Consider the polynomial matrices

$$P(x) = \begin{pmatrix} x^2 & -1 \\ -x & x^2 \end{pmatrix}$$

and

$$R(x) = \begin{pmatrix} x & -x \\ 0 & 1 \end{pmatrix}$$

To find a gcd  $G_R(x)$  of  $P(x)$  and  $R(x)$  we reduce the composite matrix

$$\begin{bmatrix} P(x) \\ R(x) \end{bmatrix} = \begin{bmatrix} x^2 & -1 \\ -x & x^2 \\ x & -x \\ 0 & 1 \end{bmatrix}$$

to upper right triangular form .

It is clear that by multiplying the last row  $[0 \ 1]$  of composite matrix by the appropriate monomial and adding the resultant expressions to the remaining rows, all other elements in the second column can be zeroed. The first column terms can also be set equal to zero, with the exception of an  $x$ , by employing an analogous procedure. Therefore, it is clear that

$$G_R(x) = \begin{bmatrix} x & 0 \\ 0 & 1 \end{bmatrix}$$

is a gcd of  $P(x)$  and  $R(x)$ .

To find a gcd  $G_L(x)$  of  $P(x)$  and  $R(x)$  we reduce the composite matrix

$[P(x) \ R(x)] = \begin{bmatrix} x^2 & -1 & x & -x \\ -x & x^2 & 0 & 1 \end{bmatrix}$  to lower left triangular form by adding the third and fourth column of the composite matrix, we obtain the column vector  $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

which can be used to zero all other second row entries. The column vector  $\begin{bmatrix} -1 \\ 0 \end{bmatrix}$  is left as the second column of the remaining matrix and can be used to zero the remaining first row entries. It is therefore clear that the two dimensional identity matrix is a gcd of  $P(x)$  and  $R(x)$ .

## 2.2 Equivalences in 1-D polynomial matrices

---

### 2.2.1 Equivalence relation

Initially, it is important for next chapters to understand better the notion of equivalence. In this chapter, we will define when we have an equivalence relation.

#### **Definition 2.12** [10]

Let  $F \in \{\mathbb{R}, \mathbb{C}\}$  and  $F \times F$  the set of ordered pairs of elements in  $F$ . A subset  $S$  of  $F \times F$  is called a relation on  $F$  and we write  $a \sim_S b$  if the pair  $(a, b) \in S$ .

#### **Definition 2.13** [10]

A relation  $S$  is called an equivalence relation if

1.  $a \sim_S a$  for all  $a$  in  $F$  (reflexivity)
2.  $a \sim_S b \Rightarrow b \sim_S a$  (symmetry)
3.  $a \sim_S b$  and  $b \sim_S c \Rightarrow a \sim_S c$  (transitivity)

Whenever  $\sim$  is an equivalence relation on  $F$ , then

$$[a] = \{b \in F | b \sim a\} \quad (2.18)$$

is called the equivalence class of  $a$  with respect to  $\sim$ , and the set  $\{[a] | a \in F\}$  of all equivalence classes is denoted by  $F/\sim$ .

#### **Definition 2.14**

If  $T$  is another set then a function  $f: S \rightarrow T$  is called *invariant element of  $R$*  when

$$a \sim_S b \Rightarrow f(a) = f(b) \quad (2.19)$$

So, the  $f: S \rightarrow T$  is an invariant element of  $R$  if all the elements  $b \in T$  such that  $(a, b) \in R$  have the same image through  $f$ .

and *complete invariant element of  $R$*  when

$$a \sim_S b \Leftrightarrow f(a) = f(b) \quad (2.20)$$



### **2.2.2 Unimodular Equivalence**

#### **Definition 2.15** [13]

Two polynomial matrices  $A(x)$  and  $B(x)$  are called 1)left-equivalent, 2)right equivalent, 3)equivalent if one of them can be obtained from the other by means of 1)left-elementary, 2)right-elementary, 3)left and right elementary operations, respectively.

Let  $B(x)$  be obtained from  $A(x)$  by means of the left-elementary operations corresponding to  $S_1, S_2, \dots, S_p$ . Then

$$B(x) = S_p S_{p-1} \dots S_1 A(x)$$

Denoting the product  $S_p S_{p-1} \dots S_1$  by  $P(x)$ , we write (2.22) in the form

$$B(x) = P(x)A(x) \tag{2.21}$$

where  $P(x)$ , like each of the matrices  $S_1, S_2, \dots, S_p$  has a constant nonzero determinant.

In the case of right equivalence of the polynomial matrices  $A(x)$  and  $B(x)$  we shall have instead of (2.21) the equation

$$B(x) = A(x)Q(x) \tag{2.22}$$

And finally in the case of equivalence we have the equation:

$$B(x) = P(x)A(x)Q(x) \tag{2.23}$$

Knowing relations (2.21), (2.22) and (2.23) we say again Definition 2.13, a bit different:

#### **Definition 2.16** [13]

Two rectangular matrices  $A(x)$  and  $B(x)$  are called 1)left-equivalent, 2)right equivalent, 3)equivalent(unimodular equivalence) if

- 1)  $B(x) = P(x)A(x)$
- 2)  $B(x) = A(x)Q(x)$
- 3)  $B(x) = P(x)A(x)Q(x)$

respectively, where  $P(x)$  and  $Q(x)$  are polynomial square matrices with constant nonzero determinants.

**Definition 2.17** [15]

Two  $n \times m$  polynomial matrices  $P_1(x), P_2(x)$  are said to be *unimodular equivalent* (u.e) if there exist unimodular matrices  $L(x), R(x)$  such that

$$P_2(x) = L(x)P_1(x)R(x) \quad (2.24)$$

**Theorem 2.6**

The relation generated by (2.24) is an equivalence relation.

**Proof.**

a) *Reflexive law:*

Let  $P_1(x) \in R[z]^{n \times m}$ . We have  $I_n P_1(x) I_m = P_1(x)$ , which is valid.

a) *Symmetric law:*

Let  $P_1(x), P_2(x) \in R[x]^{n \times m}$  polynomial matrices which are unimodular equivalent so we have  $L(x) \in R[x]^{n \times n}, R(x) \in R[x]^{m \times m}$  such that  $P_2(x) = L(x)P_1(x)R(x)$ , then we will have  $L^{-1}(x)P_2(x) = P_1(x)R(x) \Leftrightarrow L^{-1}(x)P_2(x)R^{-1}(x) = P_1(x)$ , where  $L^{-1}(x), R^{-1}(x) \in R[x]^{n \times n}$  because they are unimodular.

b) *Transitive law:*

Let  $P_1(x), P_2(x) \in R[x]^{n \times m}$  unimodular equivalent matrices, so we have  $L(x) \in R[x]^{n \times n}, R(x) \in R[x]^{m \times m}$  unimodular matrices such that

$$P_2(x) = L(x)P_1(x)R(x) \quad (2.25)$$

and  $P_2(x), P_3(x) \in R[x]^{n \times m}$  unimodular equivalent matrices, so we have  $L'(x) \in R[x]^{n \times n}, R'(x) \in R[x]^{m \times m}$  unimodular matrices such that

$$P_3(x) = L'(x)P_2(x)R'(x) \quad (2.26)$$

from (2.26), we have

$$L'^{-1}(x)P_3(x) = P_2(x)R'(x) \Leftrightarrow L'^{-1}(x)P_3(x)R'^{-1}(x) = P_2(x) \quad (2.27)$$

$$\begin{aligned} (2.25) &\stackrel{(2.27)}{\Leftrightarrow} L(x)P_1(x)R(x) = L'^{-1}(x)P_3(x)R'^{-1}(x) \\ &\Leftrightarrow P_1(x)R(x) = L^{-1}(x)L'^{-1}(x)P_3(x)R'^{-1}(x) \end{aligned}$$

$$\begin{aligned} \Leftrightarrow P_1(x) &= L^{-1}(x)L'^{-1}(x)P_3(x)R'^{-1}(x)R^{-1}(x) \\ \Leftrightarrow P_1(x) &= [L(x)L'(x)]^{-1}P_3(x)[R'(x)R(x)]^{-1} \end{aligned}$$

where  $[L(x)L'(x)]^{-1}$  and  $[R'(x)R(x)]^{-1} \in R[x]^{n \times n}$  because they are unimodular.

Consequently, unimodular equivalence is an equivalence relation.  $\square$

*Note: Invariant polynomials and the determinantal divisors of a polynomial matrix  $P(x) \in R[z]^{n \times m}$  are defined uniquely, so the Smith form in  $\mathbb{C}$  of a polynomial matrix is unique. Thus, we can define as an algebraic structure of a polynomial matrix in  $\mathbb{C}$  the structure of the Smith form in  $\mathbb{C}$  of this matrix, which its characteristics are zeros of this polynomial matrix.*

It will be proved that the unimodular equivalence in  $\mathbb{C}$  between two polynomial matrices  $P_1(x)$  and  $P_2(x) \in R[z]^{n \times m}$  has the property retain:

1. the Smith form in  $\mathbb{C}$  of  $P_1(x)$  and  $P_2(x)$
2. the invariant polynomials of  $P_1(x)$  and  $P_2(x)$
3. the determinantal divisors of  $P_1(x)$  and  $P_2(x)$ .

**Proof.**

Let  $P_1(x)$  and  $P_2(x)$  be two equivalent polynomial matrices. Then they are obtained from one another by the means of elementary operations. But an easy verification shows immediately that the elementary operations change neither the rank of  $P_2(x)$  nor the polynomials  $D_1(x), D_2(x), \dots, D_r(x)$ . For when we apply to the identity (2.24) the formula that expresses a minor of a product of matrices by the minors of the factors, we obtain for an arbitrary minor of  $P_2(x)$  the expression

$$\begin{aligned} &P_2 \left( \begin{matrix} j_1 & j_2 & \dots & j_p \\ k_1 & k_2 & \dots & k_p \end{matrix}; x \right) \\ &= \sum_{\substack{1 \leq a_1 < a_2 < \dots < a_p \leq m \\ 1 \leq \beta_1 < \beta_2 < \dots < \beta_p \leq m}} L \left( \begin{matrix} j_1 & j_2 & \dots & j_p \\ a_1 & a_2 & \dots & a_p \end{matrix} \right) P_2 \left( \begin{matrix} \alpha_1 & \alpha_2 & \dots & \alpha_p \\ \beta_1 & \beta_2 & \dots & \beta_p \end{matrix}; x \right) R \left( \begin{matrix} \beta_1 & \beta_2 & \dots & \beta_p \\ k_1 & k_2 & \dots & k_p \end{matrix} \right) \\ &\quad (p = 1, 2, \dots, \min(m, n)). \end{aligned}$$

Hence it follows that all minors of order  $r$  or greater of the matrix  $P_2(x)$ , is divisible by  $D_p(x)$  ( $p = 1, 2, \dots, \min(m, n)$ ). But the matrices  $P_1(x)$  and  $P_2(x)$  can exchange roles.

Therefore  $r \leq r^*$  and  $D_p(x)$  is divisible by  $D_p^*(x)$  ( $p = 1, 2, \dots, \min(m, n)$ ).

Hence

$$r \leq r^*, D_1^*(x) = D_1(x), D_2^*(x) = D_2(x), \dots, D_r^*(x) = D_r(x).$$

Since elementary operations do not change the polynomials

$D_1^*(x), D_2^*(x), \dots, D_r^*(x)$ , they also leave the polynomials  $i_1(x), i_2(x), \dots, i_r(x)$  unchanged.

Thus, the polynomials  $i_1(x), i_2(x), \dots, i_r(x)$  remain invariant on transition from one matrix to another equivalent one.

If the polynomial matrix has the canonical diagonal form (Smith form), then it is easy to see that for this matrix

$$D_1(x) = a_1(x), D_2(x) = a_1(x)a_2(x), \dots, D_r(x) = a_1(x)a_2(x), \dots, a_r(x).$$

But then, the diagonal polynomial in canonical diagonal form coincide with the invariant polynomials

$$i_1(x) = a_1(x), i_2(x) = a_2(x), \dots, i_r(x) = a_r(x).$$

Here  $i_1(x), i_2(x), \dots, i_r(x)$  are at the same time the invariant polynomials of the original matrix  $P_2(x)$ , because it is equivalent to canonical diagonal form.

□

### **2.2.3 Equivalence of Binomials**

We consider two square matrices  $A(x)$  and  $B(x)$  of order  $n$  in which all the elements are of degree not higher than 1 in  $x$ . These polynomial matrices may be represented in the form of matrix binomials:

$$A(x) = A_0x + A_1, B(x) = B_0x + B_1 \quad (2.28)$$

We shall assume that these binomials are of degree 1 and regular, i.e. that  $|A_0| \neq 0$ ,  $|B_0| \neq 0$ .

Below is given a criterion for the equivalence of such binomials:

#### **Theorem 2.7** [13]

If two regular binomials of the first degree  $A_0x + A_1$  and  $B_0x + B_1$  are equivalent, then they are strictly equivalent, i.e., in the identity

$$B_0x + B_1 = P(x)(A_0x + A_1)Q(x) \quad (2.29)$$

the matrices  $P(x)$  and  $Q(x)$ - with constant non-zero determinants- can be replaced by constant non-singular matrices  $P$  and  $Q$ :

$$B_0x + B_1 = P(A_0x + A_1)Q \quad (2.30)$$

**Proof.** Since the determinant of  $P(x)$  does not depend on  $x$  and is different from zero, the inverse matrix  $M(x) = P^{-1}(x)$  is also a polynomial matrix. With the help of this matrix we write (2.29) in the form

$$M(x)(B_0x + B_1) = (A_0x + A_1)Q(x) \quad (2.31)$$

Regarding  $M(x)$  and  $Q(x)$  as matrix polynomials, we divide  $M(x)$  on the left by  $A_0x + A_1$  and  $Q(x)$  on the right by  $B_0x + B_1$ :

$$M(x) = (A_0x + A_1)S(x) + M \quad (2.32)$$

$$Q(x) = T(x)(B_0x + B_1) + Q \quad (2.33)$$

here  $M$  and  $Q$  are constant square matrices (independent of  $x$ ) of order  $n$ . We substitute these expressions for  $M(x)$  and  $Q(x)$  in (2.31). After a few small transformations, we obtain

$$(A_0x + A_1)[T(x) - S(x)](B_0x + B_1) = M(B_0x + B_1) - (A_0x + A_1)Q. \quad (2.34)$$

The difference in the brackets must be identically equal to zero; for otherwise the product on the left-hand side of (2.34) would be of degree  $\geq 2$ , while the polynomial on the right-hand side of the equation is of degree not higher than 1. Therefore

$$S(x) = T(x) \quad (2.35)$$

But then we obtain from (2.34):

$$M(B_0x + B_1) = (A_0x + A_1)Q \quad (2.36)$$

We shall now show that  $M$  is a non-singular matrix. For this purpose we divide  $P(x)$  on the left by  $B_0x + B_1$ :

$$P(x) = (B_0x + B_1)U(x) + P \quad (2.37)$$

From (2.31), (2.32) and (2.37) we deduce:

$$\begin{aligned} E &= M(x)P(x) = M(x)(B_0x + B_1)U(x) + M(x)P \\ &= (A_0x + A_1)Q(x)U(x) + (A_0x + A_1)S(x)P + MP \\ &= (A_0x + A_1)[Q(x)U(x) + S(x)P] + MP. \end{aligned} \quad (2.38)$$

Since the last term of this chain of equations must be of degree zero in  $x$  (because it is equal to  $E$ ), the expression in brackets must be identically equal to zero. But then from (2.38)

$$MP = E. \quad (2.39)$$

so that  $|M| \neq 0$  and  $M^{-1} = P$ .

Multiplying both sides of (2.36) on the left by  $P$  we obtain:

$$B_0x + B_1 = P(A_0x + A_1)Q.$$

The fact that  $P$  is non-singular follows from (2.39). That  $P$  and  $Q$  are non-singular also follows directly from (2.30), since this identity implies

$$B_0 = PA_0Q$$

and therefore

$$|P||A_0||Q| = |B_0| \neq 0.$$

This completes the proof of the theorem.  $\square$

### **2.2.4 Extended Unimodular Equivalence**

Firstly, we will see an important lemma, which will help us in the next theorem.

**Lemma 2.1** [18]

Let the partitioned square polynomial matrix

$$\begin{pmatrix} A(x) & B(x) \\ C(x) & D(x) \end{pmatrix}$$

be unimodular, then

- (i)  $A(x), B(x)$  (respectively  $C(x), D(x)$ ) are relatively left prime.
- (ii)  $A(x), C(x)$  (respectively  $B(x), D(x)$ ) are relatively right prime.

**Definition 2.18** [18]

Let  $\mathcal{P}(m, l)$  be the class of  $(r + m) \times (r + l)$  polynomial matrices where  $l$  and  $m$  are fixed integers and  $r$  ranges over all integers which are greater than  $\max(-m, -l)$ .

Let  $P(x), P_1(x) \in \mathcal{P}(m, l)$  and consider the relation generated by

$$M(x)P(x) = P_1(x)N(x) \tag{2.40}$$

where  $P_1$  and  $M$  are relatively left prime and  $P$  and  $N$  are relatively right prime.

**Theorem 2.8** [18]

The relation generated by (2.40) is an equivalence relation.

**Proof.**

- (i) Reflexive law:

In the first place the relation is reflexive, since (2.40) holds for  $P \equiv P_1$  with  $M$  and  $N$  identity matrices of the appropriate sizes.  $M$  and  $N$  thus have the correct relative primeness properties.

- (ii) Symmetric law:

Secondly, for symmetry suppose,  $P$  and  $P_1$  satisfy (2.40), which may be written as

$$(M, P_1) \begin{bmatrix} P \\ -N \end{bmatrix} = 0 \tag{2.41}$$

Since  $M, P$  are relatively left prime there exist polynomial matrices  $X_1$  and  $X_2$  (which are themselves relatively right prime), such that

$$(M, P_1) \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = I_{r_1+m} \quad (2.42)$$

Similarly, there exist relatively left prime polynomial matrices  $\hat{X}_3, \hat{X}_4$  such that

$$(\hat{X}_3, \hat{X}_4) \begin{bmatrix} P \\ -N \end{bmatrix} = I_{r+l} \quad (2.43)$$

Consider now the matrix pair  $X_3, X_4$  defined by

$$(X_3, X_4) = (\hat{X}_3, \hat{X}_4) \left[ I_{r_1+m+r+l} - \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} (M P_1) \right] \quad (2.44)$$

Then

$$\begin{aligned} (X_3, X_4) \begin{bmatrix} P \\ -N \end{bmatrix} &= (\hat{X}_3, \hat{X}_4) \begin{bmatrix} P \\ -N \end{bmatrix} - (\hat{X}_3, \hat{X}_4) \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} (M P_1) \begin{bmatrix} P \\ -N \end{bmatrix} \\ &= I_{r+l}, \text{ from (2.43) and (2.41)} \end{aligned} \quad (2.45)$$

Also,

$$\begin{aligned} (X_3, X_4) \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} &= (\hat{X}_3, \hat{X}_4) \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} - (\hat{X}_3, \hat{X}_4) \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} (M P_1) \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \\ &= 0, \text{ from (2.42)} \end{aligned} \quad (2.46)$$

The relations (2.41), (2.42), (2.45) and (2.46) may be assembled as:

$$\begin{bmatrix} M & P_1 \\ X_3 & X_4 \end{bmatrix} \begin{bmatrix} X_1 & P \\ X_2 & -N \end{bmatrix} = \begin{bmatrix} I_{r_1+m} & 0 \\ 0 & I_{r+l} \end{bmatrix} \quad (2.47)$$

But the two matrices on the left-hand side of (2.47) are both square and of the same dimensions. Consequently the one is the inverse of the other. Also, since both are polynomial, they are both unimodular. By the properties of an inverse

$$\begin{bmatrix} X_1 & P \\ X_2 & -N \end{bmatrix} \begin{bmatrix} M & P_1 \\ X_3 & X_4 \end{bmatrix} = \begin{bmatrix} I_{r_1+m} & 0 \\ 0 & I_{r+l} \end{bmatrix}$$

From the (1,2) block equations we obtain:

$$(X_1, P) \begin{bmatrix} P_1 \\ X_4 \end{bmatrix} = 0 \quad (2.48)$$



Also by Lemma 2.1  $X_1$  and  $X_4$  have the correct co-primeness properties as required by the relation of Lemma 2.1, since

$$\begin{bmatrix} X_1 & P \\ X_2 & -N \end{bmatrix}; \begin{bmatrix} M & P_1 \\ X_3 & X_4 \end{bmatrix}$$

are unimodular. Thus the symmetry of (2.40) is proved.

(iii) Transitive law:

Finally for the transitivity of (2.40) suppose:

$$\left. \begin{array}{l} MP = P_1 N \\ M' P_1 = P_2 N' \end{array} \right\} \quad (2.49)$$

with the usual relative primeness properties holding.

From the first of these equations,

$$M' MP = M' P_1 N$$

and substituting from the second gives

$$M' MP = P_2 N' N \quad (2.50)$$

It must now be shown that the required relative primeness conditions are met.

From the relative primeness properties associated with (3.36), there exist polynomial matrices  $Q_1, Q_2, Q_3, Q_4$  such that

$$MQ_1 + P_1 Q_2 = I_{r_1+m} \quad (2.51)$$

$$M' Q_3 + P_2 Q_4 = I_{r_2+m} \quad (2.52)$$

From (2.51),

$$M' M Q_1 + M' P_1 Q_2 = M'$$

i.e.

$$M' M Q_1 + P_2 N' Q_2 = M', \text{ from (2.49)}$$

Post-multiplying by  $Q_3$  gives:

$$\begin{aligned} M' M Q_1 Q_3 + P_2 N' Q_2 Q_3 &= M' Q_3 \\ &= I_{r_2+m} - P_2 Q_4, \text{ from (2.52)} \end{aligned}$$

$$M' M Q_1 Q_3 + P_2 (N' Q_2 Q_3 - Q_4) = I_{r+m} \quad (2.53)$$

(2.53) thus proves that  $M'M$  and  $P_2$  are relative left prime. Similarly  $P$  and  $N'N$  are relative right prime, and transitivity follows from (2.50). This completes the proof of the theorem.

Note: The equivalence relation generated by equation (2.40) will be called *extended unimodular equivalence* (eue). Note that the definition of eue contains as a special case the usual *unimodular equivalence* ue of polynomial matrices. Thus we have:

**Proposition 2.1** [18]

Two polynomial matrices of the same dimensions that are *ue* are also *eue*.

As a consequence of this result and Theorem 2.8 we have

**Proposition 2.2** [18]

If two polynomial matrices in  $\mathcal{P}(m, l)$  are eue then so are their respective Smith form.

**Proof.** Suppose  $P(x), P_1(x) \in \mathcal{P}(m, l)$  are eue, then

$$MP = P_1N \quad (2.54)$$

with the usual relative primeness conditions holding.

Let  $S_{P(x)}(x)$  and  $S_{P_1(x)}(x)$  denote the respective Smith forms of  $P(x)$  and  $P_1(x)$  respectively, then

$$P = LS_{P(x)}R, \quad P_1 = L_1S_{P_1(x)}R_1 \quad (2.55)$$

for unimodular matrices  $L, L_1, R, R_1$ .

Using (2.55) in (2.54) gives

$$MLS_{P(x)}R = L_1S_{P_1(x)}R_1N$$

i.e.

$$L_1^{-1}MLS_{P(x)} = S_{P_1(x)}R_1NR_1^{-1} \quad (2.56)$$

Now  $L_1^{-1}ML$  and  $S_{P_1(x)}$  are relative left prime since

$$(L_1^{-1}ML, S_{P_1(x)}) = L_1^{-1} \begin{bmatrix} L & 0 \\ 0 & R_1^{-1} \end{bmatrix}$$

with the first and third matrices on the right-hand side being unimodular, and  $M, P_1$  are relative left prime. Similarly,  $S_{P(x)}$  and  $R_1NR_1^{-1}$  are relative right prime, and so (2.56) proves the proposition.  $\square$

## 2.3 2-D Polynomial Matrices

---

In this chapter, we will see what it happens in 2-D polynomial matrices, where we have polynomials with two indeterminates. *Are the notions the same as in one indeterminate? If they aren't, what's different?*

In 1-D theory, we saw that elementary matrices play a crucial role and they are subclass of unimodular matrices over a ring unimodular. Unimodular matrices can be formed as a product of elementary matrices. In this chapter, we will study systems with more than one variables, where not all unimodular matrices can be formed as a product of elementary matrices. This happens due to the absence of a division algorithm in  $R[x_1, x_2, \dots, x_n]$ .

The existence of a division algorithm for Euclidean polynomial rings forms the basis for the algorithmic derivation of many canonical forms and solution techniques at the heart of 1-D polynomial equations. In case of  $n \geq 2$  progress is possible by noting that any polynomial ring can be regarded as a subring of a larger ring with a division algorithm. The exact mechanism is to favor one of the indeterminates and consider elements of the ring to be polynomial in this indeterminate with coefficients rational in the others. If, for example,  $x_n$  is the favored indeterminate the resulting ring is denoted  $R(x_1, x_2, \dots, x_{n-1})[x_n]$ .

Consider the 2-D system matrix in the general form

$$P(x_1, x_2) = \begin{pmatrix} T(x_1, x_2) & U(x_1, x_2) \\ -V(x_1, x_2) & W(x_1, x_2) \end{pmatrix} \quad (2.57)$$

where  $T(x_1, x_2)$ ,  $U(x_1, x_2)$ ,  $V(x_1, x_2)$  and  $W(x_1, x_2)$  are respectively  $r_1 \times r_2$ ,  $r_1 \times l$ ,  $m \times r_2$  and  $m \times l$  polynomial, where  $r_1 < r_2 + l$  and  $r_2 < r_1 + m$ .

Let  $P(x_1, x_2)$  be an  $m \times l$  polynomial matrix, then  $P(x_1, x_2)$  can be written as:

$$P(x_1, x_2) = P_0(x_2) + P_1(x_2)x_1 + \dots + P_q(x_2)x_1^q \quad (2.58)$$

where  $P_i(x_2)$ ,  $i = 0, 1, \dots, q$  are  $m \times l$  polynomial matrices over  $R[x_2]$  and

$$P_q(x_2) \neq 0 \quad (2.59)$$

Obviously, the matrix  $P(x_1, x_2)$  can be expressed in a similar fashion in terms of powers of  $x_2$ , i.e.

$$P(x_1, x_2) = Q_0(x_1) + Q_1(x_1)x_2 + \cdots + Q_p(x_1)x_2^p \quad (2.60)$$

where  $Q_i(x_1)$ ,  $i = 0, 1, \dots, p$  are  $m \times l$  polynomial matrices over  $R[x_1]$  and

$$Q_p(x_1) \neq 0 \quad (2.61)$$

Such as in Definition 2.2 for 1-D case, now we will present 2-D elementary row (respectively column) over the ring of two-variable polynomials, as we said, it is not Euclidean and there is not the direct dependence between the equivalence of two-variable monomial matrices.

**Definition 2.19** [25]

The following three elementary row (column) operations on the polynomial matrix  $P(x_1, x_2)$  with coefficients in  $R[x_1, x_2]$  are defined

- (i) the multiplication of the  $i - th$  row (respectively column) by the scalar  $c \in R$
- (ii) the addition to the  $i - th$  row (respectively column) of the  $j - th$  row (respectively column) multiplied by the polynomial  $b(x_1, x_2)$
- (iii) the interchange of the  $i - th$  and row  $j - th$  (respectively column)

### **2.3.1 Notions of Coprimeness**

In case of 2-D polynomial matrices are two distinct definitions, minor and zero coprimeness and in the case of  $n \geq 3$  we have one more different definition, factor coprimeness. In the case of 2-D polynomial matrices the notions minor and factor coprimeness are identical. In this thesis, we will examine only the case with two indeterminates.

#### **Definition 2.20** [23]

The matrices  $T, U$  in (2.57) are said to be minor left coprime (mlc) in case the  $r \times r$  minors of the compound matrix  $[T \ U]$  have no non-trivial common factors in  $R[x_1, x_2]$ . Similarly, the matrices  $T, V$  in (2.57) are said to be minor right coprime (mrc) in case the  $r \times r$  minors of the compound matrix  $\begin{bmatrix} T^T \\ -V^T \end{bmatrix}$  have no non-trivial common factors in  $R[x_1, x_2]$ .

#### **Lemma 2.2** [23]

The following statements are equivalent

- (i)  $T, U$  are mlc
- (ii) Any polynomial factorization  $[T \ U] = A[T^* \ U^*]$  with  $A$  being a square matrix, implies that  $A$  is unimodular over  $R[x_1, x_2]$ .

#### **Definition 2.21** [23]

The matrices  $T, U$  in (3.1) are said to be zero left coprime (zlc) in case the compound matrix  $[T \ U]$  has rank  $r$  for all values of the indeterminate pair  $(x_1, x_2)$  over  $\mathbb{C}^2$ . Similarly the matrices  $T, V$  in (3.1) are said to be zero right coprime (zrc) in case the compound matrix  $\begin{bmatrix} T^T \\ -V^T \end{bmatrix}$  has rank  $r$  for all values of the indeterminate pair  $(x_1, x_2)$  over  $\mathbb{C}^2$ .

In 1-D system theory a polynomial matrix with rank degeneracies can be viewed as the product of two polynomial matrices, one with full rank and the other containing the rank degeneracies. Consider now a  $n \times m$  2-D polynomial matrix  $A(x)$  with  $n \leq m$ . Then there are three different notions of relative primeness for this matrix. These are termed minor and zero coprimeness and are defined as follows:

**Definition 2.22** [16]

Let  $A(x)$  and  $B(x)$  denote, respectively, an  $m \times q$  and  $m \times l$  polynomial matrix,  $q + 1 \geq m \geq 1$ , and let

$$C(x) = [A(x) \quad B(x)] \tag{2.62}$$

Then, the pair  $A(x), B(x)$  is said to be

1. *zero left coprime* (zlc) if there exists no 2-tuple  $x = (x_1, x_2)$  which is a zero of all the  $m \times m$  minors of  $C(x)$ ,
2. *minor left coprime* (mlc) if these minors are relative prime, and
3. *factor left coprime* (flc) if in any polynomial decomposition  $C(x) = C_1(x)C_2(x)$  in which  $C_1(x)$  is square,  $C_1(x)$  is necessarily elementary.

*Note:* In dual fashion,  $A(x), B(x)$  are *zero right coprime* (zrc), etc if the matrix transposed pair  $A^T(x), B^T(x)$  is *zero left coprime*.

► *Transpose* is a matrix where it has as columns the rows of matrix  $A(x)$  and as rows the columns of  $A(x)$ , for example, we have the matrix

$$A(x) = \begin{pmatrix} a_{11}(x) & a_{12}(x) & \dots & a_{1m}(x) \\ a_{21}(x) & a_{22}(x) & \dots & a_{2m}(x) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(x) & a_{n2}(x) & \dots & a_{nm}(x) \end{pmatrix}$$

and its transpose is the matrix

$$A^T(x) = \begin{pmatrix} a_{11}(x) & a_{21}(x) & \dots & a_{n1}(x) \\ a_{12}(x) & a_{22}(x) & \dots & a_{n2}(x) \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m}(x) & a_{2m}(x) & \dots & a_{nm}(x) \end{pmatrix}$$



**Example 2.5**

This example demonstrates the different types of coprimeness for polynomial matrices over the polynomial ring  $R[x_1, x_2]$ .

- (i) The greatest common divisor of the second order minors of the compound matrix  $[A_1 \quad B_1]$  formed from the polynomial matrices

$$A_1[x_1, x_2] = \begin{bmatrix} x_1 & 0 \\ 0 & x_2 \end{bmatrix}, \quad B_1[x_1, x_2] = \begin{bmatrix} x_2 & 0 \\ 0 & x_1 \end{bmatrix}$$

is equal to 1, we have six different second order minors and these are

$$\begin{vmatrix} x_1 & 0 \\ 0 & x_2 \end{vmatrix}, \begin{vmatrix} x_1 & x_2 \\ 0 & 0 \end{vmatrix}, \begin{vmatrix} x_1 & 0 \\ 0 & x_1 \end{vmatrix}, \begin{vmatrix} 0 & x_2 \\ x_2 & 0 \end{vmatrix}, \begin{vmatrix} 0 & 0 \\ x_2 & x_1 \end{vmatrix}, \begin{vmatrix} x_2 & 0 \\ 0 & x_1 \end{vmatrix}$$

However, for the values  $x_1 = x_2 = x_3 = 0$  the compound matrix loses rank, and therefore  $A_1$  and  $B_1$  are examples of minor left coprime matrices.

(ii) Finally the compound  $[A_2 \ B_2]$  formed from the polynomial matrices

$$A_2[x_1, x_2] = \begin{bmatrix} 1 + x_1x_2^2 & x_1 \\ x_2^2 & 1 \end{bmatrix}, \quad B_2[x_1, x_2] = \begin{bmatrix} x_1^2 & 0 \\ 1 + x_2 & x_1 \end{bmatrix}$$

has the second order minor

$$\begin{vmatrix} 1 + x_1x_2^2 & x_1 \\ x_2^2 & 1 \end{vmatrix} = 1$$

which is equal to 1, and therefore  $A_2$  and  $B_2$  are examples of zero left coprime matrices.

*Note :* Let  $A \equiv A[x_1, x_2]$  denote an  $m \times m$  polynomial matrix in the 2 variables  $x_i$ ,  $i = 1, 2$ , and let  $d(x) = \det A(x) \neq 0$ . Suppose that  $d(x) = d_1(x)d_2(x)$  where  $d_1(x)$  and  $d_2(x)$  are both polynomials. Then, for  $n \geq 3$  it is not always possible to find two polynomial  $m \times m$  matrices  $A_1(x)$  and  $A_2(x)$  such that  $\det A_i(x) = d_i(x)$ ,  $i = 1, 2$  and

$$A(x) = A_1(x)A_2(x) \tag{2.63}$$

**Theorem 2.9** [16]

For  $n = 1$  the three definitions are equivalent, i.e.  $zlc \equiv mlc \equiv flc$ . For  $n = 2$ ,  $zlc \not\equiv mlc \equiv flc$  and for  $n = 3$ ,  $zlc \not\equiv mlc \not\equiv flc$ . Always  $zlc \rightarrow mlc \rightarrow flc$ .

**Proof.** That the three definitions are equivalent for  $n = 1$  is well known [Rosenbrock,1970] and is directly attributable to the fact that every ideal of polynomials in one variable is principal [Vander Waerden, 1950]. Since the polynomials  $A(x) = x_1$  and  $B(x) = x_2$  possess to common zero  $x = (0,0)$  but are ely



prime, the zlc and mlc concepts differ for  $n \geq 2$ . Nevertheless, it is proved in Theorem 3.3 that for  $n = 2$ , a pair  $A(x), B(x)$  is mlc if and only if it is flc.

The proof of  $n = 3$  is in Youla and Gnani [16].

Furthermore, it is obvious that for all  $n \geq 1$ ,  $\text{zlc} \rightarrow \text{mlc} \rightarrow \text{flc}$ .  $\square$

**Theorem 2.10** [16]

1. The  $m \times q$  and  $m \times l$  polynomial matrices  $A(x)$  and  $B(x)$ ,  $q + l \geq m + l$  are zlc if and only if there exist two polynomial matrices  $X(x)$  and  $Y(x)$  such that

$$A(x)X(x) + B(x)Y(x) = I_m \quad (2.64)$$

2. They are mlc if and only if for every  $i = 1, \dots, n$ , there exist polynomial matrices  $X_i(x)$  and  $Y_i(x)$  such that

$$A(x)X_i(x) + B(x)Y_i(x) = \psi_i(x)I_m \quad (2.65)$$

where  $\psi_i(x)$  is a nontrivial scalar polynomial independent of the variable  $x_i$ . Moreover, if  $A(x)$  and  $B(x)$  are real,  $X(x)$ ,  $Y(x)$  and  $X_i(x)$ ,  $Y_i(x)$ ,  $\psi_i(x)$   $i = 1, \dots, n$ , can always be constructed.

**Before the Proof of the Theorem 3.2, we need the Cauchy-Binet theorem[13]:**

Suppose that a square matrix  $C = (c_{ij})$  is the product of two rectangular matrices  $A = (a_{ik})$  and  $B = (b_{kj})$  of dimension  $n \times m$  and  $m \times n$ , respectively:

$$\begin{pmatrix} c_{11} & \dots & c_{1n} \\ c_{21} & \dots & c_{2n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \dots & c_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ a_{21} & \dots & a_{2m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1n} \\ b_{21} & \dots & b_{2n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} \quad (2.66)$$

i.e.,

$$c_{ij} = \sum_{k=1}^m a_{ik} b_{kj} \quad (i, j = 1, \dots, n) \quad (2.67)$$

We shall establish the important Binet-Cauchy formula, which expresses the determinant  $|C|$  in terms of the minors of  $A$  and  $B$ :

$$\begin{pmatrix} c_{11} & \dots & c_{1n} \\ c_{21} & \dots & c_{2n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \dots & c_{nn} \end{pmatrix} = \sum_{1 \leq k_1 < k_2 < \dots < k_n \leq m} \begin{pmatrix} a_{1k_1} & \dots & a_{1k_n} \\ a_{2k_1} & \dots & a_{2k_n} \\ \vdots & \ddots & \vdots \\ a_{nk_1} & \dots & a_{nk_n} \end{pmatrix} \begin{pmatrix} b_{k_1 1} & \dots & b_{k_1 n} \\ b_{k_2 1} & \dots & b_{k_2 n} \\ \vdots & \ddots & \vdots \\ b_{k_n 1} & \dots & b_{k_n n} \end{pmatrix} \quad (2.68)$$

According to this formula the determinant of  $C$  is the sum of the products of all possible minors of the maximal  $(n - th)$  order of  $A$  into the corresponding minors of the same order of  $B$ .

The Binet-Cauchy formula enables us, in the general case also, to express the minors of the product of two rectangular matrices in terms of the minors of the factors. Let

$$A = (a_{ik}), B = (b_{kj}), C = (c_{ij})$$

$$(i = 1, \dots, n, k = 1, \dots, m, j = 1, \dots, q)$$

and  $C = AB$ .

We consider an arbitrary minor of  $C$ :

$$C \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix} \quad (1 \leq i_1 < i_2 < \dots < i_p \leq n \\ \leq j_1 < j_2 < \dots < j_p \leq q ; \\ p \leq n \text{ and } p \leq q)$$

The matrix formed from the elements of this minor is the product of two rectangular matrices

$$\begin{pmatrix} a_{i_1 1} & \dots & a_{i_1 m} \\ a_{i_2 1} & \dots & a_{i_2 m} \\ \vdots & \ddots & \vdots \\ a_{i_p 1} & \dots & a_{i_p m} \end{pmatrix}, \quad \begin{pmatrix} b_{1 j_1} & \dots & b_{1 j_p} \\ b_{2 j_1} & \dots & b_{2 j_p} \\ \vdots & \ddots & \vdots \\ b_{m j_1} & \dots & b_{m j_p} \end{pmatrix}$$

Therefore, by applying the Binet-Cauchy formula, we obtain:

$$C \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix} = \sum_{1 \leq k_1 < k_2 < \dots < k_p \leq m} A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} B \begin{pmatrix} k_1 & k_2 & \dots & k_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix} \quad (2.69)$$

The rank of the product of two rectangular matrices does not exceed the rank of either factor.

*Note:* If  $C = AB$  and  $\rho(A)$ ,  $\rho(B)$ ,  $\rho(C)$  are the ranks of  $A, B, C$  then  

$$\rho(C) \leq \min(\rho(A), \rho(B))$$

**Proof. (Theorem 2.10)**

1. Clearly, (2.64) guarantees that  $rank C = rank[A \ B = m]$  for all  $x$  and this implies that no  $x = (x_1, x_2, \dots, x_n)$  is a common zero of all the  $m \times m$  minors of  $C(x)$ . Thus (2.64) is sufficient for zlc. To prove necessity we employ a novel technique which succeeds in isolating each individual  $m \times m$  minor of  $C(x)$ .

Let the pair  $A(x), B(x)$  be zlc and let  $\Delta_{i_1, i_2, \dots, i_m}(x)$  denote the  $m \times m$  minor of  $C(x)$  formed with the given  $m$  rows and the  $m$  columns numbered  $i_1, i_2, \dots, i_m$ . From the Definition of zlc, these  $C_m^{q+l}$  polynomials are devoid of any common zeros and invoking a classical result due to Hilbert [Vander Waerden, 1950], there exist polynomials  $a_{i_1, i_2, \dots, i_m}(x)$  such that

$$1 = \sum_{(i)} a_{i_1, i_2, \dots, i_m}(x) \Delta_{i_1, i_2, \dots, i_m}(x) \tag{2.70}$$

In addition, the  $\alpha$ 's can all be chosen real if all the  $\Delta$ 's are real.

Pick  $K$  to be any  $(q + l) \times m$  real constant matrix whose  $m \times m$  minors

$K \begin{pmatrix} i_1 & i_2 & \dots & i_m \\ 1 & 2 & \dots & m \end{pmatrix}$  are all nonzero, introduce  $q + l$  extra independent variables,  $\lambda_1, \lambda_2, \dots, \lambda_{q+l}$  and let

$$\Lambda(x) = diag[\lambda_1, \lambda_2, \dots, \lambda_{q+l}] \tag{2.71}$$

The polynomial matrix

$$D(x, \lambda) = C(x)\Lambda(x)K \tag{2.72}$$

is  $m \times m$  and from Cauchy-Binet theorem,

$$\begin{aligned} \Delta(x, \lambda) \equiv det \Delta(x, \lambda) = \\ \sum_{(i)} \lambda_{i_1} \lambda_{i_2} \dots \lambda_{i_\mu} \Delta_{i_1, i_2, \dots, i_m}(x) K \begin{pmatrix} i_1 & i_2 & \dots & i_m \\ 1 & 2 & \dots & m \end{pmatrix} \end{aligned} \tag{2.73}$$

Thus for every one of the  $-tuples (i) = i_1, i_2, \dots, i_m$ ,

$$\Delta_{i_1, i_2, \dots, i_m}(x) K \begin{pmatrix} i_1 & i_2 & \dots & i_m \\ 1 & 2 & \dots & m \end{pmatrix} = \frac{\partial^m \Delta(x, \lambda)}{\partial \lambda_{i_1} \partial \lambda_{i_2} \dots \partial \lambda_{i_m}} \Big|_{(\lambda)=0} \quad (2.74)$$

Let  $D_a(x, \lambda)$  denote the  $m \times m$  polynomial matrix adjugate to  $D(x, \lambda)$ . Since

$$\Delta(x, \lambda) I_m = D(x, \lambda) D_a(x, \lambda) \quad (2.75)$$

multiplication of both sides of (2.73) on the right with  $D_a(x, \lambda)$  yields

$$\Delta(x, \lambda) I_m = C(x) \Lambda(\lambda) K D_a(x, \lambda) \quad (2.76)$$

In view of (2.74), (2.76) permits the identifications

$$\Delta_{i_1, i_2, \dots, i_m}(x) I_m = C(x) Z_{i_1, i_2, \dots, i_m}(x) \quad (2.77)$$

where for all  $(i)$ ,

$$Z_{i_1, i_2, \dots, i_m}(x) = \frac{1}{K \begin{pmatrix} i_1 & i_2 & \dots & i_m \\ 1 & 2 & \dots & m \end{pmatrix}} \frac{\partial^m (A(\lambda) K D_a(x, \lambda))}{\partial \lambda_{i_1} \partial \lambda_{i_2} \dots \partial \lambda_{i_m}} \Big|_{(\lambda)=0} \quad (2.78)$$

is  $(q + l) \times m$  and polynomial. Finally, by combining (2.66) and (2.77) we reach the desired result (2.64),

$$C(x) Z(x) = A(x) X(x) + B(x) Y(x) = I_m \quad (2.80)$$

where

$$Z(x) = \sum_{(i)} a_{i_1, i_2, \dots, i_m}(x) Z_{i_1, i_2, \dots, i_m}(x) \equiv \begin{bmatrix} X(x) \\ Y(x) \end{bmatrix} \quad (2.81)$$

An examination of the above procedure reveals that  $Z(x)$  is always real if  $C(x)$  is real.

2. Suppose that the pair  $A(x), B(x)$  satisfies (2.65) for every  $i = 1, \dots, n$ . Then, by Cauchy-Binet, the gcd of the  $m \times m$  minors of  $C = [A \ B]$  must divide every  $\psi_i(x)$ . Since  $\psi_i(x)$  is nontrivial and independent of  $x_i$ ,  $i = 1, \dots, n$ , this gcd must be a nonzero constant and the  $C_m^{q+1} \Delta$ 's are therefore relative prime. Hence,  $A(x)$  and  $B(x)$  are mlc. The necessity of (2.65) is also easily established with the aid of (2.77).

By definition,  $A(x)$  and  $B(x)$  are mlc if the  $\Delta$ 's form a relatively prime set of polynomials. But then, according to another classical result [Vander Waerden, 1950], for every  $i = 1, \dots, n$  there exist polynomial  $a_{i_1, i_2, \dots, i_m}(x; i)$  such that

$$\sum_{(i)} a_{i_1, i_2, \dots, i_m}(x; i) \Delta_{i_1, i_2, \dots, i_m}(x) = \psi_i(x) \quad (2.82)$$

where  $\psi_i(x)$  is nontrivial and independent of  $x_i$ .

As before, the  $a$ 's can be chosen real if all  $\Delta$ 's are real. Thus combining (2.77) and (2.82),

$$C(x)Z_i(x) = A(x)X_i(x) + B(x)Y_i(x) = \psi_i(x)I_m \quad (2.83)$$

where

$$Z_i(x) = \sum_{(i)} a_{i_1, i_2, \dots, i_m}(x; i) Z_{i_1, i_2, \dots, i_m}(x) \equiv \begin{bmatrix} X_i(x) \\ Y_i(x) \end{bmatrix} \quad (2.84) \square$$

As we said before, for case  $n=2$  notions minor coprimeness and factor coprimeness are identically. Let's see why this happens with the following Theorem:

**Theorem 2.11** [16]

For  $n = 2$ , a polynomial pair  $A(x), B(x)$  is minor left coprime if and only if it is factor left coprime.

**Proof.** Let the pair  $A(x), B(x)$  be mlc but not flc. Then,  $C(x) = [A(x) \ B(x)]$  admits a polynomial decomposition  $C(x) = C_1(x)C_2(x)$  where in  $C_1(x)$  is square and non elementary. Since mlc implies that normal rank  $C(x) = m$ ,  $\det C_1(x)$  is a non constant polynomial which divides all the  $m \times m$  minors of  $C(x)$ , a contradiction.  $\square$

A set of polynomials  $a_i(x_1), 1 \leq i \leq n$ , in one indeterminate are said to be factor coprime provided there is no value  $\bar{x}_1 \in \mathbb{C}$  such that they are not identically zero. If such a value exists then  $x_1 - \bar{x}_1$  is a factor of all the polynomials in the set. In the case of  $n \geq 2$ , this no longer holds and hence it is a necessary to distinguish between zero coprimeness and factor coprimeness. The following fundamental results, termed Hilbert's Nullstellensatz.

**Example 2.6**

Consider the 2-D polynomial matrices given in part (i) of Example 2.5

$$A_1[x_1, x_2] = \begin{bmatrix} x_1 & 0 \\ 0 & x_2 \end{bmatrix}, \quad B_1[x_1, x_2] = \begin{bmatrix} x_2 & 0 \\ 0 & x_1 \end{bmatrix}$$

It can be easily verified that the following identities hold

$$\begin{bmatrix} x_1 & 0 \\ 0 & x_2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & x_2 \end{bmatrix} + \begin{bmatrix} x_2 & 0 \\ 0 & x_1 \end{bmatrix} \begin{bmatrix} x_2 & 0 \\ 0 & 0 \end{bmatrix} = x_2^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{aligned} & \begin{bmatrix} x_1 & 0 \\ 0 & x_3 \end{bmatrix} \begin{bmatrix} x_1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} x_2 & 0 \\ 0 & x_1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & x_1 \end{bmatrix} = x_1^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ & \begin{bmatrix} x_1 & 0 \\ 0 & x_2 \end{bmatrix} \begin{bmatrix} x_2(x_2 + 1) & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} x_2 & 0 \\ 0 & x_1 \end{bmatrix} \begin{bmatrix} x_1(1 + x_2) & 0 \\ 0 & x_1 + x_2 \end{bmatrix} \\ & \quad = x_1 x_2 (1 + x_2) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

which demonstrate the Bezout identities for minor left coprimeness. Now consider the two polynomial matrices given in part (ii) of example 2.5

$$A_2[x_1, x_2] = \begin{bmatrix} 1 + x_1 x_2^2 & x_1 \\ x_2^2 & 1 \end{bmatrix}, \quad B_2[x_1, x_2] = \begin{bmatrix} x_1^2 & 0 \\ 1 + x_2 & x_1 \end{bmatrix}$$

It can be easily verified that the following identities hold

$$\begin{bmatrix} 1 + x_1 x_2^2 & x_1 \\ x_2^2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -x_2^2 & 0 \end{bmatrix} + \begin{bmatrix} x_1^2 & 0 \\ 1 + x_2 & x_1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{x_1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

which demonstrates the Bezout identity for zero left coprimeness.

### 2.3.2 Invariant Polynomials and Zeros

In 2-D matrix theory we distinguish between two types of invariants; those associated with factors of the matrix and those associated with isolated points of  $\mathbb{C}^2$ . The factors have counterparts in 1-D theory and are described by the Smith form. In 1-D the Smith form can be obtained by pre- and post-multiplication by unimodular matrices, but in 2-D this is not possible and we adopt the alternative (and in 1-D equivalent) definition.

**Definition 2.23** [24]

An  $m \times n$  polynomial matrix  $P(x_1, x_2)$  has *Smith form*:

$$\begin{aligned} S(x_1, x_2) &= (Q(x_1, x_2) \quad 0_{m,n-m}), \quad n > m \\ S(x_1, x_2) &= Q(x_1, x_2), \quad n = m \\ S(x_1, x_2) &= \begin{pmatrix} Q(x_1, x_2) \\ 0_{m-n,n} \end{pmatrix}, \quad n < m \end{aligned} \quad (2.85)$$

where  $Q(x_1, x_2)$  is a diagonal matrix having the invariant polynomials  $\varepsilon_i(x_1, x_2)$  as its non-zero elements. If the rank of  $P$  is  $r$  then there are  $r$  non-zero elements occupying the leading  $r$  positions and the remaining invariant polynomials are zero. Each  $\varepsilon_i(x_1, x_2)$  divides  $\varepsilon_{i+1}(x_1, x_2)$ ,  $i = 1, \dots, r$ . The  $\varepsilon_i(x_1, x_2)$  are given by

$$\varepsilon_i(x_1, x_2) = \frac{D_i(x_1, x_2)}{D_{i-1}(x_1, x_2)}, \quad i = 1, \dots, r \quad (2.86)$$

where  $D_0(x_1, x_2) = 1$  and  $D_i(x_1, x_2)$  is the *gcd* of minors of order  $i$  in  $P(x_1, x_2)$ .

Invariant polynomials are unique modulo a multiplicative constant, each of which, in 2-D, correspond to a matrix factor of  $P(x_1, x_2)$ . Thus invariant polynomials can be factored out of the matrix. Invariant polynomials can be decomposed into irreducible factors, which we term invariant factors, each having an associated multiplicity and a number of degrees.

Like we saw, for many reasons it is frequently necessary in systems analysis to transform a polynomial matrix to a simpler but equivalent form. One basic equivalence transformation in the 2-D context is *zero coprime equivalent*. But, before the Definition of *zero coprime equivalent*, firstly, we have to say something else[19]:

There are various zero structures one can define for  $P(x)$ , but all definitions are based on the property that zero is associated with a rank reduction of the matrix. The  $i$  – *th* determinantal divisor is unique module a multiplicative constant  $c \in \mathbb{C}/\{0\}$  and so this definition is the direct extension of the 1-D case where it

characterizes exactly the situation in which a 1-D matrix loses rank. The simple example  $P(x) = (x_1, x_2)$  reveals the inadequacy of this definition for general n-D polynomial matrices. This matrix has  $d_1(x) = 1$  and so has no determinantal zeros, nevertheless  $P(x)$  loses rank for  $x_1 = x_2 = 0$ . We required a more encompassing definition.

For any  $p \times q$  n-D polynomial matrix  $P(x)$ , let  $m_{(i,j)}$  denote an individual  $i \times i$  minor of  $P(x)$  where  $j = 1, \dots, k_i = \binom{p!}{i!(p-i)!} \binom{q!}{i!(q-i)!}$ . Denote the ideal generated by the  $i \times i$  minors of  $P(x)$  by  $I_i^{[P]}$  and write  $I_i^{[P]} = d_i J_i^{[P]}$ , where  $J_i^{[P]}$  is the ideal generated by the set of polynomials which result from the  $i \times i$  minors of  $P(x)$  when the  $i - th$  determinantal divisor  $d_i(x)$  is removed. Clearly each ideal  $J_i^{[P]}$  is generated by a set of factor coprime polynomials. This set, however, may not be additionally zero coprime which is the distinctive feature of n-D ( $n > 1$ ), and the situation which the previous simple example  $P(x) = (x_1, x_2)$  illustrates. These considerations lead us to the following definitions which find their origin in Zerz (1996).

**Definition 2.24** [19]

The  $i - th$  order invariant zeros,  $i = 1, \dots, r$ , of a polynomial matrix  $P(x)$ , are the elements of  $V(I_i^{[P]})$ , the variety defined by the ideal  $I_i^{[P]}$  and they are defined to be

$$Z_i\{P(x)\} = \eta_i\{P(x)\} - \eta_{i-1}\{P(x)\}$$

where  $\eta_i\{P(x)\}$  is the set of  $i - th$  order determinantal zeros of  $P(x)$  and  $\eta_0\{P(x)\}$  is the empty set.



## 2.4 Equivalences in 2-D polynomial matrices

---

In this chapter, we will see four different equivalences in 2-D polynomial matrices and some more results which arise.

### Definition 2.25 [24]

$P_1(x_1, x_2), P_2(x_1, x_2) \in \mathcal{P}(x_1, x_2)^{m \times l}$  are *EO-EQUIVALENT* (eoe) if one can be obtained from the other by a sequence of elementary row and column operations over  $\mathcal{P}(x_1, x_2)$ .

### Definition 2.26 [24]

$P_1(x_1, x_2), P_2(x_1, x_2)$  are *UNIMODULAR EQUIVALENT* (ue) if  $\exists$  unimodular matrices  $L(x_1, x_2), R(x_1, x_2)$  such that

$$P_2(x_1, x_2) = L(x_1, x_2)P_1(x_1, x_2)R(x_1, x_2) \quad (2.87)$$

Elementary operations are the basis for computational developments. If  $E_L(x_1, x_2)$  (resp.  $E_R(x_1, x_2)$ ) is the result of performing the elementary row (resp. column) operations on  $I_m$  (resp.  $I_l$ ) eoe can be written as

$$P_2(x_1, x_2) = E_L(x_1, x_2)P_1(x_1, x_2)E_R(x_1, x_2) \quad (2.88)$$

$E_L(x_1, x_2), E_R(x_1, x_2)$  of (2.88) are called *ELEMENTARY*. It is clear that elementary matrices are unimodular thus eoe implies ue.

*Note:* However unimodular matrices are not necessarily elementary, and so the converse is false.

### Definition 2.27 [19]

Denote the class of  $(s + p) \times (s + q)$  2-D polynomial matrices by  $\mathcal{P}(p, q)$ , where  $s > -\min(p, q)$ ,  $P_1(x_1, x_2), P_2(x_1, x_2) \in \mathcal{P}(p, q)$  are said to be *ZERO COPRIME EQUIVALENT* (zce) in case  $\exists$  polynomial matrices  $M(x_1, x_2), N(x_1, x_2)$  of appropriate dimensions such that

$$M(x_1, x_2)P_2(x_1, x_2) = P_1(x_1, x_2)N(x_1, x_2) \quad (2.89)$$

with  $M, P_1$  zlc and  $P_2, N$  zrc.

**Theorem 2.12** [27]

The relation (2.89) is an equivalence relation.

**Proof.** Let  $P_1(x_1, x_2)^{p \times q}$  and  $P_2(x_1, x_2)^{r \times s}$  two polynomial matrices (with  $p \times q = r \times s$ ) and let  $M(x_1, x_2)$  and  $N(x_1, x_2)$  polynomial matrices as to:

$$MP_2 = P_1N \quad (2.90)$$

with  $M, P_1$  zlc and  $P_2, N$  zrc.

*i. Reflexivity*

Let  $P_1 \equiv P_2$  in (2.90). Then  $p = r$  and  $q = s$ .

If  $N = I_q$  and  $M = I_r$  then  $I_r P_2 = P_1 I_q$  and  $\det I = 1$ .

$I_r, P_1$  zero left coprime and  $P_2, I_q$  zero right coprime.

*ii. Transitivity*

We suppose that

$$MP_2 = P_1N \quad (2.91)$$

$$\bar{M}P_1 = P_3N \quad (2.92)$$

The 2.91  $\Rightarrow M(MP_2) = M(P_1N) = P_3\bar{N}N$ .

So we need to prove that  $\bar{M}, M, P_3$  are zero left coprime. Similarly we will prove that  $\bar{N}, N, P_1$  are zero right coprime.

*iii. Symmetry*

$$\begin{pmatrix} \bar{Y}_1 & \bar{X}_1 \end{pmatrix} \begin{pmatrix} P_2 \\ -N \end{pmatrix} = I_q \quad (2.93)$$

$$\begin{pmatrix} M & P_1 \end{pmatrix} \begin{pmatrix} X_2 \\ Y_2 \end{pmatrix} = I_r \quad (2.94)$$

From (2.90)

$$(M \ P_1) \begin{pmatrix} P_2 \\ -N \end{pmatrix} = 0 \quad (2.95)$$

Then

$$\begin{pmatrix} M & P_1 \\ \bar{Y}_1 & X_1 \end{pmatrix} \begin{pmatrix} X_2 & P_2 \\ Y_2 & -N \end{pmatrix} = \begin{pmatrix} I_r & 0 \\ J & I_l \end{pmatrix} \quad (2.96)$$

where  $J = \bar{Y}_1 X_2 + \bar{X}_1 Y_2$

now if we multiply from the left the (2.96) with

$\begin{pmatrix} I_r & 0 \\ -J & I_l \end{pmatrix}$  we have:

$$\underbrace{\begin{pmatrix} M & P_1 \\ Y_1 & X_1 \end{pmatrix}}_A \underbrace{\begin{pmatrix} X_2 & P_2 \\ Y_2 & -N \end{pmatrix}}_B = \begin{pmatrix} I_r & 0 \\ 0 & I_l \end{pmatrix} \quad (2.97)$$

where  $X_1 = \bar{X}_1 - J P_1$  and  $Y_1 = \bar{Y}_1 - J M$

Because  $A_{(r+q) \times (p+s)}$  and  $B_{(p+s) \times (r+q)}$  square matrices of the same size, which between them are inverse polynomials, they have to be invertible and to give

$$\underbrace{\begin{pmatrix} X_2 & P_2 \\ Y_2 & -N \end{pmatrix}}_B \underbrace{\begin{pmatrix} M & P_1 \\ Y_1 & X_1 \end{pmatrix}}_A = \begin{pmatrix} I_m & 0 \\ 0 & I_s \end{pmatrix} \quad (2.98)$$

and the equations

$$X_2 M + P_2 Y_1 = I_m \quad (2.99)$$

$$X_2 P_1 + P_2 X_1 = 0$$

$$Y_2 P_1 + -N Y_1 = 0$$

So from (2.98)  $X_2, P_2$  are zero left coprime and from (2.99)  $-X_1, P_1$  are zero right coprime.

So it is symmetric.  $\square$

**Theorem 2.13** [24]

The two polynomial matrices  $P_1, P_2$ , of *Definition 2.27* are ue if-f they are zce.

**Proof.** (zce  $\Rightarrow$  ue) The Bezout identities for left and right zero coprimeness are

$$S_1\bar{Z} + P_1\bar{W} = I_{q_1}, \quad XS_2 + YP_2 = I_{q_2} \quad (2.100)$$

Hence we can write

$$\begin{pmatrix} -X & Y \\ P_1 & S_1 \end{pmatrix} \begin{pmatrix} -S_2 & \bar{W} \\ P_2 & \bar{Z} \end{pmatrix} = \begin{pmatrix} I & J \\ 0 & I \end{pmatrix} \quad (2.101)$$

where  $J = -X\bar{W} + Y\bar{Z}$ . Postmultiplying (2.91) by the inverse of the matrix on the right hand side gives

$$\begin{pmatrix} -X & Y \\ P_1 & S_1 \end{pmatrix} \begin{pmatrix} -S_2 & \bar{W} \\ P_2 & \bar{Z} \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \quad (2.102)$$

Hence the matrices on the left hand side are unimodular. Thus

$$\begin{pmatrix} -X & Y \\ P_1 & S_1 \end{pmatrix} \begin{pmatrix} I_{q_1} & 0 \\ 0 & P_2 \end{pmatrix} = \begin{pmatrix} I_{q_2} & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} -X & YP_2 \\ I & S_2 \end{pmatrix} \quad (2.103)$$

in which the first matrix on the left hand side is unimodular. Also

$$\begin{pmatrix} I & X \\ 0 & I \end{pmatrix} \begin{pmatrix} -X & YP_2 \\ I & S_2 \end{pmatrix} \begin{pmatrix} -S_2 & I \\ I & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

and so the last matrix on the right hand side of (2.103) is unimodular. Therefore (2.103) states that  $P_1$  and  $P_2$  are ue.

(ue  $\Rightarrow$  zce) Assume that the following holds

$$\underbrace{\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}}_M \begin{pmatrix} I & 0 \\ 0 & P_1 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & P_2 \end{pmatrix} \underbrace{\begin{pmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{pmatrix}}_N$$

where  $M$  and  $N$  are unimodular. Writing this as

$$\begin{pmatrix} M_{11} & M_{12}P_1 \\ M_{21} & M_{22}P_1 \end{pmatrix} = \begin{pmatrix} N_{11} & N_{12} \\ P_2N_{21} & P_2N_{22} \end{pmatrix}$$

Therefore  $M_{22}P_1 = P_2N_{22}$  and

$$M = \begin{pmatrix} M_{11} & M_{12} \\ P_2 N_{21} & M_{22} \end{pmatrix}, \quad N = \begin{pmatrix} N_{11} & M_{12} P_1 \\ N_{21} & N_{22} \end{pmatrix}$$

Hence  $P_2$  and  $M_{22}$  are zlc otherwise  $M$  is not unimodular. Similarly  $P_1$  and  $N_{22}$  must zrc.  $\square$

**Theorem 2.14** [19]

Suppose that  $P_1(x) \in \mathcal{P}(p, q)$  of rank  $r_i$  and with dimensions  $p_1 - q_1 = p_2 - q_2 (= p - q)$  are Z.C.E. according to the relation (2.89). Then

$$I_{r_1-i}^{[p_1]} = I_{r_2-i}^{[p_2]}, \quad i = 0, \dots, r-1 \quad (2.104)$$

where  $r = \min(r_1 - r_2)$ . For any  $i \geq r$ ,  $I_{r_1-i}^{[p_1]} = \langle 1 \rangle$  in case  $r_1 - i \geq 0$  or  $I_{r_2-i}^{[p_2]} = \langle 1 \rangle$  in case  $r_2 - i \geq 0$ .

**Proof.**

Suppose that  $h_1 = \min(p_1, q_1)$ ,  $h_2 = \min(p_2, q_2)$  and let  $i = \{1, 2\}$  and its complement  $i'$  in  $\{1, 2\}$  be such that  $h_i \leq h_{i'}$ .

Let

$$P_i(z) \triangleq \begin{pmatrix} I_{h_i-h_i} & 0 \\ 0 & P_i(z) \end{pmatrix}$$

It is clear that the ideals generated by the various minors of  $P_i(z), P_{i'}(z)$  are related as

$$\begin{aligned} I_{h_i}^{[P_i]} &= I_{h_i}^{[P_i]} \\ &\vdots \\ I_{h_i-h_{i+1}}^{[P_i]} &= I_1^{[P_i]} \\ I_{h_i-h_i}^{[P_i]} &= \langle 1 \rangle \\ &\vdots \\ I_1^{[P_i]} &= \langle 1 \rangle \end{aligned} \quad (2.105)$$

Now  $P_i(z), P_{i'}(z)$  are Z.C.E. as either of the statements

$$\begin{pmatrix} 0 & I_{p_i} \end{pmatrix} P_i'(z) = P_i'(z) \begin{pmatrix} 0 & I_{q_i} \end{pmatrix}$$

$$\begin{pmatrix} 0 \\ I_{p_i} \end{pmatrix} P_i'(z) = P_i'(z) \begin{pmatrix} 0 \\ I_{q_i} \end{pmatrix}$$

confirms. Hence from the transitivity of Z.C.E. relation it follows that  $P_i'(z), P_i(z)$  are Z.C.E.

We establish the theorem for  $P_i'(z), P_i(z)$  which are of identical dimensions.

Let  $h = \min(p', q')$  then from the coprimeness requirements of Z.C.E.  $\exists$  polynomial matrices  $X(z), Y(z), W(z), Z(z)$  of appropriate dimensions such that

$$\begin{aligned} MX + P_1'Y &= I_{p'} \\ WP_2' + ZN &= I_{q'} \end{aligned} \tag{2.106}$$

From (2.89) and (2.104) it follows that

$$\begin{pmatrix} W & -Z \\ M & P_1' \end{pmatrix} \begin{pmatrix} P_2' & X \\ -N & Y \end{pmatrix} = \begin{pmatrix} I_{q'} & J \\ 0 & I_{p'} \end{pmatrix} \tag{2.107}$$

where  $J = WX - ZY$ .

For any matrix  $Q$  let  $Q_{j_1, \dots, j_k}^{i_1, \dots, i_k}$  denote the  $k \times k$  submatrix formed from rows  $i_1, \dots, i_k$  and columns  $j_1, \dots, j_k$ . Consider then the following equation formed from (2.107)

$$\underbrace{\begin{pmatrix} E^{i_1, \dots, i_k} & 0 \\ M & P_1' \end{pmatrix}}_A \underbrace{\begin{pmatrix} P_2'_{j_1, \dots, j_k} & X \\ -N_{j_1, \dots, j_k} & Y \end{pmatrix}}_B = \begin{pmatrix} P_2'^{i_1, \dots, i_k} & X^{i_1, \dots, i_k} \\ 0 & I_p' \end{pmatrix} \quad (2.108)$$

where  $1 \leq k \leq h$  and  $E^{i_1, \dots, i_k}$  is the matrix whose  $t, s^{th}$  element is 1 if  $s = i_t$  and zero otherwise.

Take determinants of both sides of (2.108), and use the Cauchy-Binet theorem to expand the left hand side. This gives

$$\sum_m \left| A_{m_1, \dots, m_{p+k}}^{1, \dots, p+k} \right| \left| B_{1, \dots, p+k}^{m_1, \dots, m_{p+k}} \right| = \left| P_2'^{i_1, \dots, i_k} \right| \quad (2.109)$$

Now the form of A indicates that any factor of A of the type occurring in the left hand side of (2.109) for which  $\{i_1, \dots, i_k\}$  is not a subset of  $\{m_1, \dots, m_{p+k}\}$  is zero. Thus all minors of A which occur in the left hand side of (2.109) contain the columns  $\{i_1, \dots, i_k\}$ . Such a factor is then expressible via Laplace expansion in terms of products of minors of  $M \ \kappa \alpha \ \alpha \ P_1'$ . Thus  $\left| P_2'^{i_1, \dots, i_k} \right|$  is expressible as a linear combination of minors  $P_1'$  of order  $k$  and greater.

Since any minor can be expanded in terms of lower order minors, it follows that  $\left| P_2'^{i_1, \dots, i_k} \right|$  can be written as a linear combination of the order  $k$  minors of  $P_1'$ . It thus follows that  $I_k^{[P_2']} \subset I_k^{[P_1']}$ ,  $k = 1, \dots, h$

whre  $h = \min(p', q')$ .

By the symmetry property of the Z.C.E. relation  $\exists p' \times q', q' \times p'$  polynomial matrices  $M'(z), N'(z)$  such that

$$M' P_1' = P_2' N'$$

where  $M'(z), P_2'(z)$  are zero left coprime, and  $P_1'(z), N'(z)$  are zero right coprime.

Applying the same procedure as above gives

$$I_k^{[P_2]} \subset I_k^{[P_1]}$$

where  $k = 1, \dots, h$ . Hence

$$I_k^{[P_2]} = I_k^{[P_1]} \quad \forall k = 1, \dots, h.$$

Let  $i = \{1, 2\}$  and its complement  $i'$  in  $\{1, 2\}$  be such that  $h_i \leq h_{i'}$ . It follows from the above and the relation (2.105) that in terms of the original matrices  $P_1(x), P_2(x)$  we have

$$\begin{aligned} I_{h_i}^{[P_i]} &= I_{h_i}^{[P_i]} \\ &\vdots \\ I_{h_i - h_{i+1}}^{[P_i]} &= I_1^{[P_i]} \\ I_{h_i - h_i}^{[P_i]} &= \langle 1 \rangle \\ &\vdots \\ I_1^{[P_i]} &= \langle 1 \rangle \end{aligned} \tag{2.110}$$

Now since  $r_{i'} = \text{rank} P_{i'}$  we have

$$I_{h_i}^{[P_{i'}]} = \dots = I_{h_i - h_{i'} + r_{i'} + 1}^{[P_{i'}]} = \{0\} \neq I_{h_i - h_{i'} + r_{i'}}^{[P_{i'}]}.$$

Hence

$$r_{i'} = h_i - h_{i'} + r_{i'}$$

and so the relation (2.110) reduce to



$$\begin{aligned}
 I_{r_i}^{[P_i]} &= I_{r_i}^{[P_i]} \\
 &\vdots \\
 I_{r_i-r_{i+1}}^{[P_i]} &= I_1^{[P_i]} \\
 I_{r_i-r_i}^{[P_i]} &= \langle 1 \rangle \\
 &\vdots \\
 I_1^{[P_i]} &= \langle 1 \rangle
 \end{aligned}$$

which completes the proof. □

**Corollary 2.2** [24]

Suppose that two polynomial matrices  $P_1, P_2$  are related by *Z.C.E.* then the invariant zeros are related by

$$Z_{q-n}\{P_1\} = Z_{p-n}\{P_2\} \tag{2.111}$$

where  $p = \min(p_2, q_2), q = \min(p_1, q_1), 0 \leq n \leq \min(p, q)$ .

**Proof.** [28]

Since  $N(x), P_2(x)$  are zero right coprime and  $M(x), P_1(x)$  are zero left coprime there exist polynomial matrices  $X(x), Y(x), W(x)$  and  $Z(x)$  of appropriate dimensions such that

$$MX + P_1Y = I_{p_1} \tag{2.112}$$

$$WP_2 + ZN = I_{q_2}$$

From (2.89) and (2.105) it follows that

$$\begin{pmatrix} W & -Z \\ M & P_1 \end{pmatrix} \begin{pmatrix} P_2 & X \\ -N & Y \end{pmatrix} = \begin{pmatrix} I_{q_2} & J \\ 0 & I_{p_1} \end{pmatrix}$$

where  $J = WX - ZY$ . Now, replace  $[W \quad -Z]$  with  $[E \quad 0]$  which gives

$$\begin{pmatrix} E & 0 \\ M & P_1 \end{pmatrix} \begin{pmatrix} P_2 & X \\ -N & Y \end{pmatrix} = \begin{pmatrix} P_2^{l_1, \dots, l_{q_2}} & X^{l_1, \dots, l_{q_2}} \\ 0 & I_{p_1} \end{pmatrix}. \tag{2.113}$$

where, in the case  $p_2 \geq q_2$  (and hence  $p_1 \geq q_1$ ) the constant matrix  $E_{q_2 \times p_2}$  is the unit matrix  $I_{q_2}$  with  $p_2 - q_2$  zero columns to form a  $q_2 \times p_2$  matrix  $l_1, \dots, l_{q_2}$  correspond to the rows of the matrices  $P_2$  and  $X$  selected by multiplication by  $E$ .

For any matrix  $R$  let  $R_{j_1, \dots, j_h}^{i_1, \dots, i_k}$  denote the  $k \times h$  submatrix formed from rows  $i_1, \dots, i_k$  and columns  $j_1, \dots, j_h$ . Consider the following  $(p_1 + k) \times (p_1 + k)$  submatrix formed from (2.113)

$$\begin{pmatrix} E^{i_1, \dots, i_k} & 0_{k \times q_1} \\ M & P_1 \end{pmatrix} \begin{pmatrix} P_2_{j_1, \dots, j_k} & X \\ -N_{j_1, \dots, j_k} & Y \end{pmatrix} = \begin{pmatrix} (P_2^{l_1, \dots, l_{q_2}})_{j_1, \dots, j_k}^{i_1, \dots, i_k} & X^{l_1, \dots, l_{q_2}}_{j_1, \dots, j_k} \\ 0 & I_{p_1} \end{pmatrix}. \quad (2.114)$$

Take determinants of both sides and using the Cauchy-Binet Theorem, (2.114) shows that

$$\sum_m \left| A_{m_1, \dots, m_{p_1+k}}^{1, \dots, p_1+k} \right| \left| B_{1, \dots, p_1+k}^{m_1, \dots, m_{p_1+k}} \right| = \left| P_2^{l_1, \dots, l_{q_2}}_{j_1, \dots, j_k}^{i_1, \dots, i_k} \right|.$$

Noting that  $l_1, \dots, l_{q_2}$  are arbitrary, and by considering all combinations of the columns of  $E^{i_1, \dots, i_k}$  and the form of  $A$ , it can be seen that the only non-zero minors are those involving columns  $i_1, \dots, i_k$  of the first block column. Such a factor can then be expressed via the Laplace expansion in terms of minors of  $M$  and  $P_1$ . The smallest minor of  $P_1$  occurring in the Laplace expansion is of order  $q + k - p$ , where  $q = \min(p_1, q_1)$  and  $p = \min(p_2, q_2)$ . Therefore the minors of  $P_2$  of order  $k$  are linear combinations of the minors of  $P_1$  of order  $q + k - p$ . Hence the determinantal divisors of  $P_1(x)$  and  $P_2(x)$  are related by the equation

$$\eta_{q+k-n}\{P_1\} \subseteq \eta_n\{P_2\}$$

and letting  $k = p - i$  gives

$$\eta_{q-i}\{P_1\} \subseteq \eta_p\{P_2\}$$

In an analogous manner it is possible to write

$$\begin{pmatrix} W & -Z \\ M & P_1 \end{pmatrix} \begin{pmatrix} P_2 & 0 \\ -N & E \end{pmatrix} = \begin{pmatrix} I_{q_2} & -Z_{l_1, \dots, l_{p_1}} \\ 0 & I_{p_1 l_1, \dots, l_{p_1}} \end{pmatrix}$$

and therefore it can be deduced that the determinantal divisors of  $P_1(x)$  and  $P_2(x)$  are also related by the equation

$$\eta_k\{P_2\} \subseteq \eta_{q+k-p}\{P_1\}$$

Again letting  $k = p - i$  gives

$$\eta_{p-i}\{P_2\} \subseteq \eta_{q-i}\{P_1\}$$

Therefore the following is true

$$\eta_{p-i}\{P_2\} = \eta_{q-i}\{P_1\} \quad (2.115)$$

Now since the determinantal divisors are equal this implies that the invariant zeros must also be equal, giving

$$Z_{q-n}\{P_1\} = Z_{p-n}\{P_2\} \quad (2.116)$$

which is the required result. □

*Note:* So we see that invariant polynomials and invariant zeros are invariants of Z.C.E.

**Definition 2.28** [26]

$P_1(x_1, x_2), P_2(x_1, x_2) \in \mathcal{P}(m, l)$  are said to be FACTOR COPRIME EQUIVALENT (F.C.E.) if there exists polynomial matrices  $M(x_1, x_2), N(x_1, x_2)$  such that

$$[M(x_1, x_2) \quad P_1(x_1, x_2)] \begin{bmatrix} P_2(x_1, x_2) \\ -N(x_1, x_2) \end{bmatrix} = 0 \quad (2.117)$$

where the compound matrices

$$[M(x_1, x_2) \quad P_1(x_1, x_2)] \quad ; \quad \begin{bmatrix} P_2(x_1, x_2) \\ -N(x_1, x_2) \end{bmatrix} \quad (2.118)$$

are factor coprime i.e. if all the  $(r + m) \times (r + m)$  (resp.  $(r + l) \times (r + l)$ ) minors of  $[M(x_1, x_2) \quad P_1(x_1, x_2)]$  (resp.  $\begin{bmatrix} P_2(x_1, x_2) \\ -N(x_1, x_2) \end{bmatrix}$ ) have no polynomial factor.

**Corollary 2.3** [27]

Suppose that two polynomial matrices  $P_1(x_1, x_2), P_2(x_1, x_2)$  with sizes  $p_1 \times q_1$  and  $p_2 \times q_2$  respectively and  $p_1 - q_1 = p_2 - q_2$ , are related by a polynomial equation of the form

$$N_1(x_1, x_2)P_2(x_1, x_2) = N_2(x_1, x_2)P_1(x_1, x_2) \quad (2.119)$$

where  $N_1(x_1, x_2), N_2(x_1, x_2)$  are  $p_1 \times p_2, q_1 \times q_2$  polynomial matrices and  $N_1(x_1, x_2), P_1(x_1, x_2)$  are minor right coprime,  $N_2(x_1, x_2), P_2(x_1, x_2)$  are minor left coprime.

- (i). Let  $d_1^{[1]}(x_1, x_2), d_2^{[1]}(x_1, x_2), \dots, d_q^{[1]}(x_1, x_2)$ , where  $q = \min(p_1, q_1)$ , denote the invariant polynomials of the polynomial matrix  $P_1(x_1, x_2)$  and  $d_1^{[2]}(x_1, x_2), d_2^{[2]}(x_1, x_2), \dots, d_p^{[2]}(x_1, x_2)$ , where  $p = \min(p_2, q_2)$  denote the invariant polynomials of the polynomial matrix  $P_2(x_1, x_2)$  then

$$d_{q-i}^{[1]} = c_i d_{p-i}^{[2]} \quad \text{for } i = 0, 1, \dots, \max(p-1, q-1)$$

where  $d_j^{[1]} = 1, d_j^{[2]} = 1$  for  $j < 1, c_i \in \mathbb{R} \setminus \{0\}$ .

- (ii). Let  $e_1^{[1]}(x_1, x_2), e_2^{[1]}(x_1, x_2), \dots, e_r^{[1]}(x_1, x_2)$ , where  $r = \min(p_1, p_2)$ , denote the invariant polynomials of the polynomial matrix  $S_1(x_1, x_2)$  and  $e_1^{[2]}(x_1, x_2), e_2^{[2]}(x_1, x_2), \dots, e_t^{[2]}(x_1, x_2)$ , where  $t = \min(q_1, q_2)$  denote the invariant polynomials of the polynomial matrix  $S_2(x_1, x_2)$  then

$$e_{r-i}^{[1]} = c_i e_{t-i}^{[2]} \quad \text{for } i = 0, 1, \dots, \max(r-1, t-1)$$

where  $e_j^{[1]} = 1, e_j^{[2]} = 1$  for  $j < 1, c_i \in \mathbb{R} \setminus \{0\}$ .

**Proof.**

Since  $N_1(x_1, x_2), P_1(x_1, x_2)$  are minor right coprime and  $N_2(x_1, x_2), P_2(x_1, x_2)$  are minor left coprime there exist polynomial matrices  $X_i(x_1, x_2), Y_i(x_1, x_2),$

$W_i(x_1, x_2), Z_i(x_1, x_2)$  for  $i=1,2$  of appropriate dimensions such that

$$X_1 P_1 + Y_1 N_1 = \psi_i(x_1) I_q \quad \text{for } i=1,2 \quad (2.120)$$

$$N_2 Z_1 + P_2 W_1 = \varphi_i(x_1) I_p$$

where  $\psi_i(x_1), \varphi_i(x_1)$  are polynomials. From (2.119) and (2.120) it follows that

$$\begin{bmatrix} X_1 & Y_1 \\ N_2 & -P_2 \end{bmatrix} \begin{bmatrix} P_1 & Z_1 \\ N_1 & -W_1 \end{bmatrix} = \begin{bmatrix} \psi_i(x_1) I_q & J_1 \\ 0 & \varphi_i(x_1) I_p \end{bmatrix} \quad (2.121)$$

where  $J_1 = N_2 Z_1 + Y_1 W_1$ . Now take  $i=1$  and replace  $\begin{bmatrix} Z_2 \\ -W_1 \end{bmatrix}$  with  $\begin{bmatrix} 0 \\ J \end{bmatrix}$ . Then (2.121) gives

$$\begin{bmatrix} X_1 & Y_1 \\ N_2 & -P_2 \end{bmatrix} \begin{bmatrix} P_1 & 0 \\ N_1 & I_p \end{bmatrix} = \begin{bmatrix} \psi_i(x_1) I_q & Y_1 \\ 0 & -P_2 \end{bmatrix} \quad (2.122)$$

For any matrix  $Q$  let  $Q_{j_1, \dots, j_h}^{i_1, \dots, i_k}$  denoted the  $k \times h$  submatrix formed from rows  $i_1, \dots, i_k$  and columns  $j_1, \dots, j_h$ . Consider the following  $(q+k) \times (q+k)$  submatrix formed from (2.122)

$$\underbrace{\begin{bmatrix} X_1 & Y_1 \\ N_1^{i_1, \dots, i_k} & -P_2^{i_1, \dots, i_k} \end{bmatrix}}_A \underbrace{\begin{bmatrix} P_1 & 0_{q \times k} \\ N_1 & I_{p_1 j_1, \dots, j_h} \end{bmatrix}}_B = \begin{bmatrix} \psi_1 I_q & Y_1^{j_1, \dots, j_h} \\ 0_{k \times q} & -P_2^{i_1, \dots, i_k} \end{bmatrix} \quad (2.123)$$

Take determinants of both sides and use the Cauchy-Binet Theorem and (2.123) to show that

$$\sum_{1 \leq i_1 < \dots < i_{q+k}} \left| A_{l_1, \dots, l_{q+k}}^{1, \dots, q+k} \right| \left| B_{l_1, \dots, l_{q+k}}^{1, \dots, l_{q+k}} \right| = -\psi_1^q \left| P_2^{i_1, \dots, i_k} \right|. \quad (2.124)$$

Now the form of B indicates that any factor of B of the type occurring in the left-hand-side of (2.124), for which  $\{q + j_1, \dots, q + j_k\}$  is not a subset of  $\{l_1, \dots, l_{q+k}\}$  is zero.

Thus all the non-zero minors of B which occur in the left-hand-side of (2.124) contain the rows  $q + j_1, \dots, q + j_k$ . Such a factor is then expressible via Laplace expansion in terms of minors of  $N_1$  and  $P_1$ . The smallest minor of  $P_1$  occurring in this Laplace expansion is of order  $q + k - p$ . Therefore if  $g_i^{[1]}(x_1, x_2)$  for  $i = 1, \dots, q$  denotes the greatest common divisor of the  $i \times i$  minors  $P_1$ , it follows that

$$g_{q+k-p}^{[1]} \left| \psi_1^q \left| P_2^{i_1, \dots, i_k} \right| \right| \quad (2.125)$$

where  $g_{q+k-p}^{[1]} = 1$  if  $q + k - p \leq 0$ .

If then  $g_i^{[2]}(x_1, x_2)$  for  $i = 1, \dots, p$  denotes the greatest common divisor of the  $i \times i$  minors  $P_2$ , it follows from (2.125) and the fact  $i_1, \dots, i_k$  and  $j_1, \dots, j_k$  are arbitrary that

$$g_{q+k-p}^{[1]} \left| \psi_1^q g_k^{[2]} \right|. \quad (2.126)$$

On the other hand if we take  $i = 2$  then the same argument shows that

$$g_{q+k-p}^{[1]} \left| \psi_2^q g_k^{[2]} \right| \quad k = 1, \dots, p \quad (2.127)$$

Statements (2.126) and (2.127) then imply since  $\psi_1, \psi_2$  are factor coprime

$$g_{q+k-p}^{[1]} \left| g_k^{[2]} \right| \quad k = 1, \dots, p$$

or, on writing  $k = p - j$ ,

$$g_{q-j}^{[1]} \left| g_{p-j}^{[2]} \right| \quad j = 0, \dots, \max(p - 1, q - 1) \quad (2.128)$$

where, if necessary,  $g_j^{[1]} = 1, g_j^{[2]} = 1$  for  $j < 1$ .

Now in (2.121) replace  $[X_1 \ Y_1]$  with  $[I_q \ 0]$  to give

$$\begin{bmatrix} I_q & 0 \\ N_2 & -P_2 \end{bmatrix} \begin{bmatrix} P_1 & Z_1 \\ N_1 & -W_1 \end{bmatrix} = \begin{bmatrix} P_1 & Z_1 \\ 0 & \varphi_1 I_p \end{bmatrix}. \quad (2.129)$$

The same argument surrounding (2.122) may now be used in the case of (2.122) to show that

$$g_{p-j}^{[2]} | g_{q-j}^{[1]} \quad j = 0, \dots, \max(p-1, q-1). \quad (2.130)$$

Statements (2.128) and (2.130) then yield, modulo a constant non-zero factor,

$$g_{q-j}^{[1]}(x_1, x_2) = g_{p-j}^{[2]}(x_1, x_2), \quad j = 0, \dots, \max(p-1, q-1) \quad (2.131)$$

Now  $g_h^{[1]}(x_1, x_2)$ ,  $g_h^{[2]}(x_1, x_2)$  are the determinantal divisors of  $D_1(x_1, x_2)$ ,  $D_2(x_1, x_2)$  respectively, and so from the relationship between the determinantal divisors and their invariant polynomials the result (i) follows.

In the case of  $N_1(x_1, x_2)$  and  $N_2(x_1, x_2)$  the argument presented above will carry through with some minor modifications. Specifically, in the case  $p \leq q$  for example, the equation corresponding to (2.121), for  $i = 1, 2$ , is

$$\begin{bmatrix} X_1 & Y_1 \\ N_2 & -P_2 \end{bmatrix} \begin{bmatrix} P_1 & E \\ N_1 & 0 \end{bmatrix} = \begin{bmatrix} \psi_1 I_q & X_1^{l_1, \dots, l_p} \\ 0 & l_1, \dots, l_p \end{bmatrix} \quad (2.132)$$

where the constant matrix  $E_{q \times p}$  is the unit matrix  $I_p$  with  $q - p$  zero rows to form a  $q \times p$  matrix and  $l_1, \dots, l_p$  correspond to the columns of the matrices  $N_2$  and  $X_1$  selected by multiplication by  $E$ . The analogue of (2.122) is obtained by selecting rows  $i_1, \dots, i_k$  from the second block row and columns  $j_1, \dots, j_k$  from the second block column. By considering all combinations of the rows of  $E_{j_1, \dots, j_k}$  and the form of the second matrix on the right-hand-side of (2.125), i.e. the only non-zero minors are those involving rows  $j_1, \dots, j_k$  of the first block row, it is seen by taking all  $1 \leq i_1 < \dots < i_k \leq p$  and  $1 \leq j_1 < \dots < j_k \leq q$  that

$$h_k^{[1]} | \psi_1^q h_k^{[2]}$$

where  $h_k^{[1]} h_k^{[2]}$  are the greatest common divisors of the  $k \times k$  order minors of  $N_1(x_1, x_2)$ ,  $N_2(x_1, x_2)$  respectively. Also by considering  $i = 2$

$$h_k^{[1]} | \psi_2^q h_k^{[2]}.$$

Therefore, by similar reasoning surrounding (2.128)

$$h_k^{[1]} | h_k^{[2]} \quad \text{for } k = 1, \dots, p.$$

Now the equation corresponding to (2.129) is

$$\begin{bmatrix} 0 & E \\ N_2 & -P_2 \end{bmatrix} \begin{bmatrix} P_1 & Z_1 \\ N_1 & -W_1 \end{bmatrix} = \begin{bmatrix} N_1^{l_1, \dots, l_p} & W_1^{l_1, \dots, l_p} \\ 0 & \varphi_1 l_p \end{bmatrix}$$

where  $E$  and  $l_1, \dots, l_p$  are defined in (2.132). Thus the result

$$h_k^{[2]} | h_k^{[1]} \text{ for } k = 1, \dots, p$$

is obtained by a similar discussion to that following (2.125). Therefore

$$h_k^{[1]} = c_k h_k^{[2]} \text{ for } k = 1, \dots, p$$

where  $c_k \in R \setminus \{0\}$  for  $k = 1, \dots, p$  and (ii) is established.

□

*Note:* So we see that invariant polynomials are invariants of M.C.E, since M.C.E.  $\equiv$  F.C.E. in 2-D polynomial matrices, the invariant polynomials are also invariants of F.C.E.

## References

---

- [1] J. B. Fraleigh, “A first course in Abstract Algebra”, University of Rhode Island, 1989
- [2] E. Ψωμόπουλος, “Αλγεβρικές Δομές II” Θεσσαλονίκη, Δεκέμβριος 2009
- [3] Θ. Θεοχάρη- Αποστολίδη κ.α., “Εισαγωγή στη Γραμμική Άλγεβρα”, Θεσσαλονίκη, 2006
- [4] T. Becker, Volker Weispfenning, “Gröbner Base: A computational Approach to Commutative Algebra”, USA, 1993
- [5] W. W. Adams, Philippe Loustau, “An introduction to Gröbner bases”, USA, 1994
- [6] R. H. Vilarreal, “Monomial Algebras”, New York, 2001
- [7] J. Herzog – Takayki Hibi, “Monomial Ideals”, New York, 2011
- [8] B.Buchberger – F.Winker, “Gröbner Bases and Applications”, Cambridge University Press, 1998
- [9] R. Froberg, “An introduction to Gröbner Bases”, England, 1997
- [10] Wikipedia : monomial
- [11] Π.-Χ.Γ. Βασιλείου, Γ. Τσακλίδης, “Εφαρμοσμένη Θεωρία Πινάκων”, Θεσσαλονίκη, 2003
- [12] W.A. Wolovich, “Linear Multivariable Systems”, Springer – Verlag New York, 1974
- [13] F.R. Gantmacher, “The Theory of Matrices”, Vols 1 and 2 Chelsea Publishing Co. New York, 1974
- [14] W. J. Rugh, “Linear System Theory”, Prentice hall New Jersey, 1996
- [15] A.C. Pugh, P. Fretwell, G.E. Hayton, “Some transformations of Matrix Equivalence arising from linear systems theory”, American Control Conference, 1983
- [16] D.C. Youla, G. Gnani, “Notes on n-dimensional systems.”, IEEE Transactions on Circuits and Systems, 26, 105-111, 1979
- [17] P.S. Johnson ET. AL., “A polynomial matrix theory for a certain class of two-Dimensional Linear Systems.”, Eisevier Science Inc., 1996
- [18] A.C. Pugh, A.K. Shelton, “On a new definition of strict system equivalence”, International Journal of Control, 27,657-672, 1978
- [19] A.C. Pugh, S.J. McInerney, El-Nabrawy, “Zero structures of n-D systems”, International Journal of Control, 78, 277-285, 2005
- [20] E. Rogers ET. AL., “Control Systems Theory and Applications for Linear Repetitive Processes”, Springer, 2007
- [21] C. Chen, “Linear System Theory and Design”, Oxford University Press, 1999
- [22] H. Blomberg, R. Ylinen, “Algebraic Theory for Multivariable Linear Systems”, Academic Press, 1983



- [23] A.C. Pugh ET. AL. , “A transformation for 2-D linear systems and a generalization of a theorem of Rosenbrock”, *International Journal of Control*, 71, 491-503, 1998
- [24] A.C. Pugh ET. AL. , “A Transformation for 2-D Systems and its Invariants”, 35<sup>th</sup> Conference in Japan, 1996
- [25] K. Galkowski, “Elementary Operations and Equivalence of two-dimensional system”, 63:6, 1129-1148, Institute of Telecommunication and Acoustic, Poland 2007
- [26] N.P. Karampetakis, S. Vologiannidis, “Infinite elementary divisor structure-preserving transformations for polynomial matrices”, *International Journal of applied Mathematics and Computer Science*, August 2003
- [27] D.S. Johnson, 1993, “Coprime-ness in Multidimensional System Theory and Symbolic Computation”, P.h.D. Thesis, Loughborough University of Technology, U.K.
- [28] S.J. McInerney, “Representations and Transformations for Multi-dimensional Systems”, P.h.D Thesis, Doctor of Philosophy of Loughborough University, August 1999