

Gröbner Bases and Their Applications

M.Sc. Thesis

by

Vikas Bharti



DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY INDORE

May 2026

Gröbner Bases and Their Applications

A THESIS

Submitted in partial fulfillment of the requirements for the award of
the degree of

Master of Science

by

Vikas Bharti

(Roll No. 2403141017)

Under the guidance of

Dr. Dibyendu Mondal and Dr. Sumit Chandra Mishra



**DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY INDORE**

May 2026

INDIAN INSTITUTE OF TECHNOLOGY INDORE
CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled “**Gröbner Bases and Their Applications**” in the partial fulfillment of the requirements for the award of the degree of **MASTER OF SCIENCE** and submitted in the **DEPARTMENT OF MATHEMATICS, INDIAN INSTITUTE OF TECHNOLOGY INDORE**, is an authentic record of my own work carried out during the time period from July 2025 to May 2026 under the supervision of **Dr. Dibyendu Mondal** and **Dr. Sumit Chandra Mishra**, Assistant Professors, Department of Mathematics, IIT Indore. The matter presented in this thesis by me has not been submitted for the award of any other degree of this or any other institute.

Vikas
21/05/2026

Signature of the student with date

(Vikas Bharti)

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Dibyendu Mondal

21/05/2026

Sumit Chandra Mishra
21/05/2026

Signatures of Thesis Supervisors with date

Dr. Dibyendu Mondal and Dr. Sumit Chandra Mishra

Vikas Bharti has successfully defended his M.Sc. thesis in the Oral Examination held on 13th May, 2026.

Dibyendu Mondal Sumit Chandra Mishra

Signature of Supervisors of M.Sc. Thesis

Date: 21/05/2026

Rash...

Signature of Convener, DPGC

Date: 21-5-2026

Dedicated to my
Father

Acknowledgements

I would like to express my sincere gratitude to my thesis supervisors, **Dr. Dibyendu Mondal** and **Dr. Sumit Chandra Mishra**, for their constant guidance, encouragement, and valuable suggestions throughout the course of this work. Their insightful discussions and support have been instrumental in the completion of this thesis.

I sincerely thank the members of my committee, **Prof. Md. Aquil Khan**, **Prof. Bapan Ghosh**, **Dr. Bibekananda Maji**, and **Dr. Charitha Cherugondi**, for their valuable suggestions, constructive feedback, and continuous academic support throughout the course of this work. Their insightful comments and careful evaluation have greatly contributed to improving the quality of this work.

I would also like to thank the all faculty members of the **Department of Mathematics, Indian Institute of Technology Indore**, for providing a stimulating academic environment and continuous support during my studies.

Finally, I express my heartfelt gratitude to my friends and family for their unwavering love, encouragement, and support throughout this journey. Their constant motivation has been a strong source of strength and inspiration.

Vikas Bharti

ABSTRACT

This thesis develops the fundamental theory of computational algebraic geometry with emphasis on polynomial ideals and Gröbner bases. The preliminary part of the work is based on standard material from *Ideals, Varieties, and Algorithms* [1], and includes the study of polynomial ideals, affine varieties, monomial orderings, the division algorithm, Dickson's Lemma, and the Hilbert Basis Theorem, which establish the algebraic and algorithmic foundations of the subject.

The central theme of the thesis is Gröbner bases and their computational properties. We present the theory of Gröbner bases along with key algorithmic tools such as Buchberger's algorithm and elimination theory. The role of Gröbner bases in solving systems of polynomial equations is highlighted through applications of the Hilbert Nullstellensatz and methods for computing elimination ideals and minimal polynomials.

In addition, the thesis explores polynomial mappings and their algebraic structure using Gröbner basis techniques, illustrating how algebraic information can be extracted computationally from ideals.

Finally, applications in combinatorial algebra are discussed, particularly in graph theory. Algebraic encodings of the k -coloring problem are used to study 3-colorable and general k -colorable graphs through associated polynomial ideals. These results demonstrate how Gröbner basis methods provide a powerful bridge between algebraic geometry and combinatorial problems.

Overall, the thesis emphasizes Gröbner bases as an effective computational framework for translating geometric and combinatorial problems into algebraic form.

1 Preliminaries	1
1.1 Monomials and Polynomials	1
1.2 Affine Space and Affine Varieties	3
1.3 Ideals	8
2 Introduction to Gröbner Bases	13
2.1 Ordering on the Monomials in $k[x_1, \dots, x_n]$	13
2.2 A Division Algorithm in $k[x_1, \dots, x_n]$	17
2.3 Monomial Ideals and Dickson's Lemma	19
2.4 The Hilbert Basis Theorem and Gröbner Bases	22
2.5 Properties of Gröbner Bases	24
3 Hilbert's Nullstellensatz, Elimination and the Eigenvector Theorems	33
3.1 Hilbert's Nullstellensatz	33
3.2 Elimination Theory	39
3.3 Eigenvalue Methods for Solving Polynomial Systems	45
3.3.1 Number of Solutions	47
3.3.2 The Eigenvector Theorems	49
4 Minimal Polynomials of Elements in Field Extensions	54
4.1 Polynomial Maps	54
4.2 The Minimal Polynomials of Elements	58
5 Coloring of Graphs	68
5.1 Graphs and Graph Coloring	68
5.2 3-Colorability of Graphs	70
5.3 k-Colorability of Graphs	75
Bibliography	82

Chapter 1

Preliminaries

In this chapter, we introduce the basic algebraic and geometric concepts that will be used throughout this thesis. We review polynomial rings, affine varieties, ideals, and the correspondence between algebraic sets and ideals. These notions form the foundational language of algebraic geometry and provide the necessary background for the study of Gröbner bases and their applications in later chapters. Most of the results and definitions in this chapter are standard; for a detailed treatment, see Chapter 1 of [1].

1.1 Monomials and Polynomials

Before defining monomials and polynomials, we introduce the ambient ring in which these objects lie.

Definition 1.1.1. Let k be a field and let $x_1, x_2, \dots, x_n$ be indeterminates over k . The set of all polynomials in the variables $x_1, x_2, \dots, x_n$ with coefficients in k is called the *polynomial ring in n variables over k* and is denoted by

$$k[x_1, x_2, \dots, x_n].$$

Definition 1.1.2. A *monomial* in $x_1, \dots, x_n$ is a product of the form

$$x_1^{\alpha_1} \cdots x_n^{\alpha_n},$$

where all of the exponents $\alpha_1, \dots, \alpha_n$ are nonnegative integers. The *total degree* of this monomial is the sum $\alpha_1 + \cdots + \alpha_n$.

We can simplify the notation for monomials as follows: let $\alpha = (\alpha_1, \dots, \alpha_n)$ be an n -tuple of nonnegative integers. Then we set

$$x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}.$$

When $\alpha = (0, \dots, 0)$, note that $x^\alpha = 1$. We also let $|\alpha| = \alpha_1 + \cdots + \alpha_n$ denote the total degree of the monomial x^α .

Definition 1.1.3. A *polynomial* in $x_1, \dots, x_n$ with coefficients in a field k is a finite linear combination (with coefficients in k) of monomials. We will write a polynomial f in the form

$$f = \sum_{\alpha} a_{\alpha} x^{\alpha}, \quad a_{\alpha} \in k,$$

where the sum is over a finite number of n -tuples $\alpha = (\alpha_1, \dots, \alpha_n)$.

When dealing with polynomials in a small number of variables, we will usually dispense with subscripts. Thus, polynomials in one, two, and three variables lie in $k[x]$, $k[x, y]$, and $k[x, y, z]$, respectively.

Example 1.1.4. For example,

$$f = 7x^4y^2 + y^3z^3 - \frac{3}{4}xyz + 5y^2$$

is a polynomial in $\mathbb{Q}[x, y, z]$.

Definition 1.1.5. Let

$$f = \sum_{\alpha} a_{\alpha} x^{\alpha}$$

be a polynomial in $k[x_1, \dots, x_n]$.

- (i) We call a_{α} the *coefficient* of the monomial x^{α} .
- (ii) If $a_{\alpha} \neq 0$, then we call $a_{\alpha}x^{\alpha}$ a *term* of f .
- (iii) The *total degree* of $f \neq 0$, denoted $\deg(f)$, is the maximum $|\alpha|$ such that the coefficient a_{α} is nonzero. The total degree of the zero polynomial is undefined.

Example 1.1.6. Consider the polynomial

$$f = 7x^4y^2 + y^3z^3 - \frac{3}{4}xyz + 5y^2.$$

in $\mathbb{Q}[x, y, z]$. Then:

- (i) The coefficients of the monomials are

$$7, \quad 1, \quad -\frac{3}{4}, \quad 5.$$

- (ii) The terms of f are

$$7x^4y^2, \quad y^3z^3, \quad -\frac{3}{4}xyz, \quad 5y^2.$$

- (iii) The total degrees of the monomials are

$$\deg(x^4y^2) = 6, \quad \deg(y^3z^3) = 6, \quad \deg(xyz) = 3, \quad \deg(y^2) = 2.$$

Hence,

$$\deg(f) = 6.$$

1.2 Affine Space and Affine Varieties

Definition 1.2.1. Given a field k and a positive integer n , we define the n -dimensional affine space over k to be the set

$$k^n = \{(a_1, \dots, a_n) \mid a_1, \dots, a_n \in k\}.$$

For an example of affine space, consider the case $k = \mathbb{R}$. Here we get the familiar space $\mathbb{R}^n$ from calculus and linear algebra. In general, we call $k^1 = k$ the *affine line* and k^2 the *affine plane*.

Let us next see how polynomials relate to affine space. The key idea is that a polynomial

$$f = \sum_{\alpha} a_{\alpha} x^{\alpha} \in k[x_1, \dots, x_n]$$

gives rise to a function

$$f : k^n \longrightarrow k$$

defined as follows: given $(a_1, \dots, a_n) \in k^n$, replace every x_i by a_i in the expression for f . Since all of the coefficients also lie in k , this operation gives an element

$$f(a_1, \dots, a_n) \in k.$$

The ability to regard a polynomial as a function is what makes it possible to link algebra and geometry.

A polynomial possesses both an algebraic and a functional interpretation. Consequently, the statement $f = 0$ may mean either that f is the zero polynomial (that is, all coefficients of f are zero) or that f defines the zero function on k^n . In general, these two notions are not equivalent. For instance, over the finite field $\mathbb{F}_2$, the nonzero polynomial $x^2 - x$ vanishes at every point of $\mathbb{F}_2$. However, as long as k is infinite, there is no problem.

Proposition 1.2.2 ([1, Proposition 5, Section 1, p. 3]). *Let k be an infinite field, and let $f \in k[x_1, \dots, x_n]$. Then $f = 0$ in $k[x_1, \dots, x_n]$ if and only if $f : k^n \rightarrow k$ is the zero function.*

Proof. Since the zero polynomial clearly defines the zero function, one implication is immediate. For the converse, suppose that $f(a_1, \dots, a_n) = 0$ for all $(a_1, \dots, a_n) \in k^n$. We prove that f is the zero polynomial by induction on n .

If $n = 1$, then $f \in k[x]$ is a polynomial in one variable. Since a nonzero polynomial of degree d has at most d roots, and k is infinite, the assumption that $f(a) = 0$ for all $a \in k$ implies that f must be the zero polynomial.

Now assume the result holds for polynomials in $n-1$ variables, and let $f \in k[x_1, \dots, x_n]$ vanish at every point of k^n . Viewing f as a polynomial in x_n with coefficients in $k[x_1, \dots, x_{n-1}]$, we may write

$$f = \sum_{i=0}^N g_i(x_1, \dots, x_{n-1}) x_n^i,$$

where $g_i \in k[x_1, \dots, x_{n-1}]$.

Fix $(a_1, \dots, a_{n-1}) \in k^{n-1}$. Then the polynomial $f(a_1, \dots, a_{n-1}, x_n) \in k[x_n]$ vanishes for every $x_n \in k$. By the case $n = 1$, it follows that $f(a_1, \dots, a_{n-1}, x_n) = 0$ as a polynomial in $k[x_n]$. Hence all of its coefficients vanish, so $g_i(a_1, \dots, a_{n-1}) = 0$ for all i .

Since $(a_1, \dots, a_{n-1})$ was arbitrary, each g_i defines the zero function on k^{n-1} . By the induction hypothesis, each g_i is the zero polynomial. Therefore $f = 0$ in $k[x_1, \dots, x_n]$. $\square$

Definition 1.2.3. Let k be a field, and let $f_1, \dots, f_s$ be polynomials in $k[x_1, \dots, x_n]$. Then we set

$$V(f_1, \dots, f_s) = \{(a_1, \dots, a_n) \in k^n \mid f_i(a_1, \dots, a_n) = 0, \text{ for all } 1 \leq i \leq s\}.$$

We call $V(f_1, \dots, f_s)$ the *affine variety* defined by $f_1, \dots, f_s$.

Thus, an affine variety $V(f_1, \dots, f_s) \subseteq k^n$ is the set of all solutions of the system of equations

$$f_1(x_1, \dots, x_n) = \dots = f_s(x_1, \dots, x_n) = 0.$$

Example 1.2.4. We will use $k = \mathbb{R}$ so that we can draw pictures. Consider the plane $\mathbb{R}^2$ with the variety

$$V(x^2 + y^2 - 1),$$

which is the unit circle centered at the origin [1, p. 6]:

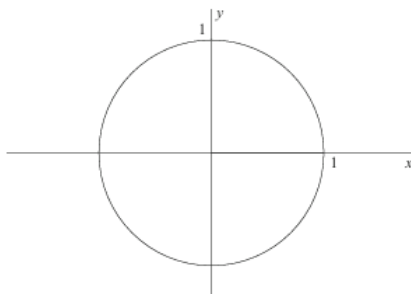


Figure 1.1: The affine variety $V(x^2 + y^2 - 1)$.

The conic sections studied in analytic geometry (circles, ellipses, parabolas, and hyperbolas) are affine varieties.

Example 1.2.5. Now consider the three-dimensional space $\mathbb{R}^3$. A nice affine variety is given by the paraboloid of revolution

$$V(z - x^2 - y^2),$$

which is obtained by rotating the parabola $z = x^2$ about the z -axis [1, p. 7]. This gives us the picture:

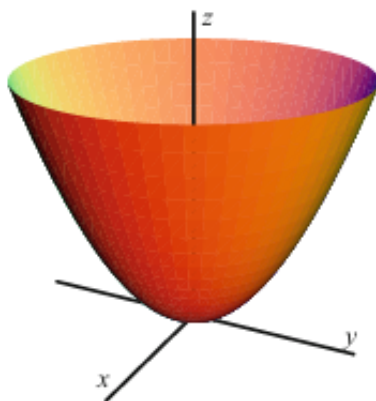


Figure 1.2: The affine variety $V(z - x^2 - y^2)$.

Consider another affine variety

$$V(z^2 - x^2 - y^2),$$

which corresponds to the cone $z^2 = x^2 + y^2$ [1, p. 7]. This gives the following picture:

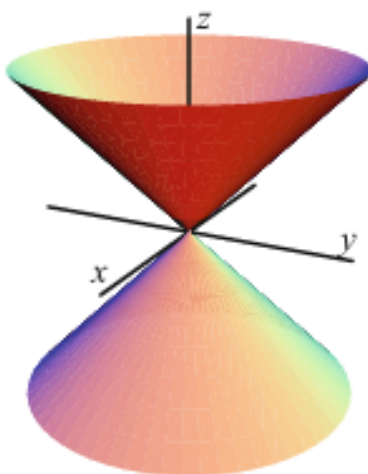


Figure 1.3: The affine variety $V(z^2 - x^2 - y^2)$.

Example 1.2.6. An interesting example of a curve in $\mathbb{R}^3$ is the *twisted cubic*, which is the affine variety

$$V(y - x^2, z - x^3)$$

[1, p. 8].

For simplicity, we will restrict to the portion that lies in the first octant. To begin,

we draw the surfaces $y = x^2$ and $z = x^3$ separately:

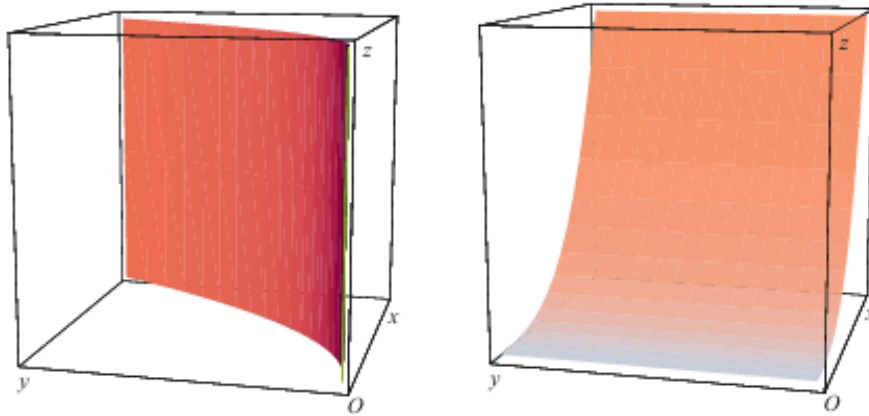


Figure 1.4: The surfaces $y = x^2$ and $z = x^3$ in $\mathbb{R}^3$.

Then their intersection gives the twisted cubic, i.e., the required variety:

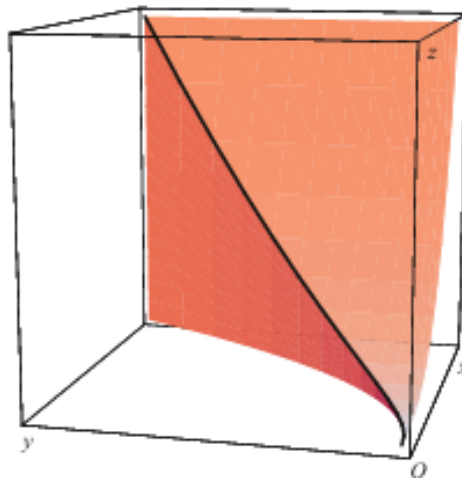


Figure 1.5: Twisted cubic curve given by $V(y - x^2, z - x^3)$.

Lemma 1.2.7 ([1, Lemma 2, Section 2, p. 11]). *If $V, W \subset k^n$ are affine varieties, then $V \cup W$ and $V \cap W$ are also affine varieties.*

Proof. Suppose that $V = V(f_1, \dots, f_s)$ and $W = V(g_1, \dots, g_t)$. Then our claim is

$$\begin{aligned} V \cap W &= V(f_1, \dots, f_s, g_1, \dots, g_t), \\ V \cup W &= V(f_i g_j \mid 1 \leq i \leq s, 1 \leq j \leq t). \end{aligned}$$

An n -tuple in k^n being in $V \cap W$ implies that all $f_1, \dots, f_s$ and $g_1, \dots, g_t$ vanish at that point, which is the same as $f_1, \dots, f_s, g_1, \dots, g_t$ vanishing there. Therefore, the first equality is proved.

Now let $(a_1, \dots, a_n) \in V$. Then all of the f_i vanish at this point, which implies that all of the $f_i g_j$ also vanish at $(a_1, \dots, a_n)$. Thus,

$$V \subseteq V(f_i g_j \mid 1 \leq i \leq s, 1 \leq j \leq t).$$

Similarly,

$$W \subseteq V(f_i g_j \mid 1 \leq i \leq s, 1 \leq j \leq t).$$

This proves that

$$V \cup W \subseteq V(f_i g_j \mid 1 \leq i \leq s, 1 \leq j \leq t).$$

For the other inclusion, let $(a_1, \dots, a_n) \in V(f_i g_j \mid 1 \leq i \leq s, 1 \leq j \leq t)$. If this point belongs to V , then we are done. If not, then $f_{i_0}(a_1, \dots, a_n) \neq 0$ for some i_0 ($1 \leq i_0 \leq s$). Since $f_{i_0} g_j$ vanishes at $(a_1, \dots, a_n)$ for all j , it follows (since k is a field) that all g_j must vanish at this point. Hence,

$$(a_1, \dots, a_n) \in W.$$

This shows that

$$V(f_i g_j \mid 1 \leq i \leq s, 1 \leq j \leq t) \subseteq V \cup W.$$

This completes the proof. $\square$

Corollary 1.2.8. *Finite intersections and unions of affine varieties are again affine varieties.*

Proof. The result follows by induction from Lemma 1.2.7. $\square$

As an example concerning unions, consider the union of the (x, y) -plane and the z -axis in affine 3-space. By the above formula, we have $V(z) \cup V(x, y) = V(zx, zy)$.

In the case of intersections, notice that the twisted cubic was given as the intersection of two surfaces (see Example 1.2.6).

The examples given in this section lead to some interesting questions concerning affine varieties. Suppose that we have $f_1, \dots, f_s \in k[x_1, \dots, x_n]$. Then:

- **(Consistency)** Can we determine if $V(f_1, \dots, f_s) = \emptyset$, i.e., do the equations $f_1 = \dots = f_s = 0$ always have a common solution?
- **(Finiteness)** Can we determine if $V(f_1, \dots, f_s)$ is finite, and if so, can we find all of the solutions explicitly?
- **(Dimension)** Can we determine the “dimension” of $V(f_1, \dots, f_s)$?

The answer to these questions is yes, although care must be taken in choosing the field k that we work over. Nevertheless, we will give complete solutions to all three problems.

1.3 Ideals

Definition 1.3.1. A subset $I \subset k[x_1, \dots, x_n]$ is an *ideal* if it satisfies the following:

- (i) $0 \in I$.
- (ii) If $f, g \in I$, then $f + g \in I$.
- (iii) If $f \in I$ and $h \in k[x_1, \dots, x_n]$, then $hf \in I$.

The first natural example of an ideal is the ideal generated by a finite number of polynomials.

Definition 1.3.2. Let $f_1, \dots, f_s$ be polynomials in $k[x_1, \dots, x_n]$. Then we set

$$\langle f_1, \dots, f_s \rangle = \left\{ \sum_{i=1}^s h_i f_i \mid h_1, \dots, h_s \in k[x_1, \dots, x_n] \right\}.$$

One can prove that $\langle f_1, \dots, f_s \rangle$ is an ideal. We will call $\langle f_1, \dots, f_s \rangle$ the *ideal generated by* $f_1, \dots, f_s$.

We say that an ideal I is *finitely generated* if there exist $f_1, \dots, f_s \in k[x_1, \dots, x_n]$ such that

$$I = \langle f_1, \dots, f_s \rangle,$$

and we say that $f_1, \dots, f_s$ are a *basis* of I .

In our later discussion, we will prove the amazing fact that every ideal of $k[x_1, \dots, x_n]$ is finitely generated (this is known as the *Hilbert Basis Theorem*).

Note that a given ideal may have many different bases. In later chapters, we will show that one can choose an especially useful type of basis, called a *Gröbner basis*.

Definition 1.3.3. Let I be an ideal in $k[x_1, \dots, x_n]$. The *variety of* I is defined as

$$V(I) = \{(a_1, \dots, a_n) \in k^n \mid f(a_1, \dots, a_n) = 0, \forall f \in I\}.$$

Remark 1.3.4. If I, J are two ideals of $k[x_1, \dots, x_n]$ such that

$$I \subset J,$$

then it follows directly from the definition that

$$V(I) \supset V(J).$$

Proposition 1.3.5 ([1, Proposition 9, Section 5, p. 81]). *If $f_1, \dots, f_s$ are a basis for an ideal $I \subset k[x_1, \dots, x_n]$, so that $I = \langle f_1, \dots, f_s \rangle$, then $V(I) = V(f_1, \dots, f_s)$.*

Proof. Let $(a_1, \dots, a_n) \in V(I)$. Then $f(a_1, \dots, a_n) = 0$, for all $f \in I$. Since $f_i \in I$, so $f_i(a_1, \dots, a_n) = 0$, for all $1 \leq i \leq s$, that is, $(a_1, \dots, a_n) \in V(f_1, \dots, f_s)$.

Conversely, let $(a_1, \dots, a_n) \in V(f_1, \dots, f_s)$. Then $f_i(a_1, \dots, a_n) = 0$, $\forall 1 \leq i \leq s$. Let $f \in I = \langle f_1, \dots, f_s \rangle$, then

$$f = \sum_{i=1}^s u_i f_i,$$

for some $u_i \in k[x_1, \dots, x_n]$. Thus $f(a_1, \dots, a_n) = 0$. Hence, $(a_1, \dots, a_n) \in V(I)$. Thus we have that $V(I) = V(f_1, \dots, f_s)$. $\square$

Example 1.3.6. Consider the variety $V(2x^2 + 3y^2 - 11, x^2 - y^2 - 3) \subset \mathbb{R}^2$. Let I be the ideal generated by $f_1 = 2x^2 + 3y^2 - 11$ and $f_2 = x^2 - y^2 - 3$.

Now, $f_1 = 2(x^2 - 4) + 3(y^2 - 1)$, $f_2 = (x^2 - 4) - (y^2 - 1)$. So, $I = \langle f_1, f_2 \rangle \subset \langle x^2 - 4, y^2 - 1 \rangle$.

Also, clearly,

$$\frac{f_1 + 3f_2}{5} = x^2 - 4,$$

and

$$y^2 - 1 = \frac{f_1 - 2(x^2 - 4)}{3} = \frac{f_1 - 2f_2}{5}.$$

Thus, $\langle x^2 - 4, y^2 - 1 \rangle \subset \langle f_1, f_2 \rangle = I$.

Therefore,

$$I = \langle f_1, f_2 \rangle = \langle x^2 - 4, y^2 - 1 \rangle.$$

Hence, by above Proposition,

$$V(2x^2 + 3y^2 - 11, x^2 - y^2 - 3) = V(I) = V(x^2 - 4, y^2 - 1) = \{(\pm 2, \pm 1)\}.$$

Definition 1.3.7. Let $V \subset k^n$ be an affine variety. Then we set

$$I(V) = \{f \in k[x_1, \dots, x_n] \mid f(a_1, \dots, a_n) = 0 \text{ for all } (a_1, \dots, a_n) \in V\}.$$

Lemma 1.3.8 ([1, Lemma 6, Section 4, p. 32]). *If $V \subset k^n$ is an affine variety, then $I(V) \subset k[x_1, \dots, x_n]$ is an ideal. We will call $I(V)$ the ideal of V .*

Proof. Clearly, $0 \in I(V)$, since the zero polynomial vanishes on all of k^n , and so, in particular, it vanishes on V . Next, suppose that $f, g \in I(V)$ and $h \in k[x_1, \dots, x_n]$. Let $(a_1, \dots, a_n)$ be an arbitrary point of V . Then

$$(f + g)(a_1, \dots, a_n) = f(a_1, \dots, a_n) + g(a_1, \dots, a_n) = 0 + 0 = 0,$$

and

$$(hf)(a_1, \dots, a_n) = h(a_1, \dots, a_n)f(a_1, \dots, a_n) = h(a_1, \dots, a_n) \cdot 0 = 0.$$

So, $f + g, hf \in I(V)$, since $(a_1, \dots, a_n)$ was chosen arbitrarily. Hence, $I(V)$ is an ideal. $\square$

Example 1.3.9. For an example of the ideal of a variety, consider the variety $\{(0, 0)\}$ consisting of the origin in k^2 . Then its ideal $I(\{(0, 0)\})$ consists of all polynomials in $k[x, y]$ that vanish at the origin, and we claim that

$$I(\{(0, 0)\}) = \langle x, y \rangle.$$

Any polynomial of the form $A(x, y)x + B(x, y)y$ vanishes at the origin. For the converse, suppose that

$$f = \sum_{i,j} a_{ij} x^i y^j$$

vanishes at the origin. Then

$$a_{00} = f(0, 0) = 0,$$

and consequently,

$$f = \sum_{(i,j) \neq (0,0)} a_{ij} x^i y^j.$$

Separating the terms according to whether $i > 0$ or $j > 0$, we may write

$$f = \left(\sum_{\substack{(i,j) \neq (0,0) \\ i > 0}} a_{ij} x^{i-1} y^j \right) x + \left(\sum_{\substack{(i,j) \neq (0,0) \\ j > 0}} a_{ij} x^i y^{j-1} \right) y.$$

Hence $f \in \langle x, y \rangle$, and therefore

$$I(\{(0, 0)\}) = \langle x, y \rangle.$$

Example 1.3.10. Consider the case when $V = k^n$. Then $I(k^n)$ consists of all polynomials that vanish at every point of k^n . Hence, by Proposition [1.2.2](#), we have

$$I(k^n) = \{0\},$$

whenever k is infinite.

Example 1.3.11. Now let's give a more interesting example. Consider the twisted cubic

$$V = V(y - x^2, z - x^3) \subset \mathbb{R}^3.$$

We claim that

$$I(V) = \langle y - x^2, z - x^3 \rangle.$$

To prove this, we will first show that given a polynomial $f \in \mathbb{R}[x, y, z]$, we can write f in the form

$$f = h_1(y - x^2) + h_2(z - x^3) + r, \tag{1.1}$$

where $h_1, h_2 \in \mathbb{R}[x, y, z]$ and r is a polynomial in the variable x alone. First, consider the case when f is a monomial $x^\alpha y^\beta z^\gamma$. Then the binomial theorem tells us that

$$\begin{aligned} x^\alpha y^\beta z^\gamma &= x^\alpha (x^2 + (y - x^2))^\beta (x^3 + (z - x^3))^\gamma \\ &= x^\alpha (x^{2\beta} + \text{terms involving } (y - x^2))(x^{3\gamma} + \text{terms involving } (z - x^3)), \end{aligned}$$

and multiplying this out shows that

$$x^\alpha y^\beta z^\gamma = h_1(y - x^2) + h_2(z - x^3) + x^{\alpha+2\beta+3\gamma},$$

for some polynomials $h_1, h_2 \in \mathbb{R}[x, y, z]$. Thus, (1.1) is true in this case. Since an arbitrary $f \in \mathbb{R}[x, y, z]$ is an $\mathbb{R}$ -linear combination of monomials, it follows that (1.1) holds in general.

We now prove $I(V) = \langle y - x^2, z - x^3 \rangle$. By the definition of the twisted cubic V , we have $y - x^2, z - x^3 \in I(V)$, and since $I(V)$ is an ideal, it follows that $h_1(y - x^2) + h_2(z - x^3) \in I(V)$. This proves that

$$\langle y - x^2, z - x^3 \rangle \subset I(V).$$

To prove the opposite inclusion, let $f \in I(V)$ and let

$$f = h_1(y - x^2) + h_2(z - x^3) + r$$

be the decomposition given by (1.1). To prove that r is zero, we will use the parametrization (t, t^2, t^3) of the twisted cubic. Since f vanishes on V , we obtain

$$0 = f(t, t^2, t^3) = 0 + 0 + r(t).$$

Since t can be any real number, $r \in \mathbb{R}[x]$ must be the zero polynomial by Proposition 1.2.2. But $r = 0$ shows that $f \in \langle y - x^2, z - x^3 \rangle$, and therefore $I(V) = \langle y - x^2, z - x^3 \rangle$.

The construction of the ideal $I(V)$ above is a very important one. It provides the bridge from geometry back to algebra since, in addition to the map

$$\begin{aligned} \{\text{Subsets of } k[x_1, \dots, x_n]\} &\longrightarrow \{\text{Varieties of } k^n\} \\ S &\longmapsto V(S), \end{aligned}$$

we now have a map

$$\begin{aligned} \{\text{Varieties of } k^n\} &\longrightarrow \{\text{Ideals of } k[x_1, \dots, x_n]\} \\ V &\longmapsto I(V). \end{aligned}$$

Understanding the relationship between these two maps allows us to pass back and forth between algebraic and geometric questions. In particular, we will be interested in the relationship between an ideal I and the ideal $I(V(I))$.

It is easy to see that

$$I \subseteq I(V(I)),$$

but equality does not always hold. Thus, one may ask whether

$$I(V(f_1, \dots, f_s)) = \langle f_1, \dots, f_s \rangle.$$

Unfortunately, the answer is not always affirmative. The following result gives the best general statement that can be made at this stage.

Lemma 1.3.12 ([1] Lemma 7, Section 4, p. 34]. *Let $I \subset k[x_1, \dots, x_n]$ be an ideal. Then*

$$I \subseteq I(V(I)).$$

Equality need not hold in general.

Proof. Let $f \in I$. Suppose that $(a_1, \dots, a_n) \in V(I)$. Then, by definition of the variety of an ideal,

$$g(a_1, \dots, a_n) = 0 \quad \text{for all } g \in I.$$

In particular, since $f \in I$, we have

$$f(a_1, \dots, a_n) = 0.$$

Since $(a_1, \dots, a_n) \in V(I)$ was arbitrary, it follows that f vanishes on all of $V(I)$. Hence,

$$f \in I(V(I)).$$

Therefore $I \subseteq I(V(I))$. □

For showing that equality may not hold, we need an example where $I(V(I))$ is strictly bigger than I . For this, consider $I = \langle x^2, y^2 \rangle$. By Proposition 1.3.5, $I(V(\langle x^2, y^2 \rangle)) = I(V(x^2, y^2))$. So, we first compute $I(V(x^2, y^2))$. The equations $x^2 = 0$ and $y^2 = 0$ imply that $V(x^2, y^2) = \{(0, 0)\}$. But Example 1.3.9 showed that the ideal of $\{(0, 0)\}$ is $\langle x, y \rangle$, so that $I(V(x^2, y^2)) = \langle x, y \rangle$. To see that this is strictly larger than $\langle x^2, y^2 \rangle$, note that $x \notin \langle x^2, y^2 \rangle$, since for polynomials of the form $h_1(x, y)x^2 + h_2(x, y)y^2$, every monomial has total degree at least two.

Proposition 1.3.13 ([1] Proposition 4, Section 4, p. 34]. *Let V and W be affine varieties in k^n . Then:*

(i) $V \subset W$ if and only if $I(V) \supset I(W)$.

(ii) $V = W$ if and only if $I(V) = I(W)$.

Proof. To prove (i), first suppose that $V \subset W$. Then any polynomial vanishing on W must vanish on V , which proves that $I(W) \subset I(V)$.

Next, assume that $I(W) \subset I(V)$. We know that W is the variety defined by some polynomials $g_1, \dots, g_t \in k[x_1, \dots, x_n]$. Then $g_1, \dots, g_t \in I(W) \subset I(V)$, and hence the g_i 's vanish on V . Since W consists of all common zeros of the g_i 's, it follows that $V \subset W$.

The proof of (ii) follows immediately from part (i). □

Chapter 2

Introduction to Gröbner Bases

The theory of Gröbner bases, introduced by Bruno Buchberger in 1965, provides an algorithmic framework for studying ideals in polynomial rings and solving systems of polynomial equations. In this chapter, we develop the basic theory of Gröbner bases, beginning with monomial orderings and the division algorithm in multivariate polynomial rings. We then establish the Hilbert Basis Theorem, introduce Gröbner bases and their fundamental properties, and conclude with Buchberger's Criterion, which gives a practical characterization for determining whether a given basis is a Gröbner basis.

2.1 Ordering on the Monomials in $k[x_1, \dots, x_n]$

Definition 2.1.1. A *monomial ordering* $>$ on $k[x_1, \dots, x_n]$ is a relation $>$ on $\mathbb{Z}_{\geq 0}^n$, or equivalently, a relation on the set of monomials x^α , $\alpha \in \mathbb{Z}_{\geq 0}^n$, satisfying:

- (i) $>$ is a total ordering on $\mathbb{Z}_{\geq 0}^n$.
- (ii) If $\alpha > \beta$ and $\gamma \in \mathbb{Z}_{\geq 0}^n$, then
$$\alpha + \gamma > \beta + \gamma.$$
- (iii) $>$ is a well-ordering on $\mathbb{Z}_{\geq 0}^n$. This means that every nonempty subset of $\mathbb{Z}_{\geq 0}^n$ has a smallest element under $>$.

The following lemma clarifies the meaning and significance of the well-ordering condition in part (iii) of the above definition.

Lemma 2.1.2 ([1, Lemma 2, Section 2, p. 56]). *An order relation $>$ on $\mathbb{Z}_{\geq 0}^n$ is a well-ordering if and only if every strictly decreasing sequence in $\mathbb{Z}_{\geq 0}^n$,*

$$\alpha^{(1)} > \alpha^{(2)} > \alpha^{(3)} > \dots$$

eventually terminates.

Proof. We will prove this in contrapositive form: $>$ is not a well-ordering if and only if there is an infinite strictly decreasing sequence in $\mathbb{Z}_{\geq 0}^n$.

Suppose that $>$ is not a well-ordering. Then some non-empty subset $S \subset \mathbb{Z}_{\geq 0}^n$ has no least element. Now pick $\alpha^{(1)} \in S$. Since $\alpha^{(1)}$ is not the least element, we can find $\alpha^{(2)} \in S$

such that $\alpha^{(1)} > \alpha^{(2)}$. Then $\alpha^{(2)}$ is also not the least element, so there exists $\alpha^{(3)} \in S$ such that $\alpha^{(2)} > \alpha^{(3)}$. Continuing in this way, we obtain an infinite strictly decreasing sequence

$$\alpha^{(1)} > \alpha^{(2)} > \alpha^{(3)} > \dots$$

Now assume that there exists an infinite strictly decreasing sequence

$$\alpha^{(1)} > \alpha^{(2)} > \alpha^{(3)} > \dots$$

Then the set $\{\alpha^{(1)}, \alpha^{(2)}, \alpha^{(3)}, \dots\}$ is a non-empty subset of $\mathbb{Z}_{\geq 0}^n$ having no least element. Hence $>$ is not a well-ordering. $\square$

Definition 2.1.3 (Lexicographic Order). Let $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ and $\beta = (\beta_1, \beta_2, \dots, \beta_n) \in \mathbb{Z}_{\geq 0}^n$. We say that

$$\alpha >_{\text{lex}} \beta$$

if, in the vector difference $\alpha - \beta \in \mathbb{Z}^n$, the leftmost nonzero entry is positive. We write

$$x^\alpha >_{\text{lex}} x^\beta$$

if $\alpha >_{\text{lex}} \beta$.

Example 2.1.4. Some examples of the lexicographic order are as follows:

(a) Since the leftmost nonzero entry of

$$(5, 3, 0) - (4, 4, 2) = (1, -1, -2)$$

is positive, we have

$$(5, 3, 0) >_{\text{lex}} (4, 4, 2).$$

(b) Since the leftmost nonzero entry of

$$(8, 7, 4) - (8, 7, 1) = (0, 0, 3)$$

is positive, we have

$$(8, 7, 4) >_{\text{lex}} (8, 7, 1).$$

(c) The variables $x_1, \dots, x_n$ are ordered in the usual way by the lex ordering:

$$(1, 0, 0, \dots, 0) >_{\text{lex}} (0, 1, 0, \dots, 0) >_{\text{lex}} \dots >_{\text{lex}} (0, \dots, 0, 1).$$

So,

$$x_1 >_{\text{lex}} x_2 >_{\text{lex}} \dots >_{\text{lex}} x_n.$$

Definition 2.1.5 (Graded Lexicographic Order). Let $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ and $\beta = (\beta_1, \beta_2, \dots, \beta_n) \in \mathbb{Z}_{\geq 0}^n$. We say that

$$\alpha >_{\text{grlex}} \beta$$

if either

$$|\alpha| > |\beta|,$$

or

$$|\alpha| = |\beta| \quad \text{and} \quad \alpha >_{\text{lex}} \beta.$$

Example 2.1.6. Some examples of the graded lexicographic order are as follows:

(a) Since

$$|(2, 4, 3)| = 9 > 3 = |(1, 2, 0)|,$$

we have

$$(2, 4, 3) >_{\text{grlex}} (1, 2, 0).$$

(b) Since

$$|(7, 3, 4)| = 14 = |(5, 1, 8)|$$

and

$$(7, 3, 4) >_{\text{lex}} (5, 1, 8),$$

we have

$$(7, 3, 4) >_{\text{grlex}} (5, 1, 8).$$

(c) The variables are ordered according to the lex order, i.e.,

$$x_1 >_{\text{grlex}} x_2 >_{\text{grlex}} \cdots >_{\text{grlex}} x_n.$$

Proposition 2.1.7 ([1, Proposition 4, Section 2, p. 57]). *The lex, grlex, and grevlex orderings on $\mathbb{Z}_{\geq 0}^n$ are monomial orderings.*

Proof. See the proof of Proposition 4 in [1, Section 2, p. 57]. □

We conclude this section with a discussion of how a monomial ordering can be applied to polynomials. If

$$f = \sum_{\alpha} a_{\alpha} x^{\alpha} \in k[x_1, \dots, x_n]$$

and a monomial ordering $>$ has been selected, then the monomials of f can be arranged uniquely in decreasing order with respect to $>$. For example, let

$$f = 4xy^2z + 4z^2 - 5x^3 + 7x^2z^2 \in k[x, y, z].$$

Reordering the terms of f in decreasing order, we obtain:

(a) With respect to the lex order,

$$f = -5x^3 + 7x^2z^2 + 4xy^2z + 4z^2.$$

(b) With respect to the grlex order,

$$f = 7x^2z^2 + 4xy^2z - 5x^3 + 4z^2.$$

(c) With respect to the grevlex order,

$$f = 4xy^2z + 7x^2z^2 - 5x^3 + 4z^2.$$

We will use the following terminology:

Definition 2.1.8. Let

$$f = \sum_{\alpha} a_{\alpha} x^{\alpha}$$

be a nonzero polynomial in $k[x_1, \dots, x_n]$, and let $>$ be a monomial ordering.

(i) The *multidegree* of f is

$$\text{multideg}(f) = \max\{\alpha \in \mathbb{Z}_{\geq 0}^n \mid a_{\alpha} \neq 0\},$$

where the maximum is taken with respect to $>$.

(ii) The *leading coefficient* of f is

$$LC(f) = a_{\text{multideg}(f)} \in k.$$

(iii) The *leading monomial* of f is

$$LM(f) = x^{\text{multideg}(f)},$$

with coefficient 1.

(iv) The *leading term* of f is

$$LT(f) = LC(f) \cdot LM(f).$$

To illustrate, Let $f = 4xy^2z + 4z^2 - 5x^3 + 7x^2z^2 \in k[x, y, z]$ and let $>$ denote the lex order. Then

$$\text{multideg}(f) = (3, 0, 0), \quad LC(f) = -5, \quad LM(f) = x^3, \quad LT(f) = -5x^3.$$

2.2 A Division Algorithm in $k[x_1, \dots, x_n]$

To study the division process in the multivariable case, we formulate a division algorithm for polynomials in $k[x_1, \dots, x_n]$, which extends the usual division algorithm in $k[x]$.

In the general setting, the goal is to divide a polynomial $f \in k[x_1, \dots, x_n]$ by polynomials $f_1, \dots, f_s \in k[x_1, \dots, x_n]$. As we will see, this means expressing f in the form

$$f = q_1 f_1 + \dots + q_s f_s + r,$$

where the quotients $q_1, \dots, q_s \in k[x_1, \dots, x_n]$ and the remainder $r \in k[x_1, \dots, x_n]$.

The basic idea of the algorithm is the same as in the one-variable case. We aim to cancel the leading term of f (with respect to a fixed monomial ordering) by multiplying one of the f_i by a suitable monomial and subtracting it from f . This process is then repeated, and each such step contributes to the corresponding quotient a_i .

We illustrate this procedure with an example:

Example 2.2.1. We will first divide $f = xy^2 + 1$ by $f_1 = xy + 1$ and $f_2 = y + 1$, using lex order with $x > y$. Listing the divisors f_1, f_2 and the quotients q_1, q_2 vertically, we have the following setup:

$$\begin{array}{r} q_1 : \\ q_2 : \\ xy + 1 \\ y + 1 \end{array} \bigg) \overline{xy^2 + 1}$$

The leading terms $LT(f_1) = xy$ and $LT(f_2) = y$ both divide the leading term $LT(f) = xy^2$. Since f_1 is listed first, we will use it. Thus, we divide xy into xy^2 , leaving y , and then subtract $y \cdot f_1$ from f :

$$\begin{array}{r} q_1 : y \\ q_2 : \\ xy + 1 \\ y + 1 \end{array} \bigg) \begin{array}{r} \overline{xy^2 + 1} \\ xy^2 + y \\ \hline -y + 1 \end{array}$$

Now we repeat the same process on $-y + 1$. This time we must use f_2 since $LT(f_1) = xy$ does not divide $LT(-y + 1) = -y$. We obtain:

$$\begin{array}{r} q_1 : y \\ q_2 : -1 \\ xy + 1 \\ y + 1 \end{array} \bigg) \begin{array}{r} \overline{xy^2 + 1} \\ xy^2 + y \\ \hline -y + 1 \\ -y - 1 \\ \hline 2 \end{array}$$

Since $LT(f_1)$ and $LT(f_2)$ do not divide 2, the remainder is $r = 2$. Thus:

$$xy^2 + 1 = y \cdot (xy + 1) + (-1) \cdot (y + 1) + 2$$

Theorem 2.2.2. [1, Theorem 3, Section 3, p. 64] (The Division Algorithm) Fix a monomial order $>$ on $\mathbb{Z}_{\geq 0}^n$, and let $F = (f_1, \dots, f_s)$ be an ordered s -tuple of polynomials in $k[x_1, \dots, x_n]$. Then every $f \in k[x_1, \dots, x_n]$ can be written as

$$f = q_1 f_1 + \dots + q_s f_s + r,$$

where $q_i, r \in k[x_1, \dots, x_n]$, and either $r = 0$ or r is a k -linear combination of monomials, none of which is divisible by any of

$$LT(f_1), \dots, LT(f_s).$$

We call r a remainder of f on division by F .

Furthermore, if $q_i f_i \neq 0$, then

$$\text{multideg}(f) \geq \text{multideg}(q_i f_i).$$

Proof. See the proof of Theorem 3 in [1, Section 3, p. 64]. □

Motivating Problems:

We now describe several fundamental problems in computational algebra that motivate the theory of Gröbner bases.

- (1) **The Ideal Description Problem:** Does every ideal $I \subset k[x_1, \dots, x_n]$ have a finite generating set? In other words, can we write

$$I = \langle f_1, \dots, f_s \rangle$$

for some $f_i \in k[x_1, \dots, x_n]$?

- (2) **The Ideal Membership Problem:** Given $f \in k[x_1, \dots, x_n]$ and an ideal $I = \langle f_1, \dots, f_s \rangle$, determine whether $f \in I$. Geometrically, this is closely related to determining whether the variety $V(f_1, \dots, f_s)$ lies inside the hypersurface $V(f)$.
- (3) **The Problem of Solving Polynomial Equations:** Find all common solutions in k^n of a system of polynomial equations

$$f_1(x_1, \dots, x_n) = \dots = f_s(x_1, \dots, x_n) = 0.$$

Equivalently, this is the problem of finding the points in the affine variety $V(f_1, \dots, f_s)$.

- (4) **The Implicitization Problem:** Let $V \subset k^n$ be given parametrically by

$$x_1 = g_1(t_1, \dots, t_m), \quad \dots, \quad x_n = g_n(t_1, \dots, t_m).$$

If the g_i are polynomials (or rational functions) in the parameters t_j , then V defines an affine variety or a part of one. The goal is to find a system of polynomial equations in the variables x_i that defines the variety V .

One implication is an easy corollary of Theorem 2.2.2. If, after division of f by $F = (f_1, \dots, f_s)$, we obtain a remainder $r = 0$, then

$$f = a_1 f_1 + \dots + a_s f_s,$$

so that $f \in \langle f_1, \dots, f_s \rangle$. Thus, $r = 0$ is a sufficient condition for ideal membership.

However, as the following example shows, $r = 0$ is not a necessary condition for membership in the ideal.

Example 2.2.3. Let $f_1 = xy + 1$, $f_2 = y^2 - 1 \in k[x, y]$ with the lexicographic order.

Dividing $f = xy^2 - x$ by $F = (f_1, f_2)$, we obtain

$$xy^2 - x = y \cdot (xy + 1) + 0 \cdot (y^2 - 1) + (-x - y).$$

With $F = (f_2, f_1)$, however, we have

$$xy^2 - x = x \cdot (y^2 - 1) + 0 \cdot (xy + 1) + 0.$$

The second calculation shows that $f \in \langle f_1, f_2 \rangle$. The first calculation shows that even if $f \in \langle f_1, f_2 \rangle$, it is still possible to obtain a nonzero remainder on division by $F = (f_1, f_2)$.

In dealing with a collection of polynomials $f_1, \dots, f_s \in k[x_1, \dots, x_n]$, it is frequently desirable to pass to the ideal $I = \langle f_1, \dots, f_s \rangle$ they generate. This allows the possibility of replacing $f_1, \dots, f_s$ with a different generating set for I . Hence, we may ask whether there exists a “good” generating set for I .

For such a set, we would want the remainder r obtained on division by these “good” generators to be uniquely determined, and the condition $r = 0$ should be equivalent to membership in the ideal.

Later, we will see that Gröbner bases have exactly these desirable properties.

2.3 Monomial Ideals and Dickson’s Lemma

In this section, we consider the ideal description problem for the special case of monomial ideals. We begin by defining monomial ideals in $k[x_1, \dots, x_n]$.

Definition 2.3.1. An ideal $I \subset k[x_1, \dots, x_n]$ is called a *monomial ideal* if there exists a subset $A \subset \mathbb{Z}_{\geq 0}^n$ (possibly infinite) such that I consists of all polynomials which are finite sums of the form

$$\sum_{\alpha \in A} h_{\alpha} x^{\alpha},$$

where $h_{\alpha} \in k[x_1, \dots, x_n]$.

In this case, we write

$$I = \langle x^{\alpha} \mid \alpha \in A \rangle.$$

Example 2.3.2. An example of a monomial ideal is given by

$$I = \langle x^4 y^2, x^3 y^4, x^2 y^5 \rangle \subset k[x, y].$$

Lemma 2.3.3 ([1] Lemma 2, Section 4, p. 70). *Let $I = \langle x^{\alpha} \mid \alpha \in A \rangle$ be a monomial ideal. Then a monomial x^{β} lies in I if and only if x^{β} is divisible by x^{α} for some $\alpha \in A$.*

Proof. If x^β is a multiple of x^α for some $\alpha \in A$, then $x^\beta \in I$ by the definition of an ideal. Conversely, suppose that $x^\beta \in I$. Then

$$x^\beta = \sum_{i=1}^s h_i x^{\alpha(i)},$$

where $h_i \in k[x_1, \dots, x_n]$ and $\alpha(i) \in A$. Expanding each h_i as a linear combination of monomials, we see that every term on the right-hand side is divisible by some $x^{\alpha(i)}$. Hence, the left side x^β must also be divisible by some $x^{\alpha(i)}$. $\square$

Note that x^β is divisible by x^α if and only if

$$x^\beta = x^\alpha \cdot x^\gamma$$

for some $\gamma \in \mathbb{Z}_{\geq 0}^n$. This is equivalent to

$$\beta = \alpha + \gamma.$$

Thus, the set

$$\alpha + \mathbb{Z}_{\geq 0}^n = \{\alpha + \gamma \mid \gamma \in \mathbb{Z}_{\geq 0}^n\}$$

consists of the exponents of all monomials divisible by x^α .

This observation allows us to draw pictures of the monomials in a given monomial ideal.

For example, if

$$I = \langle x^6, x^5y^4, x^3y^6 \rangle,$$

then the exponents of the monomials in I form the set

$$((6, 0) + \mathbb{Z}_{\geq 0}^2) \cup ((5, 4) + \mathbb{Z}_{\geq 0}^2) \cup ((3, 6) + \mathbb{Z}_{\geq 0}^2).$$

We can visualize this set as the union of the integer points in three translated copies of the first quadrant in the plane:

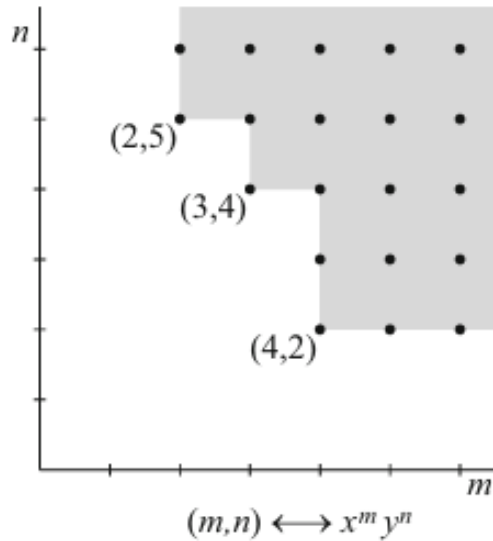


Figure 2.1: Exponent set of the monomial ideal $I = \langle x^6, x^5y^4, x^3y^6 \rangle$.

Some books use different terminology. In Eisenbud [5], the leading term $LT(f)$ becomes the initial term $in_>(f)$. A more substantial difference appears in Becker and Weispfenning [6], where the meanings of “monomial” and “term” are interchanged. For them, the leading term $LT(f)$ is the head monomial $HM(f)$, while the leading monomial $LM(f)$ is the head term $HT(f)$. Page 10 of Kreuzer and Robbiano [7] contains a summary of the terminology used in different books. Our advice when reading other texts is to check the definitions carefully.

Theorem 2.3.4 (Dickson’s Lemma [8, p. 414]). *Let $I = \langle x^\alpha \mid \alpha \in A \rangle \subset k[x_1, \dots, x_n]$ be a monomial ideal. Then I can be written in the form*

$$I = \langle x^{\alpha(1)}, \dots, x^{\alpha(s)} \rangle,$$

where $\alpha(1), \dots, \alpha(s) \in A$.

In particular, every monomial ideal has a finite generating set.

Proof. We proceed by induction on n , the number of variables.

If $n = 1$, then I is generated by the monomials x_1^α , where $\alpha \in A \subset \mathbb{Z}_{\geq 0}$. Let β be the smallest element of $A \subset \mathbb{Z}_{\geq 0}$. Then $\beta \leq \alpha$ for all $\alpha \in A$, so that x_1^β divides all other generators x_1^α . Therefore,

$$I = \langle x_1^\beta \rangle.$$

Now assume that $n > 1$ and that the theorem is true for $n - 1$ variables. We will write the variables as $x_1, \dots, x_{n-1}, y$, so that monomials in $k[x_1, \dots, x_{n-1}, y]$ can be written as $x^\alpha y^m$, where $\alpha = (\alpha_1, \dots, \alpha_{n-1}) \in \mathbb{Z}_{\geq 0}^{n-1}$ and $m \in \mathbb{Z}_{\geq 0}$.

Suppose that $I \subset k[x_1, \dots, x_{n-1}, y]$ is a monomial ideal. To find generators for I , let J be the ideal in $k[x_1, \dots, x_{n-1}]$ generated by the monomials x^α for which $x^\alpha y^m \in I$ for some $m \geq 0$. Since J is a monomial ideal in $k[x_1, \dots, x_{n-1}]$, our inductive hypothesis implies that finitely many of the x^α ’s generate J , say

$$J = \langle x^{\alpha(1)}, \dots, x^{\alpha(s)} \rangle.$$

The ideal J can be understood as the “projection” of I into $k[x_1, \dots, x_{n-1}]$.

For each i between 1 and s , the definition of J tells us that $x^{\alpha(i)} y^{m_i} \in I$ for some $m_i \geq 0$. Let m be the largest of the m_i . Then, for each k between 0 and $m - 1$, consider the monomial ideal $J_k \subset k[x_1, \dots, x_{n-1}]$ generated by the monomials x^β such that $x^\beta y^k \in I$. One can think of J_k as the “slice” of I generated by monomials containing y exactly to the k th power. Using our induction hypothesis again, J_k has a finite generating set of monomials, say

$$J_k = \langle x^{\alpha_k(1)}, \dots, x^{\alpha_k(s_k)} \rangle.$$

We claim that I is generated by the monomials in the following list:

$$\begin{aligned} \text{from } J : & \quad x^{\alpha(1)} y^m, \dots, x^{\alpha(s)} y^m, \\ \text{from } J_0 : & \quad x^{\alpha_0(1)}, \dots, x^{\alpha_0(s_0)}, \\ \text{from } J_1 : & \quad x^{\alpha_1(1)} y, \dots, x^{\alpha_1(s_1)} y, \\ & \quad \vdots \\ \text{from } J_{m-1} : & \quad x^{\alpha_{m-1}(1)} y^{m-1}, \dots, x^{\alpha_{m-1}(s_{m-1})} y^{m-1}. \end{aligned}$$

We claim that every monomial in I is divisible by one on the list. To see why, let $x^\alpha y^p \in I$. If $p \geq m$, then $x^\alpha y^p$ is divisible by some $x^{\alpha(i)} y^m$ by the construction of J . On the other hand, if $p \leq m - 1$, then $x^\alpha y^p$ is divisible by some $x^{\alpha_p(j)} y^p$ by the construction of J_p .

It follows from Lemma 2.3.3 that the above monomials generate an ideal having the same monomials as I . This forces the ideals to be the same, and our claim is proved.

To complete the proof of the theorem, we need to show that the finite set of generators can be chosen from a given set of generators for the ideal. If we switch back to writing the variables as $x_1, \dots, x_n$, then our monomial ideal is

$$I = \langle x^\alpha \mid \alpha \in A \rangle \subset k[x_1, \dots, x_n].$$

We need to show that I is generated by finitely many of the x^α 's, where $\alpha \in A$. By the previous paragraph, we know that

$$I = \langle x^{\beta(1)}, \dots, x^{\beta(s)} \rangle$$

for some monomials $x^{\beta(i)} \in I$. Since $x^{\beta(i)} \in I = \langle x^\alpha \mid \alpha \in A \rangle$, Lemma 2.3.3 tells us that each $x^{\beta(i)}$ is divisible by $x^{\alpha(i)}$ for some $\alpha(i) \in A$. From here, it is easy to show that

$$I = \langle x^{\alpha(1)}, \dots, x^{\alpha(s)} \rangle.$$

This completes the proof. □

2.4 The Hilbert Basis Theorem and Gröbner Bases

Definition 2.4.1. Let $I \subset k[x_1, \dots, x_n]$ be a nonzero ideal.

(i) We denote by $LT(I)$ the set of leading terms of elements of I . Thus,

$$LT(I) = \{cx^\alpha \mid \text{there exists } f \in I \text{ with } LT(f) = cx^\alpha\}.$$

(ii) We denote by $\langle LT(I) \rangle$ the ideal generated by the elements of $LT(I)$.

Remark 2.4.2. If we are given a finite generating set for I , say $I = \langle f_1, \dots, f_s \rangle$, then the ideals $\langle LT(f_1), \dots, LT(f_s) \rangle$ and $\langle LT(I) \rangle$ need not be equal.

Indeed, by definition we have $LT(f_i) \in LT(I)$ for each i , which implies $\langle LT(f_1), \dots, LT(f_s) \rangle \subset \langle LT(I) \rangle$.

However, the inclusion may be strict. The following example illustrates this phenomenon.

Example 2.4.3. Let $I = \langle f_1, f_2 \rangle$, where $f_1 = x^3 - 2xy$ and $f_2 = x^2y - 2y^2 + x$, and use the graded lexicographic ordering on monomials in $k[x, y]$. Then

$$x^2 = x \cdot (x^2y - 2y^2 + x) - y \cdot (x^3 - 2xy),$$

so that $x^2 \in I$. Thus, $x^2 = LT(x^2) \in LT(I)$.

However, x^2 is not divisible by $LT(f_1) = x^3$ or $LT(f_2) = x^2y$, so that $x^2 \notin \langle LT(f_1), LT(f_2) \rangle$.

Proposition 2.4.4 ([1] Proposition 3, Section 5, p. 77]. *Let $I \subset k[x_1, \dots, x_n]$ be an ideal.*

(i) $\langle LT(I) \rangle$ is a monomial ideal.

(ii) There exist $g_1, \dots, g_t \in I$ such that $\langle LT(I) \rangle = \langle LT(g_1), \dots, LT(g_t) \rangle$.

Proof. (i) The leading monomials $LM(g)$ of elements $g \in I \setminus \{0\}$ generate the monomial ideal $\langle LM(g) \mid g \in I \setminus \{0\} \rangle$. Since $LM(g)$ and $LT(g)$ differ by a nonzero constant, this ideal equals $\langle LT(g) \mid g \in I \setminus \{0\} \rangle = \langle LT(I) \rangle$. Thus, $\langle LT(I) \rangle$ is a monomial ideal.

(ii) Since $\langle LT(I) \rangle$ is generated by the monomials $LM(g)$ for $g \in I \setminus \{0\}$, Dickson's Lemma implies that $\langle LT(I) \rangle = \langle LM(g_1), \dots, LM(g_t) \rangle$ for finitely many $g_1, \dots, g_t \in I$. Since $LM(g_i)$ differs from $LT(g_i)$ by a nonzero constant, it follows that $\langle LT(I) \rangle = \langle LT(g_1), \dots, LT(g_t) \rangle$.

This completes the proof. $\square$

We now use the above proposition together with the division algorithm to prove that every polynomial ideal has a finite generating set, thereby giving an affirmative answer to the ideal description problem.

Let $I \subset k[x_1, \dots, x_n]$ be any ideal, and consider the associated ideal $\langle LT(I) \rangle$. As always, we fix a monomial order to be used in the division algorithm and in the computation of leading terms.

Theorem 2.4.5 (Hilbert Basis Theorem [9] p. 474]. *Every ideal $I \subset k[x_1, \dots, x_n]$ has a finite generating set. That is,*

$$I = \langle g_1, \dots, g_t \rangle$$

for some $g_1, \dots, g_t \in I$.

Proof. If $I = \{0\}$, we take our generating set to be $\{0\}$, which is certainly finite.

If I contains some nonzero polynomial, then a generating set $g_1, \dots, g_t$ for I can be constructed as follows. By Proposition 2.4.4, there exist $g_1, \dots, g_t \in I$ such that $\langle LT(I) \rangle = \langle LT(g_1), \dots, LT(g_t) \rangle$. We claim that $I = \langle g_1, \dots, g_t \rangle$. Clearly, $\langle g_1, \dots, g_t \rangle \subset I$, since each $g_i \in I$. For the reverse inclusion, let $f \in I$ be arbitrary. Applying the division algorithm (Theorem 2.2.2) to divide f by $(g_1, \dots, g_t)$, we obtain

$$f = a_1 g_1 + \dots + a_t g_t + r,$$

where no term of r is divisible by any of $LT(g_1), \dots, LT(g_t)$. We claim that $r = 0$. To see this, note that

$$r = f - a_1 g_1 - \dots - a_t g_t \in I.$$

If $r \neq 0$, then $LT(r) \in LT(I) \subset \langle LT(I) \rangle = \langle LT(g_1), \dots, LT(g_t) \rangle$. Hence, by Lemma 2.3.3, it follows that $LT(r)$ must be divisible by some $LT(g_i)$. This contradicts the definition of the remainder in the division algorithm.

Therefore, $r = 0$, and so

$$f = a_1 g_1 + \dots + a_t g_t \in \langle g_1, \dots, g_t \rangle.$$

Thus,

$$I \subset \langle g_1, \dots, g_t \rangle.$$

Hence,

$$I = \langle g_1, \dots, g_t \rangle.$$

This completes the proof. $\square$

Remark 2.4.6. In addition to answering the ideal description problem, the generating set $\{g_1, \dots, g_t\}$ used in the proof of the Hilbert Basis Theorem has the special property that

$$\langle LT(I) \rangle = \langle LT(g_1), \dots, LT(g_t) \rangle.$$

We have already seen in Example 2.4.3 that not every generating set of an ideal has this property. We now give a special name to generating sets satisfying this condition.

Definition 2.4.7. Fix a monomial order. A finite subset $G = \{g_1, \dots, g_t\} \subset I$ is called a *Gröbner basis* (or *standard basis*) of I if

$$\langle LT(I) \rangle = \langle LT(g_1), \dots, LT(g_t) \rangle.$$

Equivalently, a finite subset $\{g_1, \dots, g_t\} \subset I$ is a Gröbner basis of I if and only if the leading term of every element of I is divisible by one of the $LT(g_i)$.

Corollary 2.4.8 ([1 Corollary 6, Section 5, p. 78]). *Fix a monomial order. Then every ideal*

$$I \subset k[x_1, \dots, x_n], \quad I \neq \{0\},$$

has a Gröbner basis. Furthermore, every Gröbner basis of an ideal I generates I .

Proof. Given a nonzero ideal I , the set

$$G = \{g_1, \dots, g_t\}$$

constructed in the proof of the Hilbert Basis Theorem is a Gröbner basis by definition. Thus, the first assertion follows.

For the second assertion, note that if $\langle LT(I) \rangle = \langle LT(g_1), \dots, LT(g_t) \rangle$, then the argument used in the proof of the Hilbert Basis Theorem shows that $I = \langle g_1, \dots, g_t \rangle$.

Hence, G generates I . $\square$

Remark 2.4.9. Let $I \subset k[x_1, \dots, x_n]$ be an ideal. By the Hilbert Basis Theorem, there exists a finite generating set

$$I = \langle g_1, \dots, g_t \rangle$$

for some $g_1, \dots, g_t \in I$. Hence, by Proposition 1.3.5,

$$V(I) = V(\langle g_1, \dots, g_t \rangle) = V(g_1, \dots, g_t).$$

Therefore, $V(I)$ is an affine variety.

2.5 Properties of Gröbner Bases

We first prove that the remainder is uniquely determined when we divide by a Gröbner basis.

Proposition 2.5.1 ([1, Proposition 1, Section 6, p. 83]). *Let*

$$G = \{g_1, \dots, g_t\}$$

be a Gröbner basis for an ideal $I \subset k[x_1, \dots, x_n]$, and let $f \in k[x_1, \dots, x_n]$. Then there exists a unique polynomial $r \in k[x_1, \dots, x_n]$ satisfying the following properties:

- (i) *No term of r is divisible by any of $LT(g_1), \dots, LT(g_t)$.*
- (ii) *There exists $g \in I$ such that $f = g + r$.*

In particular, r is the remainder on division of f by G , regardless of the order in which the elements of G are listed in the division algorithm.

Proof. The division algorithm gives $f = a_1g_1 + \dots + a_tg_t + r$, where r satisfies (i). We can also satisfy (ii) by setting $g = a_1g_1 + \dots + a_tg_t \in I$. This proves the existence of r . To prove uniqueness, suppose that $f = g + r = g' + r'$ where both r and r' satisfy (i) and (ii). Then $r - r' = g' - g \in I$. If $r \neq r'$, then $LT(r - r') \in LT(I) \subset \langle LT(I) \rangle = \langle LT(g_1), \dots, LT(g_t) \rangle$. By Lemma 2.3.3, it follows that $LT(r - r')$ is divisible by some $LT(g_i)$. This is impossible, since no term of either r or r' is divisible by any of $LT(g_1), \dots, LT(g_t)$. Thus, $r - r' = 0$, and hence $r = r'$. Therefore, uniqueness is proved.

The final part of the proposition follows from the uniqueness of r . □

Corollary 2.5.2 ([1, Corollary 2, Section 6, p. 84]). *Let $G = \{g_1, \dots, g_t\}$ be a Gröbner basis for an ideal $I \subset k[x_1, \dots, x_n]$, and let $f \in k[x_1, \dots, x_n]$. Then $f \in I$ if and only if the remainder on division of f by G is zero.*

Proof. If the remainder on division of f by G is zero, then by the division algorithm we have $f = a_1g_1 + \dots + a_tg_t$ for some $a_1, \dots, a_t \in k[x_1, \dots, x_n]$, and hence $f \in I$.

Conversely, suppose that $f \in I$. Then $f = f + 0$ satisfies the two conditions of previous proposition. Therefore, by uniqueness of the remainder, it follows that 0 is the remainder on division of f by G . □

Proposition 2.5.3 ([1, Proposition 3, Section 6, p. 84]). *Let $I \subset k[x_1, \dots, x_n]$ be an ideal with Gröbner basis G . Suppose $f, g \in k[x_1, \dots, x_n]$. Then*

$$f \equiv g \pmod{I} \quad \text{if and only if} \quad \bar{f}^G = \bar{g}^G.$$

Therefore, $\{\bar{f}^G \mid f \in k[x_1, \dots, x_n]\}$ is a set of coset representatives for $k[x_1, \dots, x_n]/I$. Moreover, the map

$$N_G : k[x_1, \dots, x_n] \longrightarrow k[x_1, \dots, x_n]$$

$$f \longmapsto \bar{f}^G$$

is k -linear.

Proof. Consider the map

$$N_G : k[x_1, \dots, x_n] \longrightarrow k[x_1, \dots, x_n]$$

$$f \longmapsto \bar{f}^G.$$

By the Division Algorithm (Theorem 2.2.2), there exists $q \in I$ such that $f = q + N_G(f)$, so that $f - N_G(f) \in I$. Thus, $f + I = N_G(f) + I$ in $k[x_1, \dots, x_n]/I$.

Also, for any $c_1, c_2 \in k$ and any $f_1, f_2 \in k[x_1, \dots, x_n]$, $c_1 f_1 + c_2 f_2 - (c_1 N_G(f_1) + c_2 N_G(f_2)) \in I$, and $c_1 N_G(f_1) + c_2 N_G(f_2)$ is reduced with respect to G . Therefore, $N_G(c_1 f_1 + c_2 f_2) = c_1 N_G(f_1) + c_2 N_G(f_2)$, and hence the map N_G is k -linear.

Now,

$$f \equiv g \pmod{I} \iff \text{there exists } q \in I \text{ such that } f = q + g.$$

Thus, $N_G(f) = N_G(q) + N_G(g)$. But $N_G(q) = 0$, since $q \in I$ by Corollary 2.5.2. Hence, $N_G(f) = N_G(g)$. Conversely, suppose that $N_G(f) = N_G(g)$. Then $f - g = (f - N_G(f)) - (g - N_G(g)) \in I$, by Proposition 2.5.1(ii). Hence, $f \equiv g \pmod{I}$. $\square$

Using Corollary 2.5.2, we obtain an algorithm for solving the ideal membership problem, provided that a Gröbner basis G for the ideal is known. Indeed, to determine whether $f \in I$, it suffices to compute the remainder of f upon division by G .

In the subsequent discussion, we will explain how to compute Gröbner bases and thereby obtain a complete solution to the ideal membership problem. We will use the following notation for the remainder.

Definition 2.5.4. Let $F = (f_1, \dots, f_s)$ be an ordered s -tuple of polynomials. We write $\overline{f}^F$ for the remainder on division of f by F . If F is a Gröbner basis for $\langle f_1, \dots, f_s \rangle$, then by Proposition 2.5.1, we may regard F as a set (without any particular ordering).

For instance, let

$$F = (x^2y - y^2, x^4y^2 - y^2) \subset k[x, y],$$

and use the lexicographic order. Then

$$\overline{x^5y}^F = xy^3,$$

since the division algorithm yields

$$x^5y = (x^3 + xy)(x^2y - y^2) + 0 \cdot (x^4y^2 - y^2) + xy^3.$$

Definition 2.5.5. Let $f, g \in k[x_1, \dots, x_n]$ be nonzero polynomials.

(i) If

$$\text{multideg}(f) = \alpha \quad \text{and} \quad \text{multideg}(g) = \beta,$$

then let $\gamma = (\gamma_1, \dots, \gamma_n)$, where

$$\gamma_i = \max(\alpha_i, \beta_i) \quad \text{for each } i.$$

We call x^γ the least common multiple of $LM(f)$ and $LM(g)$, written as

$$x^\gamma = \text{LCM}(LM(f), LM(g)).$$

(ii) The S -polynomial of f and g is defined by

$$S(f, g) = \frac{x^\gamma}{LT(f)} f - \frac{x^\gamma}{LT(g)} g.$$

Example 2.5.6. Let $f = x^3y^2 - x^2y^3 + x$ and $g = 3x^4y + y^2$ in $R[x, y]$, and use the graded lexicographic order. Then $\gamma = (4, 2)$, and

$$S(f, g) = \frac{x^4y^2}{x^3y^2} f - \frac{x^4y^2}{3x^4y} g.$$

Hence,

$$S(f, g) = x f - \frac{1}{3} y g$$

and therefore

$$S(f, g) = -x^2y^3 + x^2 - \frac{1}{3}y^3.$$

Consider another example with $f = 2yx - x$ and $g = 3y^2 - x$ in $R[x, y]$. Suppose $>$ is the graded lexicographic order with $y > x$. Then $LT(f) = 2yx$ and $LT(g) = 3y^2$, with $\gamma = (1, 2)$. Hence, $\text{LCM}(LM(f), LM(g)) = xy^2$.

Therefore,

$$S(f, g) = \frac{xy^2}{2yx} f - \frac{xy^2}{3y^2} g.$$

Thus,

$$S(f, g) = \frac{y}{2} f - \frac{x}{3} g,$$

and hence

$$S(f, g) = -\frac{1}{2}xy + \frac{1}{3}x^2.$$

Lemma 2.5.7 ([1] Lemma 5, Section 6, p. 87]. Suppose we have a sum $\sum_{i=1}^s c_i f_i$, where $c_i \in k$ for all i , and $\text{multideg}(f_i) = \delta \in \mathbb{Z}_{\geq 0}^n$ for all i . If $\text{multideg}(\sum_{i=1}^s c_i f_i) < \delta$, then $\sum_{i=1}^s c_i f_i$ is a k -linear combination of the S -polynomials $S(f_j, f_k)$, $1 \leq j, k \leq s$.

Furthermore, each $S(f_j, f_k)$ has multidegree strictly less than δ .

Proof. Let $d_i = LC(f_i)$, so that $c_i d_i$ is the leading coefficient of $c_i f_i$. Since the $c_i f_i$ all have multidegree δ and their sum has strictly smaller multidegree, it follows that

$$\sum_{i=1}^s c_i d_i = 0.$$

Define $p_i = f_i/d_i$, and note that p_i has leading coefficient 1. Consider the telescoping sum

$$\begin{aligned} \sum_{i=1}^s c_i f_i &= \sum_{i=1}^s c_i d_i p_i \\ &= c_1 d_1 (p_1 - p_2) + (c_1 d_1 + c_2 d_2) (p_2 - p_3) + \cdots + (c_1 d_1 + \cdots + c_{s-1} d_{s-1}) (p_{s-1} - p_s) + (c_1 d_1 + \cdots + c_s d_s) p_s. \end{aligned}$$

Using $\sum_{i=1}^s c_i d_i = 0$, the last term vanishes, and hence

$$\sum_{i=1}^s c_i f_i = c_1 d_1 (p_1 - p_2) + (c_1 d_1 + c_2 d_2) (p_2 - p_3) + \cdots + (c_1 d_1 + \cdots + c_{s-1} d_{s-1}) (p_{s-1} - p_s).$$

Assume that $LT(f_i) = d_i x^\delta$, which implies that the least common multiple of $LM(f_j)$ and $LM(f_k)$ is x^δ . Then

$$S(f_j, f_k) = \frac{x^\delta}{LT(f_j)} f_j - \frac{x^\delta}{LT(f_k)} f_k = \frac{x^\delta}{d_j x^\delta} f_j - \frac{x^\delta}{d_k x^\delta} f_k = p_j - p_k. \quad (2.1)$$

Thus each difference $p_j - p_k$ is an S -polynomial, and so

$$\sum_{i=1}^s c_i f_i$$

is a k -linear combination of $S(f_j, f_k)$'s.

Since p_j and p_k have multidegree δ and leading coefficient 1, the difference $p_j - p_k$ has multidegree strictly less than δ . By Equation (1.1), the same is true of $S(f_j, f_k)$, and the lemma is proved. $\square$

When $f_1, \dots, f_s$ satisfy the hypothesis of Lemma 2.5.7, we obtain an equation

$$\sum_{i=1}^s c_i f_i = \sum_{j,k} c_{jk} S(f_j, f_k).$$

In the sum on the right-hand side, each summand $c_{jk} S(f_j, f_k)$ has multidegree strictly less than δ .

Theorem 2.5.8 (Buchberger's Criterion [11]). *Let I be a polynomial ideal. Then a basis $G = \{g_1, \dots, g_t\}$ for I is a Gröbner basis for I if and only if, for all pairs $i \neq j$, the remainder on division of $S(g_i, g_j)$ by G (listed in some order) is zero.*

Proof. ($\Rightarrow$) If G is a Gröbner basis, then since $S(g_i, g_j) \in I$, the remainder on division by G is zero by Corollary 2.5.2.

($\Leftarrow$) Let $f \in I$ be a nonzero polynomial. We must show that if the S -polynomials all have zero remainders on division by G , then

$$LT(f) \in \langle LT(g_1), \dots, LT(g_t) \rangle.$$

Before giving the details, let us outline the strategy of the proof.

Given

$$f \in I = \langle g_1, \dots, g_t \rangle,$$

there are polynomials $h_i \in k[x_1, \dots, x_n]$ such that

$$f = \sum_{i=1}^t h_i g_i. \quad (2.2)$$

By the property of multidegree, we can say

$$\text{multideg}(f) \leq \max(\text{multideg}(h_i g_i)). \quad (2.3)$$

If equality does not occur, then some cancellation must occur among the leading terms of (2.2). Lemma 2.5.7 will enable us to rewrite this in terms of S -polynomials. Then our assumption that S -polynomials have zero remainders will allow us to replace the S -polynomials by expressions that involve less cancellation. Thus, we will get an expression for f that has less cancellation of leading terms. Continuing in this way, we will eventually find an expression (2.2) for f where equality occurs in (2.3). Then

$$\text{multideg}(f) = \max(\text{multideg}(h_i g_i))$$

for some i , and it will follow that $LT(f)$ is divisible by $LT(g_i)$. This will show that

$$LT(f) \in \langle LT(g_1), \dots, LT(g_t) \rangle,$$

which is what we want to prove.

We now give the details of the proof. Given an expression (2.2) for f , let

$$m(i) = \text{multideg}(h_i g_i),$$

and define

$$\delta = \max(m(1), \dots, m(t)).$$

Then inequality (2.3) becomes

$$\text{multideg}(f) \leq \delta.$$

Now consider all possible ways that f can be written in the form (2.2). For each such expression, we get a possibly different δ . Since a monomial order is a well-ordering, we can select an expression (2.2) for f such that δ is minimal.

We will show that once this minimal δ is chosen, we have

$$\text{multideg}(f) = \delta.$$

Then equality occurs in (2.3), and as we observed, it follows that

$$LT(f) \in \langle LT(g_1), \dots, LT(g_t) \rangle.$$

It remains to show $\text{multideg}(f) = \delta$. We will prove this by contradiction. Equality can fail only when $\text{multideg}(f) < \delta$. To isolate the terms of multidegree δ , let us write f in the following form:

$$\begin{aligned} f &= \sum_{m(i)=\delta} h_i g_i + \sum_{m(i)<\delta} h_i g_i \\ &= \sum_{m(i)=\delta} LT(h_i) g_i + \sum_{m(i)=\delta} (h_i - LT(h_i)) g_i + \sum_{m(i)<\delta} h_i g_i. \end{aligned} \tag{2.4}$$

The monomials appearing in the second and third sums on the second line all have multidegree $< \delta$. Thus, the assumption $\text{multideg}(f) < \delta$ means that the first sum also has multidegree $< \delta$:

$$LT(h_i) = c_i x^{\alpha(i)}.$$

Then the first sum

$$\sum_{m(i)=\delta} LT(h_i) g_i = \sum_{m(i)=\delta} c_i x^{\alpha(i)} g_i$$

has exactly the form described in Lemma 2.5.7. Thus Lemma 2.5.7 implies that this sum is a linear combination of the S -polynomials

$$S(x^{\alpha(i)} g_i, x^{\alpha(k)} g_k).$$

However,

$$S(x^{\alpha(i)} g_i, x^{\alpha(k)} g_k) = \frac{x^\delta}{x^{\alpha(i)} LT(g_i)} x^{\alpha(i)} g_i - \frac{x^\delta}{x^{\alpha(k)} LT(g_k)} x^{\alpha(k)} g_k = x^{\delta - \gamma_{jk}} S(g_j, g_k),$$

where $x^{\gamma_{jk}} = \text{LCM}(LM(g_j), LM(g_k))$. Thus there are constants $c_{jk} \in k$ such that

$$\sum_{m(i)=\delta} LT(h_i)g_i = \sum_{j,k} c_{jk} S(x^{\alpha(j)}g_j, x^{\alpha(k)}g_k) = \sum_{j,k} c_{jk} x^{\delta-\gamma_{jk}} S(g_j, g_k). \quad (2.5)$$

The next step is to use our hypothesis that the remainder of $S(g_j, g_k)$ on division by $g_1, \dots, g_t$ is zero. Using the division algorithm, this means that each S -polynomial can be written in the form

$$S(g_j, g_k) = \sum_{i=1}^t a_{ijk} g_i \quad (2.6)$$

where $a_{ijk} \in k[x_1, \dots, x_n]$. The division algorithm also tells us that

$$\text{multideg}(a_{ijk}g_i) \leq \text{multideg}(S(g_j, g_k)) \quad (2.7)$$

for all i, j, k .

Multiply the expression for $S(g_j, g_k)$ by $x^{\delta-\gamma_{jk}}$ to obtain

$$x^{\delta-\gamma_{jk}} S(g_j, g_k) = \sum_{i=1}^t b_{ijk} g_i,$$

where

$$b_{ijk} = x^{\delta-\gamma_{jk}} a_{ijk}.$$

Then (2.7) and Lemma 2.5.7 together imply that

$$\text{multideg}(b_{ijk}g_i) \leq \text{multideg}(x^{\delta-\gamma_{jk}} S(g_j, g_k)) < \delta. \quad (2.8)$$

Substituting the above expression into (2.5), we get

$$\sum_{m(i)=\delta} LT(h_i)g_i = \sum_{j,k} c_{jk} x^{\delta-\gamma_{jk}} S(g_j, g_k) = \sum_{j,k} c_{jk} \left(\sum_{i=1}^t b_{ijk} g_i \right) = \sum_{i=1}^t \tilde{h}_i g_i,$$

where by (2.8) we have

$$\text{multideg}(\tilde{h}_i g_i) < \delta$$

for all i .

For the final step in the proof, substitute

$$\sum_{m(i)=\delta} LT(h_i)g_i = \sum_{i=1}^t \tilde{h}_i g_i$$

into equation (2.4) to obtain an expression for f as a polynomial combination of the g_i 's where all terms have multidegree $< \delta$. This contradicts the minimality of δ and completes the proof of the theorem. $\square$

So far, it has been difficult to determine whether a basis of an ideal is a Gröbner basis. Using the S -pair criterion, however, it is now easy to determine whether a given basis is a Gröbner basis.

Example 2.5.9. As an example of how to use Buchberger's Criterion, consider the ideal $I = \langle y - x^2, z - x^3 \rangle$ of the twisted cubic in R^3 . We claim that $G = \{y - x^2, z - x^3\}$ is a Gröbner basis for the lexicographic order with $y > z > x$.

To prove this, consider the S -polynomial

$$S(y - x^2, z - x^3) = \frac{yz}{y}(y - x^2) - \frac{yz}{z}(z - x^3) = -zx^2 + yx^3.$$

Using the division algorithm, we find

$$-zx^2 + yx^3 = x^3(y - x^2) + (-x^2)(z - x^3) + 0,$$

so that

$$\overline{S(y - x^2, z - x^3)}^G = 0.$$

Thus, by Theorem 2.5.8, G is a Gröbner basis for I .

Lemma 2.5.10 ([1] Lemma 3, Section 7, p. 92]. *Let G be a Gröbner basis for an ideal I . Let $p \in G$ be a polynomial such that $\text{LT}(p) \in \text{LT}(G \setminus \{p\})$. Then $G \setminus \{p\}$ is also a Gröbner basis for I .*

Proof. We know that $\text{LT}(G) = \text{LT}(I)$. Since $\text{LT}(p) \in \text{LT}(G \setminus \{p\})$, removing p does not change the leading term ideal generated by G . Hence,

$$\text{LT}(G \setminus \{p\}) = \text{LT}(G) = \text{LT}(I).$$

By definition of a Gröbner basis, a set of polynomials whose leading term ideal equals $\text{LT}(I)$ is a Gröbner basis for I . Therefore, $G \setminus \{p\}$ is also a Gröbner basis for I . $\square$

Definition 2.5.11. A *minimal Gröbner basis* for a polynomial ideal I is a Gröbner basis G for I such that:

- (i) $\text{LC}(p) = 1$ for all $p \in G$.
- (ii) For all $p \in G$, $\text{LT}(p) \notin \langle \text{LT}(G \setminus \{p\}) \rangle$.

Definition 2.5.12. A *reduced Gröbner basis* for a polynomial ideal I is a Gröbner basis G for I such that:

- (i) $\text{LC}(p) = 1$ for all $p \in G$.
- (ii) For all $p \in G$, no monomial of p lies in $\langle \text{LT}(G \setminus \{p\}) \rangle$.

Proposition 2.5.13 ([1] Theorem 5, Section 7, p. 93]. *Let $I \neq \{0\}$ be a polynomial ideal. Then, for a given monomial ordering, I has a unique reduced Gröbner basis.*

Proof. Let G be a minimal Gröbner basis for I . We say that $g \in G$ is *reduced for G* provided that no monomial of g lies in $\langle \text{LT}(G \setminus \{g\}) \rangle$. Our goal is to modify G until all of its elements are reduced.

A first observation is that if g is reduced for G , then g is also reduced for any other minimal Gröbner basis of I that contains g and has the same set of leading terms. This follows because the notion of reduced depends only on leading terms.

Next, given $g \in G$, let $\bar{g} = g^{G \setminus \{g\}}$ be the remainder when g is divided by $G \setminus \{g\}$, and set

$$\bar{G} = (G \setminus \{g\}) \cup \{\bar{g}\}.$$

We claim that $\bar{G}$ is a minimal Gröbner basis for I . To see this, note that

$$\text{LT}(\bar{g}) = \text{LT}(g),$$

since when we divide g by $G \setminus \{g\}$, the leading term cannot be cancelled and hence remains unchanged in the remainder.

Thus,

$$\text{LT}(\bar{G}) = \text{LT}(G),$$

so $\bar{G}$ is also a Gröbner basis for I . Minimality follows since leading terms are preserved. Moreover, by construction, $\bar{g}$ is reduced for $\bar{G}$.

Now we repeat this process for each element of G . The basis may change at each step, but once an element becomes reduced, it remains reduced because we never change leading terms during the process. Hence, we eventually obtain a reduced Gröbner basis.

Finally, we prove uniqueness. Suppose G and $\tilde{G}$ are reduced Gröbner bases for I . Then both are minimal, hence

$$\text{LT}(G) = \text{LT}(\tilde{G}).$$

Let $g \in G$. Then there exists $\tilde{g} \in \tilde{G}$ such that

$$\text{LT}(g) = \text{LT}(\tilde{g}).$$

We claim that $g = \tilde{g}$. Consider $g - \tilde{g} \in I$. Since G is a Gröbner basis,

$$(g - \tilde{g})_G = 0.$$

But the leading terms cancel because $\text{LT}(g) = \text{LT}(\tilde{g})$. Since both bases are reduced, no remaining monomial is divisible by any element of $\text{LT}(G) = \text{LT}(\tilde{G})$. Hence,

$$(g - \tilde{g})_G = g - \tilde{g},$$

which implies $g - \tilde{g} = 0$. Therefore $g = \tilde{g}$, and so $G = \tilde{G}$.

This completes the proof. □

Chapter 3

Hilbert's Nullstellensatz, Elimination and the Eigenvector Theorems

In this chapter, we study some fundamental applications of Gröbner basis theory to algebraic geometry and the solution of polynomial systems. We begin with Hilbert's Nullstellensatz, which establishes a deep connection between ideals in polynomial rings and algebraic varieties. This correspondence provides the algebraic foundation for the geometric study of polynomial equations. We then discuss elimination theory, where Gröbner bases with respect to elimination orderings are used to eliminate variables and analyze projections of algebraic varieties. Finally, we introduce the eigenvector theorems, which give an effective method for computing the points of zero-dimensional varieties via linear algebra on quotient rings.

3.1 Hilbert's Nullstellensatz

Hilbert's Nullstellensatz is a fundamental result in algebraic geometry that connects ideals in polynomial rings with algebraic varieties. In this section, we present its weak and strong forms together with some important consequences. Throughout this section, let k be an uncountable algebraically closed field.

We begin with two preliminary lemmas that will be used in the proof of the Weak Hilbert's Nullstellensatz.

Lemma 3.1.1. *Let V be a vector space and let W be a subspace of V which is the linear span of a countable set of vectors $\{v_1, v_2, \dots\}$. Then any linearly independent subset of W is either finite or countable.*

Proof. For each $n \geq 1$, define $W_n = \text{Span}\{v_1, v_2, \dots, v_n\}$. Then each W_n is a finite-dimensional subspace of W with $\dim(W_n) \leq n$, and since W is spanned by the countable set $\{v_1, v_2, \dots\}$, we have

$$W = \bigcup_{n=1}^{\infty} W_n.$$

Let $S \subseteq W$ be a linearly independent set. Since $S \subseteq W$, it follows that

$$S = S \cap W = S \cap \left(\bigcup_{n=1}^{\infty} W_n \right) = \bigcup_{n=1}^{\infty} (S \cap W_n).$$

For each n , the set $S \cap W_n$ is a linearly independent subset of the finite-dimensional vector space W_n . Hence,

$$|S \cap W_n| \leq \dim(W_n) \leq n,$$

so each $S \cap W_n$ is finite.

Therefore, S is a countable union of finite sets. Consequently, S is either finite or countable. This completes the proof. $\square$

Lemma 3.1.2. *The set $\left\{ \frac{1}{x-a} \mid a \in k \right\} \subset k(x)$ is linearly independent over k .*

Proof. Suppose that

$$\frac{b_1}{x-a_1} + \frac{b_2}{x-a_2} + \cdots + \frac{b_n}{x-a_n} = 0,$$

where $b_1, b_2, \dots, b_n \in k$, the scalars $a_1, a_2, \dots, a_n \in k$ are distinct, and the equality holds in $k(x)$.

Multiplying both sides by the common denominator

$$(x-a_1)(x-a_2) \cdots (x-a_n),$$

we obtain

$$\sum_{i=1}^n b_i \prod_{\substack{1 \leq j \leq n \\ j \neq i}} (x-a_j) = 0.$$

Fix $i \in \{1, 2, \dots, n\}$ and substitute $x = a_i$. Since the a_j 's are distinct, every term in the above sum vanishes except the i -th term. Thus,

$$b_i \prod_{\substack{1 \leq j \leq n \\ j \neq i}} (a_i - a_j) = 0.$$

Since $a_i \neq a_j$ for $j \neq i$, the product is nonzero, and hence $b_i = 0$.

As this holds for every $i = 1, 2, \dots, n$, we conclude that

$$b_1 = b_2 = \cdots = b_n = 0.$$

Therefore,

$$\left\{ \frac{1}{x-a} \mid a \in k \right\}$$

is linearly independent over k . $\square$

Theorem 3.1.3 (Hilbert's Nullstellensatz: Weak Form [10, pp. 320–324]). *The maximal ideals in the polynomial ring $k[x_1, \dots, x_n]$ are precisely the ideals of the form*

$$\mathfrak{m}_a = \langle x_1 - a_1, x_2 - a_2, \dots, x_n - a_n \rangle,$$

where $a = (a_1, a_2, \dots, a_n) \in k^n$.

Proof. Let $R = k[x_1, x_2, \dots, x_n]$ and let M be a maximal ideal of R . Define

$$N = M \cap k[x_1].$$

We claim that N is a maximal ideal of $k[x_1]$. Consider the natural homomorphism

$$\pi : k[x_1] \rightarrow R/M, \quad f(x_1) \mapsto f(x_1) + M.$$

Then $\ker(\pi) = N$, so N is a prime ideal of $k[x_1]$.

Since $k[x_1]$ is a principal ideal domain, every nonzero prime ideal is maximal. Hence either $N = (0)$ or $N = (x_1 - a_1)$ for some $a_1 \in k$.

Case 1: $N \neq (0)$.

Then $N = (x_1 - a_1)$ for some $a_1 \in k$, and hence $x_1 - a_1 \in M$.

Case 2: $N = (0)$.

Then π is injective, so we obtain an embedding

$$k[x_1] \hookrightarrow R/M.$$

Since R/M is a field, this extends to an embedding of fields

$$k(x_1) \hookrightarrow R/M.$$

Now R/M is a vector space over k of countable dimension, hence $k(x_1)$ must also be countable-dimensional over k . However, the set $\left\{ \frac{1}{x_1 - a} \mid a \in k \right\}$ is an uncountable linearly independent subset of $k(x_1)$, which is a contradiction. Therefore, $N \neq (0)$.

Hence $N = (x_1 - a_1)$ for some $a_1 \in k$, so $x_1 - a_1 \in M$.

Repeating the same argument for each variable $x_2, \dots, x_n$, we obtain $a_2, \dots, a_n \in k$ such that

$$x_i - a_i \in M \quad \text{for } i = 1, 2, \dots, n.$$

Thus,

$$\langle x_1 - a_1, x_2 - a_2, \dots, x_n - a_n \rangle \subseteq M.$$

Since this ideal is maximal, we conclude

$$M = \langle x_1 - a_1, x_2 - a_2, \dots, x_n - a_n \rangle = \mathfrak{m}_a.$$

this completes the proof. □

Corollary 3.1.4. *Let $f_1, f_2, \dots, f_t \in R = k[x_1, \dots, x_n]$. Let $V \subset k^n$ be the variety defined by $f_1, f_2, \dots, f_t$, and let $I = \langle f_1, f_2, \dots, f_t \rangle$. Then there is a one-to-one correspondence between V and the set of maximal ideals of R/I .*

Proof. By the Weak Hilbert's Nullstellensatz, maximal ideals of R are of the form

$$\mathfrak{m}_a = \langle x_1 - a_1, x_2 - a_2, \dots, x_n - a_n \rangle$$

for $a = (a_1, \dots, a_n) \in k^n$.

A point $a \in k^n$ lies in V if and only if $f_i(a) = 0$ for all i , which is equivalent to

$$I \subseteq \mathfrak{m}_a.$$

Thus, maximal ideals of R containing I correspond exactly to points of V .

Since maximal ideals of R/I are in one-to-one correspondence with maximal ideals of R containing I , we obtain a bijection between V and the maximal ideals of R/I . □

Corollary 3.1.5. *Let I be an ideal of $R = k[x_1, x_2, \dots, x_n]$. Then $V(I) = \emptyset$ if and only if $I = R$.*

Proof. First suppose that $V(I) = \emptyset$. We prove that $I = R$. Assume, for contradiction, that $I \neq R$. Then I is a proper ideal and is contained in some maximal ideal $\mathfrak{m} \subset R$.

By the Weak Hilbert's Nullstellensatz, every maximal ideal of R is of the form

$$\mathfrak{m} = \langle x_1 - a_1, x_2 - a_2, \dots, x_n - a_n \rangle$$

for some $a = (a_1, a_2, \dots, a_n) \in k^n$.

Since $I \subseteq \mathfrak{m}$, every $f \in I$ vanishes at a , i.e., $f(a) = 0$. Hence $a \in V(I)$, contradicting the assumption that $V(I) = \emptyset$. Therefore, $I = R$.

Conversely, suppose $I = R$. Then $1 \in I$. If $a \in V(I)$, then every polynomial in I vanishes at a , in particular $1(a) = 0$, which is impossible. Hence $V(I) = \emptyset$. $\square$

The Weak Hilbert's Nullstellensatz gives a correspondence between maximal ideals of $k[x_1, \dots, x_n]$ and points in k^n . To extend this correspondence from points to general algebraic sets, we need to consider arbitrary ideals and their vanishing behavior. This leads naturally to the notion of radical ideals, which capture all polynomials vanishing on a given variety. Hence we now introduce radical ideals in order to prepare for the Strong Hilbert's Nullstellensatz.

Definition 3.1.6. Let I be an ideal of $R = k[x_1, x_2, \dots, x_n]$. The *radical* of I , denoted by $\sqrt{I}$, is defined as

$$\sqrt{I} = \{f \in R \mid f^m \in I \text{ for some integer } m \geq 1\}.$$

If $I = \sqrt{I}$, then I is called a *radical ideal*.

Example 3.1.7. Let $R = k[x]$ and let $I = \langle x^2 \rangle$. Then the radical of I is

$$\sqrt{I} = \langle x \rangle.$$

Indeed, if $f(x) \in k[x]$ satisfies $f(x)^m \in \langle x^2 \rangle$ for some $m \geq 1$, then $f(x)$ must be divisible by x . Hence $\sqrt{I} = \langle x \rangle$, which is a radical ideal.

Example 3.1.8. Let $R = k[x, y]$ and consider the ideal

$$I = \langle x^2, y^3 \rangle.$$

Then its radical is

$$\sqrt{I} = \langle x, y \rangle.$$

Indeed, if $f(x, y) \in R$ satisfies $f(x, y)^m \in \langle x^2, y^3 \rangle$ for some $m \geq 1$, then $f(x, y)$ must vanish at $(0, 0)$. Hence $f \in \langle x, y \rangle$, and therefore $\sqrt{I} = \langle x, y \rangle$.

Proposition 3.1.9. Let I, J be ideals of $R = k[x_1, \dots, x_n]$. Then:

- (i) $I \subseteq \sqrt{I}$,
- (ii) $\sqrt{I}$ is an ideal of R ,
- (iii) $\sqrt{\sqrt{I}} = \sqrt{I}$,
- (iv) If $I \subseteq J$, then $\sqrt{I} \subseteq \sqrt{J}$.

Proof. (i) If $f \in I$, then $f^1 \in I$, hence $f \in \sqrt{I}$. Thus $I \subseteq \sqrt{I}$.

- (ii) Let $f, g \in \sqrt{I}$, so $f^m \in I$ and $g^n \in I$ for some $m, n \geq 1$. Then $(f + g)^{m+n} \in I$, hence $f + g \in \sqrt{I}$. Also, for $h \in R$, we have $(hf)^m = h^m f^m \in I$, so $hf \in \sqrt{I}$. Thus $\sqrt{I}$ is an ideal.
- (iii) From (i), $I \subseteq \sqrt{I}$ implies $\sqrt{I} \subseteq \sqrt{\sqrt{I}}$. Conversely, if $f \in \sqrt{\sqrt{I}}$, then $f^m \in \sqrt{I}$ for some m , hence $(f^m)^n = f^{mn} \in I$ for some n , so $f \in \sqrt{I}$. Thus equality holds.
- (iv) If $I \subseteq J$ and $f^m \in I$, then $f^m \in J$, hence $f \in \sqrt{J}$. Therefore $\sqrt{I} \subseteq \sqrt{J}$.

□

Theorem 3.1.10 (Hilbert's Nullstellensatz: Strong Form [10, pp. 320–324]). *Let J be an ideal of $R = k[x_1, \dots, x_n]$ with k as an algebraically closed field. Then*

$$I(V(J)) = \sqrt{J}.$$

Proof. The inclusion $J \subseteq I(V(J))$ is clear from Lemma 1.3.12, since every polynomial in J vanishes on $V(J)$. To prove the reverse inclusion, let $0 \neq g \in I(V(J))$. We show that $g \in \sqrt{J}$. Let $J = \langle f_1, f_2, \dots, f_m \rangle$, and introduce a new variable t . Consider the ideal

$$I = \langle f_1, f_2, \dots, f_m, tg - 1 \rangle \subseteq R[t].$$

Suppose $b = (b_1, b_2, \dots, b_n, b_{n+1}) \in V(I)$. Then $f_i(b_1, \dots, b_n) = 0$ for all $i = 1, \dots, m$, so $(b_1, \dots, b_n) \in V(J)$. Since $g \in I(V(J))$, we get $g(b_1, \dots, b_n) = 0$.

Also, since $b \in V(I)$,

$$(tg - 1)(b) = b_{n+1}g(b_1, \dots, b_n) - 1 = 0,$$

which gives a contradiction $-1 = 0$. Hence $V(I) = \emptyset$.

By Corollary 3.1.5, $I = R[t]$. Therefore,

$$1 = g_1 f_1 + g_2 f_2 + \dots + g_m f_m + g_{m+1}(tg - 1),$$

for some $g_1, g_2, \dots, g_{m+1} \in R[t]$.

Now substitute $t = \frac{1}{g}$. After clearing denominators, we obtain that there exists $r \geq 1$ such that $g^r \in J$. Hence $g \in \sqrt{J}$.

Thus,

$$I(V(J)) \subseteq \sqrt{J}.$$

Combining both inclusions, we conclude

$$I(V(J)) = \sqrt{J}.$$

this completes the proof. □

Theorem 3.1.11 (Radical Membership Problem [1, Proposition 8, Section 2, p. 184]). *Let $I = \langle f_1, \dots, f_s \rangle$ be an ideal in $k[x_1, \dots, x_n]$. Then $f \in \sqrt{I}$ if and only if*

$$1 \in \langle f_1, \dots, f_s, 1 - wf \rangle \subset k[x_1, \dots, x_n, w],$$

where w is a new indeterminate.

Proof. We have $\sqrt{I} = I(V(I))$, and hence $f \in \sqrt{I}$ if and only if $f(a_1, \dots, a_n) = 0$ for all $(a_1, \dots, a_n) \in V(I)$.

Suppose $f \in \sqrt{I}$. If $(a_1, \dots, a_n, b) \in V(f_1, \dots, f_s, 1 - wf)$, then

$$f_i(a_1, \dots, a_n) = 0 \text{ for all } i = 1, \dots, s,$$

and

$$1 - bf(a_1, \dots, a_n) = 0.$$

Thus $(a_1, \dots, a_n) \in V(I)$, so $f(a_1, \dots, a_n) = 0$, which implies a contradiction in the equation $1 - bf(a_1, \dots, a_n) = 0$. Therefore,

$$V(f_1, \dots, f_s, 1 - wf) = \emptyset.$$

By Corollary 3.1.5, this implies

$$1 \in \langle f_1, \dots, f_s, 1 - wf \rangle.$$

Conversely, suppose

$$1 \in \langle f_1, \dots, f_s, 1 - wf \rangle.$$

Then there exist $h_1, \dots, h_s, h \in k[x_1, \dots, x_n, w]$ such that

$$1 = \sum_{i=1}^s h_i f_i + h(1 - wf).$$

For every $(a_1, \dots, a_n) \in V(I)$, we obtain

$$1 = (1 - wf(a_1, \dots, a_n)) h(a_1, \dots, a_n, w).$$

If $f(a_1, \dots, a_n) \neq 0$, then substituting $w = \frac{1}{f(a_1, \dots, a_n)}$ gives a contradiction. Hence $f(a_1, \dots, a_n) = 0$ for all $(a_1, \dots, a_n) \in V(I)$, so $f \in \sqrt{I}$. $\square$

The radical membership problem can be reduced to an ideal membership problem. To determine whether $f \in \sqrt{I}$, we consider the ideal $\langle f_1, \dots, f_s, 1 - wf \rangle \subset k[x_1, \dots, x_n, w]$, where $I = \langle f_1, \dots, f_s \rangle$. We then compute a reduced Gröbner basis G of this ideal with respect to a fixed monomial order. If $1 \in G$, then the ideal generated is the whole ring, which implies $1 \in \langle f_1, \dots, f_s, 1 - wf \rangle$, and hence $f \in \sqrt{I}$. If $1 \notin G$, then the ideal is proper, and consequently $f \notin \sqrt{I}$. Thus, the question of radical membership is reduced to checking whether the reduced Gröbner basis of an associated ideal contains the constant polynomial 1.

Example 3.1.12. Consider the ideal $I = \langle x^2, y^3 \rangle$ in $R = \mathbb{Q}[x, y]$. We wish to determine whether the element $f = xy$ belongs to $\sqrt{I}$. We consider the ideal $J = \langle x^2, y^3, 1 - wxy \rangle \subset \mathbb{Q}[x, y, w]$. We compute the reduced Gröbner basis G for J with respect to a lexicographic term ordering with $x > y > w$ using SINGULAR.

```

> ring A = 0, (x,y,w), lp;
characteristic : 0
number of vars : 3

block 1 : ordering lp
: names x y w
block 2 : ordering C

> poly p1 = x2;
> poly p2 = y3;
> poly p3 = 1-wxy;

> ideal J = p1,p2,p3;
> J;
J[1]=x2
J[2]=y3
J[3]=1-wxy

> option(redSB);
> ideal G = std(J);
> G;
G[1]=1

```

Since the reduced Gröbner basis contains 1, we conclude that

$$1 \in J.$$

Hence $xy \in \sqrt{I}$.

3.2 Elimination Theory

Elimination theory is a key technique in computational algebraic geometry used to eliminate variables from systems of polynomial equations. It allows us to project algebraic varieties onto coordinate subspaces by removing selected variables. This is achieved effectively using Gröbner bases with respect to suitable monomial orders, especially lexicographic order.

To illustrate the basic idea of elimination theory, consider the system of polynomial equations

$$\begin{aligned}
 x^2 + y + z &= 1, \\
 x + y^2 + z &= 1, \\
 x + y + z^2 &= 1.
 \end{aligned}$$

Let

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

A Gröbner basis for I with respect to the lexicographic order $x > y > z$ is given by

$$\begin{aligned}
 g_1 &= x + y + z^2 - 1, \\
 g_2 &= y^2 - y - z^2 + z, \\
 g_3 &= 2yz^2 + z^4 - z^2, \\
 g_4 &= z^6 - 4z^4 + 4z^3 - z^2.
 \end{aligned}$$

Since Gröbner bases preserve the solution set, the above system has the same solutions as the system $g_1 = g_2 = g_3 = g_4 = 0$. Observe that

$$g_4 = z^6 - 4z^4 + 4z^3 - z^2 = z^2(z-1)^2(z^2 + 2z - 1),$$

which depends only on z . Thus, we have eliminated the variables x and y , reducing the original system to a single equation in one variable.

Solving $g_4 = 0$, we obtain the possible values

$$z = 0, 1, -1 + \sqrt{2}, -1 - \sqrt{2}.$$

Substituting these values into

$$g_2 = y^2 - y - z^2 + z = 0 \quad \text{and} \quad g_3 = 2yz^2 + z^4 - z^2 = 0,$$

we determine the corresponding values of y , and then use

$$g_1 = x + y + z^2 - 1 = 0$$

to determine x . Hence the system has exactly five solutions:

$$\begin{aligned} (1, 0, 0), \quad (0, 1, 0), \quad (0, 0, 1), \\ (-1 + \sqrt{2}, -1 + \sqrt{2}, -1 + \sqrt{2}), \\ (-1 - \sqrt{2}, -1 - \sqrt{2}, -1 - \sqrt{2}). \end{aligned}$$

This example demonstrates the central idea of elimination theory: by computing a Gröbner basis with respect to a suitable monomial order, one can eliminate variables from a polynomial system and reduce the problem to solving equations in fewer variables.

Definition 3.2.1. ([4]) Let $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_m)$ be two sets of variables with monomial orders $<_x$ and $<_y$, respectively. For monomials $X_1 = x^\alpha$, $X_2 = x^\beta$ and $Y_1 = y^\gamma$, $Y_2 = y^\delta$, define an order $<$ on $k[x_1, \dots, x_n, y_1, \dots, y_m]$ by

$$X_1 Y_1 < X_2 Y_2 \iff (X_1 <_x X_2) \text{ or } (X_1 = X_2 \text{ and } Y_1 <_y Y_2).$$

This is called an *elimination order*, where the x -variables are larger than the y -variables.

Example 3.2.2. Consider $K[x, y]$ with the lexicographic order $x > y$. Then

$$x^2 > xy > x > y^5 > y^2 > 1.$$

Every monomial containing x is larger than any monomial involving only y . Hence, this is an elimination order with x larger than y .

Example 3.2.3. Consider $K[x_1, x_2, y_1, y_2]$ with block order

$$(x_1, x_2) \gg (y_1, y_2).$$

For example,

$$x_1^2 y_2^5 > y_1^{100}.$$

Monomials involving x -variables dominate those involving only y -variables. Hence, this is an elimination order with the x -block larger than the y -block.

Definition 3.2.4. Let $I \subseteq k[x_1, \dots, x_n, y_1, \dots, y_m]$ be an ideal in a polynomial ring. For the set of variables $x = (x_1, \dots, x_n)$, the *x-elimination ideal* of I is defined as

$$I \cap k[y_1, \dots, y_m].$$

That is, it consists of all polynomials in I that involve only the y -variables.

Theorem 3.2.5 (Elimination Theorem [1, Theorem 2, Section 1, p. 122]). *Let I be a non-zero ideal of $k[x_1, \dots, x_n, y_1, \dots, y_m]$, and let $<$ be an elimination order with the x -variables larger than the y -variables. Let $G = \{g_1, \dots, g_t\}$ be a Gröbner basis for I . Then $G \cap k[y_1, \dots, y_m]$ is a Gröbner basis for the ideal $I \cap k[y_1, \dots, y_m]$.*

Proof. Clearly,

$$G \cap k[y_1, \dots, y_m] \subseteq I \cap k[y_1, \dots, y_m],$$

since $G \subseteq I$. Now let

$$0 \neq f(y_1, \dots, y_m) \in I \cap k[y_1, \dots, y_m].$$

Since G is a Gröbner basis for I , there exists some $g_i \in G$ such that

$$\text{LT}(g_i) \mid \text{LT}(f).$$

Because f involves only the y -variables, its leading term $\text{LT}(f)$ contains no x -variables. Hence $\text{LT}(g_i)$, being a divisor of $\text{LT}(f)$, also contains no x -variables.

Suppose g_i has a term involving an x -variable. Since the monomial order is an elimination order with the x -variables larger than the y -variables, every monomial involving an x -variable is larger than every monomial involving only y -variables. Therefore, the leading term of g_i must involve an x -variable, contradicting the fact that $\text{LT}(g_i)$ contains only y -variables.

Thus every term of g_i involves only the y -variables, and hence

$$g_i \in G \cap k[y_1, \dots, y_m].$$

Therefore, for every nonzero

$$f \in I \cap k[y_1, \dots, y_m],$$

there exists

$$g_i \in G \cap k[y_1, \dots, y_m]$$

such that

$$\text{LT}(g_i) \mid \text{LT}(f).$$

Hence $G \cap k[y_1, \dots, y_m]$ is a Gröbner basis for $I \cap k[y_1, \dots, y_m]$. this completes the proof. $\square$

Having established the theoretical foundation of elimination via Gröbner bases, we now consider some important applications of elimination theory in ideal computations. In particular, elimination provides effective methods for computing intersections of ideals, least common multiples of ideals, and colon ideals. These constructions illustrate the computational power of elimination and further demonstrate its usefulness in algorithmic commutative algebra.

Proposition 3.2.6 ([1 Theorem 11, Section 3, p. 194]). *Let I, J be ideals in $k[x_1, \dots, x_n]$, and let w be a new variable. Consider the ideal $\langle wI, (1-w)J \rangle \subseteq k[x_1, \dots, x_n, w]$. Then*

$$I \cap J = \langle wI, (1-w)J \rangle \cap k[x_1, \dots, x_n].$$

Proof. Let $f \in I \cap J$. Since $f \in I$ and $f \in J$, we may write

$$f = wf + (1-w)f.$$

Hence

$$f \in \langle wI, (1-w)J \rangle \cap k[x_1, \dots, x_n].$$

Conversely, suppose that

$$f \in \langle wI, (1-w)J \rangle \cap k[x_1, \dots, x_n].$$

Then

$$f = \sum_{i=1}^s wf_i h_i + \sum_{j=1}^p (1-w)g_j h_j,$$

where $f_i \in I$, $g_j \in J$, and $h_i, h_j \in k[x_1, \dots, x_n, w]$.

Since f does not involve the variable w , we may substitute $w = 1$ to obtain

$$f = \sum_{i=1}^s f_i h_i(x_1, \dots, x_n, 1) \in I.$$

Similarly, substituting $w = 0$ gives

$$f = \sum_{j=1}^p g_j h_j(x_1, \dots, x_n, 0) \in J.$$

Therefore,

$$f \in I \cap J.$$

Hence,

$$I \cap J = \langle wI, (1-w)J \rangle \cap k[x_1, \dots, x_n].$$

this completes the poof. □

The preceding proposition yields an effective procedure for computing generators of the ideal $I \cap J$. Namely, one first computes a Gröbner basis G for the ideal $\langle wI, (1-w)J \rangle \subseteq k[x_1, \dots, x_n, w]$ with respect to an elimination order in which w is larger than $x_1, \dots, x_n$. By the Elimination Theorem, the set

$$G \cap k[x_1, \dots, x_n]$$

forms a Gröbner basis for $I \cap J$, and hence provides generators for the intersection ideal. This method extends naturally to the computation of intersections of finitely many ideals.

Lemma 3.2.7 ([1 Proposition 13, Section 3, p. 195]). *If $f, g \in k[x_1, \dots, x_n]$ are non-zero, then*

$$\langle f \rangle \cap \langle g \rangle = \langle \text{LCM}(f, g) \rangle.$$

Proof. Let $l = \text{LCM}(f, g)$. Then $l \in \langle f \rangle \cap \langle g \rangle$ by the definition of least common multiple. Hence

$$\langle l \rangle \subseteq \langle f \rangle \cap \langle g \rangle.$$

Conversely, let $h \in \langle f \rangle \cap \langle g \rangle$. Then

$$h = af = bg$$

for some $a, b \in k[x_1, \dots, x_n]$. Thus f divides h and g divides h . By the definition of $\text{LCM}(f, g)$, it follows that l divides h . Hence

$$h \in \langle l \rangle.$$

Therefore,

$$\langle f \rangle \cap \langle g \rangle = \langle l \rangle = \langle \text{LCM}(f, g) \rangle.$$

□

Example 3.2.8. ([4]) Let $f = x^2y^2 - y^2 + x^2 - 1$ and $g = xy^2 - y^2 - x + 1$ be polynomials in $\mathbb{Q}[x, y]$. To compute $\text{LCM}(f, g)$, we first compute the reduced Gröbner basis G for the ideal

$$\langle wf, (1 - w)g \rangle = \langle w(x^2y^2 - y^2 + x^2 - 1), (1 - w)(xy^2 - y^2 - x + 1) \rangle \subseteq \mathbb{Q}[x, y, w]$$

using the lexicographic term ordering with $w > x > y$.

This yields

$$G = \{x^2y^4 - x^2 - y^4 + 1, wxy^2 + wx + wy^2 - w + xy^2 - x - y^2 + 1, 2wx^2 - 2w + x^2y^2 - x^2 - y^2 + 1\}.$$

Hence,

$$\text{LCM}(f, g) = 1 - x^2 - y^4 + x^2y^4.$$

To compute $\text{GCD}(f, g)$, we divide fg by $\text{LCM}(f, g)$ using the Division Algorithm. The quotient obtained is

$$x - 1.$$

Therefore,

$$\text{GCD}(f, g) = x - 1.$$

Definition 3.2.9. Let I and J be ideals in $k[x_1, \dots, x_n]$. The *ideal quotient* $J : I$ is defined by

$$J : I = \{g \in k[x_1, \dots, x_n] \mid gI \subseteq J\}.$$

We now discuss how elimination theory can be used to compute ideal quotients. The following results reduce the computation of a general colon ideal to simpler cases and provide an effective Gröbner basis method for determining such ideals. These methods further illustrate the computational utility of elimination in ideal-theoretic constructions.

Lemma 3.2.10 ([1] Proposition 13, Section 4, p. 204]. *Let $I = \langle f_1, \dots, f_s \rangle$ and J be ideals in $k[x_1, \dots, x_n]$. Then*

$$J : I = \bigcap_{i=1}^s (J : \langle f_i \rangle).$$

Proof. Let $g \in J : I$. Then, by definition, $gI \subseteq J$. Since $f_i \in I$ for each $i = 1, \dots, s$, it follows that $gf_i \in J$ for all $i = 1, \dots, s$. Hence,

$$g \in J : \langle f_i \rangle \quad \text{for all } i = 1, \dots, s,$$

and therefore

$$g \in \bigcap_{i=1}^s (J : \langle f_i \rangle).$$

Conversely, let

$$g \in \bigcap_{i=1}^s (J : \langle f_i \rangle).$$

Then

$$gf_i \in J \quad \text{for all } i = 1, \dots, s.$$

Now every element $h \in I$ can be written as $h = \sum_{i=1}^s a_i f_i$ for some $a_i \in k[x_1, \dots, x_n]$. Multiplying by g , we obtain

$$gh = \sum_{i=1}^s a_i (gf_i) \in J,$$

since each $gf_i \in J$ and J is an ideal. Thus

$$gI \subseteq J,$$

which implies that

$$g \in J : I.$$

Hence,

$$J : I = \bigcap_{i=1}^s (J : \langle f_i \rangle).$$

□

Lemma 3.2.11 ([1, Proposition 13, Section 4, p. 204]). *Let J be an ideal and let $0 \neq f \in k[x_1, \dots, x_n]$. Then*

$$J : \langle f \rangle = \frac{1}{f} (J \cap \langle f \rangle).$$

Proof. Let $g \in \frac{1}{f} (J \cap \langle f \rangle)$. Then there exists $h \in J \cap \langle f \rangle$ such that $g = \frac{h}{f}$. Since $h \in J$, we have

$$gf = h \in J.$$

Therefore,

$$g \in J : \langle f \rangle.$$

Conversely, let $g \in J : \langle f \rangle$. Then, by definition, $gf \in J$. Clearly, $gf \in \langle f \rangle$ hence $gf \in J \cap \langle f \rangle$. Thus,

$$g = \frac{1}{f} (gf) \in \frac{1}{f} (J \cap \langle f \rangle).$$

Therefore,

$$J : \langle f \rangle = \frac{1}{f} (J \cap \langle f \rangle).$$

□

Example 3.2.12. ([4]) Let $g_1 = x(x + y)^2$, $g_2 = y$, $f_1 = x^2$, $f_2 = x + y$ in $\mathbb{Q}[x, y]$. Consider the ideals $I = \langle f_1, f_2 \rangle$ and $J = \langle g_1, g_2 \rangle$. We wish to compute $J : I$.

By Lemma 3.2.10,

$$J : I = (J : \langle f_1 \rangle) \cap (J : \langle f_2 \rangle),$$

and by Lemma 3.2.11,

$$J : I = \frac{1}{f_1}(J \cap \langle f_1 \rangle) \cap \frac{1}{f_2}(J \cap \langle f_2 \rangle).$$

First, we compute $J \cap \langle f_1 \rangle$ by finding a Gröbner basis G_1 for the ideal $\langle wg_1, wg_2, (1 - w)f_1 \rangle \subseteq \mathbb{Q}[x, y, w]$ with respect to the lexicographic order $w > x > y$. This gives $G_1 = \{x^2w - x^2, wy, x^3, x^2y\}$, so that $\frac{1}{f_1}(J \cap \langle f_1 \rangle) = \langle x, y \rangle$.

Next, we compute $J \cap \langle f_2 \rangle$ by finding a Gröbner basis G_2 for the ideal $\langle wg_1, wg_2, (1 - w)f_2 \rangle \subseteq \mathbb{Q}[x, y, w]$ with respect to the same order. This yields $G_2 = \{wx - x - y, wy, x^3 + y^3, xy + y^2\}$, so that $\frac{1}{f_2}(J \cap \langle f_2 \rangle) = \langle x^2 - xy + y^2, y \rangle$.

Finally, we compute

$$\langle x, y \rangle \cap \langle x^2 - xy + y^2, y \rangle$$

using Proposition 3.2.6 by finding a Gröbner basis for the ideal

$$\langle wx, wy, (1 - w)(x^2 - xy + y^2), (1 - w)y \rangle \subseteq \mathbb{Q}[x, y, w]$$

with respect to the lexicographic order $w > x > y$. This yields

$$G = \{wx, x^2, y\}.$$

Therefore,

$$J : I = \langle x^2, y \rangle.$$

3.3 Eigenvalue Methods for Solving Polynomial Systems

In this section, we present an alternative approach for solving polynomial systems using linear algebraic techniques. For zero-dimensional ideals, multiplication by a polynomial defines a linear transformation on the quotient ring, and the solutions of the system can be recovered from the eigenvalues of the corresponding multiplication operators. This method provides an effective bridge between Gröbner basis theory and linear algebra.

Theorem 3.3.1 (Finiteness Theorem [1, Theorem 6, Section 3, p. 251]). *Let I be a proper ideal of the ring $R = k[x_1, x_2, \dots, x_n]$, where k is an algebraically closed field. Then the following are equivalent:*

- (i) $V(I)$ is a finite set.
- (ii) R/I is a finite-dimensional k -vector space.
- (iii) I has a Gröbner basis G containing polynomials $g_1, g_2, \dots, g_n \in G$ such that

$$\text{LT}(g_i) = x_i^{d_i} \quad \text{for all } i = 1, 2, \dots, n.$$

Proof. We prove the equivalence of the three statements.

(i) $\Rightarrow$ (ii): Assume that $V(I)$ is finite. For each $i = 1, \dots, n$, let

$$a_{i1}, a_{i2}, \dots, a_{ir_i}$$

be the distinct i -th coordinates of the points of $V(I)$, and define

$$f_i(x_i) = \prod_{j=1}^{r_i} (x_i - a_{ij}).$$

Then f_i vanishes at every point of $V(I)$, and hence

$$f_i \in I(V(I)).$$

By the Strong Nullstellensatz,

$$I(V(I)) = \sqrt{I}.$$

Therefore, there exists $d_i \geq 1$ such that

$$f_i^{d_i} \in I.$$

Since $f_i^{d_i}$ is a polynomial in x_i whose leading term is a power of x_i , it follows that some power $x_i^{N_i}$ can be expressed modulo I as a linear combination of lower powers of x_i . Thus every sufficiently high power of x_i is reducible modulo I .

Hence every residue class in R/I can be represented by a linear combination of monomials

$$x_1^{a_1} \cdots x_n^{a_n}$$

with bounded exponents. Since there are only finitely many such monomials, R/I is finite-dimensional over k .

(ii) $\Rightarrow$ (iii): Assume that R/I is finite-dimensional as a k -vector space.

For each fixed $i \in \{1, \dots, n\}$, consider the infinite sequence

$$[1], [x_i], [x_i^2], [x_i^3], \dots$$

in R/I . Since R/I is finite-dimensional, these residue classes must be linearly dependent. Therefore, there exist scalars

$$\alpha_0, \alpha_1, \dots, \alpha_t \in k,$$

not all zero, such that

$$\alpha_0[1] + \alpha_1[x_i] + \cdots + \alpha_t[x_i^t] = 0.$$

Hence

$$f_i(x_i) = \alpha_0 + \alpha_1 x_i + \cdots + \alpha_t x_i^t \in I.$$

It follows that

$$x_i^t \in \text{LT}(I).$$

Repeating this argument for each $i = 1, \dots, n$, we conclude that for every variable x_i , some power $x_i^{d_i}$ lies in $\text{LT}(I)$. Hence a Gröbner basis G of I may be chosen containing polynomials

$$g_i \in G$$

such that

$$\text{LT}(g_i) = x_i^{d_i}, \quad i = 1, \dots, n.$$

(iii) $\Rightarrow$ (i): Assume that I has a Gröbner basis G containing polynomials

$$g_1, \dots, g_n$$

such that

$$\text{LT}(g_i) = x_i^{d_i}, \quad i = 1, \dots, n.$$

After reordering variables if necessary, we may assume

$$x_n > \dots > x_2 > x_1.$$

Then each g_i involves only the variables $x_1, \dots, x_i$.

Thus the equation

$$g_1(x_1) = 0$$

has only finitely many solutions for x_1 . For each such solution, the equation

$$g_2(x_1, x_2) = 0$$

has only finitely many possible values of x_2 . Proceeding inductively, each variable admits only finitely many choices. Hence $V(I)$ is finite. $\square$

Definition 3.3.2. An ideal $I \subseteq k[x_1, \dots, x_n]$ is called *zero-dimensional* if the variety $V(I)$ is finite. Equivalently, by the Finiteness Theorem, I is zero-dimensional if and only if R/I is finite-dimensional as a k -vector space.

3.3.1 Number of Solutions

For a zero-dimensional ideal, an important question is how the dimension of the quotient ring R/I relates to the number of points in the variety $V(I)$. In general, $\dim_k(R/I)$ provides an algebraic measure of the size of the solution set, counting solutions with multiplicity. The following result makes this relationship precise in the radical case.

We begin by recalling the construction of interpolating polynomials, which will be used to distinguish points in a finite variety. These polynomials play an important role in relating the number of solutions of a zero-dimensional ideal to the dimension of the corresponding quotient ring.

Proposition 3.3.3. *Let $p_1, p_2, \dots, p_m \in k^n$ be distinct points. Then there exist polynomials $g_1, g_2, \dots, g_m \in k[x_1, \dots, x_n]$ such that*

$$g_i(p_j) = 0 \quad \text{for } j \neq i,$$

and

$$g_i(p_i) = 1$$

for each $i = 1, 2, \dots, m$.

Proof. Since the points $p_1, p_2, \dots, p_m$ are distinct, for each pair $i \neq j$ there exists at least one coordinate index $r = r(i, j)$ such that the r -th coordinates of p_i and p_j are different. Write

$$p_i = (a_{i1}, a_{i2}, \dots, a_{in}).$$

For each fixed i and each $j \neq i$, define the linear polynomial

$$\ell_{ij}(x) = x_{r(i,j)} - a_{jr(i,j)}.$$

Then

$$\ell_{ij}(p_j) = 0 \quad \text{and} \quad \ell_{ij}(p_i) \neq 0.$$

Now define

$$g_i(x) = \prod_{j \neq i} \frac{\ell_{ij}(x)}{\ell_{ij}(p_i)}.$$

Since each $\ell_{ij}(p_i) \neq 0$, this is a well-defined polynomial in $k[x_1, \dots, x_n]$.

For $j \neq i$, one factor in the product vanishes at p_j , namely $\ell_{ij}(p_j) = 0$, so

$$g_i(p_j) = 0.$$

On the other hand, evaluating at p_i gives

$$g_i(p_i) = \prod_{j \neq i} \frac{\ell_{ij}(p_i)}{\ell_{ij}(p_i)} = 1.$$

Hence the polynomials $g_1, g_2, \dots, g_m$ satisfy the required properties. $\square$

Definition 3.3.4. Let $p_1, p_2, \dots, p_m \in k^n$ be distinct points, and let $g_1, g_2, \dots, g_m$ be polynomials satisfying

$$g_i(p_j) = 0 \quad \text{for } j \neq i, \quad \text{and} \quad g_i(p_i) = 1$$

for each $i = 1, 2, \dots, m$. Then the polynomials $g_1, g_2, \dots, g_m$ are called the *interpolation polynomials* for the points $p_1, p_2, \dots, p_m$.

Theorem 3.3.5 ([1, Proposition 7, Section 3, p. 253]). *Let I be a zero-dimensional ideal of $R = k[x_1, x_2, \dots, x_n]$, where k is an algebraically closed field. Then*

$$|V(I)| = \dim_k(R/\sqrt{I}),$$

and

$$|V(I)| = \dim_k(R/I) \quad \text{if and only if} \quad I = \sqrt{I}.$$

Proof. Let

$$V(I) = \{p_1, p_2, \dots, p_m\} \subseteq k^n.$$

Define a k -linear map

$$\phi : R/I \rightarrow k^m, \quad \phi([f]) = (f(p_1), f(p_2), \dots, f(p_m)).$$

This map is well-defined, since if $[f] = [g]$ in R/I , then $f - g \in I$, and hence

$$(f - g)(p_i) = 0$$

for every $p_i \in V(I)$. Therefore, $f(p_i) = g(p_i)$ for all $i = 1, \dots, m$.

Now,

$$\ker(\phi) = \{[f] \in R/I : f(p_i) = 0 \text{ for all } i = 1, \dots, m\}.$$

Hence,

$$\ker(\phi) = I(V(I))/I.$$

By the Strong Nullstellensatz,

$$I(V(I)) = \sqrt{I},$$

so that

$$\ker(\phi) = \sqrt{I}/I.$$

Next, let $g_1, \dots, g_m$ be interpolation polynomials for the points $p_1, \dots, p_m$. Then $\phi([g_i]) = e_i$, where e_i is the i -th standard basis vector of k^m . Hence ϕ is surjective.

By the Rank–Nullity Theorem,

$$m = \dim_k(k^m) = \dim_k(R/I) - \dim_k(\ker \phi).$$

Substituting $\ker(\phi) = \sqrt{I}/I$, we obtain

$$|V(I)| = m = \dim_k(R/I) - \dim_k(\sqrt{I}/I).$$

Using the short exact sequence

$$0 \rightarrow \sqrt{I}/I \rightarrow R/I \rightarrow R/\sqrt{I} \rightarrow 0,$$

it follows that

$$|V(I)| = \dim_k(R/\sqrt{I}).$$

Finally,

$$|V(I)| = \dim_k(R/I)$$

if and only if

$$\dim_k(\sqrt{I}/I) = 0,$$

which holds if and only if

$$\sqrt{I} = I.$$

Therefore,

$$|V(I)| = \dim_k(R/I) \iff I = \sqrt{I}.$$

□

3.3.2 The Eigenvector Theorems

Let I be a zero-dimensional ideal in the ring $S = k[x_1, x_2, \dots, x_n]$. Then $R = S/I$ is a finite-dimensional k -vector space. We have seen that the number of points in $V(I)$ is at most $\dim_k R$, with equality if and only if I is a radical ideal.

We now study the role played by the eigenvalues of the multiplication maps

$$m_g : R \rightarrow R.$$

We shall show that the eigenvalues of m_g are precisely the values $g(p)$, where $p \in V(I)$. This observation leads to Stickelberger's Theorem, which identifies all points of $V(I)$ from common eigenvectors of the multiplication maps

$$m_{x_1}, m_{x_2}, \dots, m_{x_n}.$$

We begin with an illustrative example in the case of polynomials in one variable.

Example 3.3.6. ([4]) Let $f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n \in S = \mathbb{C}[x]$ be a polynomial of degree n with $a_0 \neq 0$. Consider the multiplication map

$$m_x : S/\langle f(x) \rangle \rightarrow S/\langle f(x) \rangle, \quad m_x([g(x)]) = [xg(x)].$$

With respect to the basis

$$\{[1], [x], [x^2], \dots, [x^{n-1}]\},$$

the matrix of m_x is

$$M_x = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -\frac{a_n}{a_0} \\ 1 & 0 & 0 & \cdots & 0 & -\frac{a_{n-1}}{a_0} \\ 0 & 1 & 0 & \cdots & 0 & -\frac{a_{n-2}}{a_0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -\frac{a_1}{a_0} \end{bmatrix}.$$

The transpose of M_x is called the companion matrix of $f(x)$.

We now show that the eigenvalues of m_x are precisely the roots of $f(x)$.

Let $r \in \mathbb{C}$ be a root of $f(x)$. Then $f(x) = (x - r)g(x)$ for some $g(x) \in \mathbb{C}[x]$. Passing to the quotient ring, we obtain $[f(x)] = 0 = [(x - r)g(x)] = [xg(x)] - r[g(x)]$. Hence

$$m_x([g(x)]) = r[g(x)].$$

Since $\deg g < n$, we have $[g(x)] \neq 0$, and thus r is an eigenvalue of m_x with eigenvector $[g(x)]$.

Conversely, suppose that $m_x([g(x)]) = r[g(x)]$ for some nonzero $[g(x)] \in S/\langle f(x) \rangle$ and $r \in \mathbb{C}$. Then $[xg(x)] = r[g(x)]$, so $xg(x) - rg(x) \in \langle f(x) \rangle$. Thus there exists $h(x) \in \mathbb{C}[x]$ such that

$$f(x)h(x) = (x - r)g(x).$$

Suppose, for contradiction, that $(x - r)$ divides $h(x)$. Then we may write $h(x) = (x - r)h_1(x)$ for some $h_1(x) \in \mathbb{C}[x]$. Substituting into the relation $f(x)h(x) = (x - r)g(x)$, we obtain

$$f(x)h_1(x) = g(x).$$

Hence $f(x)$ divides $g(x)$, which implies that

$$[g(x)] = 0$$

in $S/\langle f(x) \rangle$, contradicting the assumption that $[g(x)] \neq 0$.

Therefore $(x - r)$ does not divide $h(x)$. Since

$$f(x)h(x) = (x - r)g(x),$$

it follows that $(x - r)$ must divide $f(x)$.

Thus each eigenvalue of m_x is a root of $f(x)$, and conversely each root of $f(x)$ is an eigenvalue of m_x .

This establishes a direct connection between factorization of polynomials and eigenvalues of the multiplication operator, which is the key idea in extending this viewpoint to multivariable polynomial systems.

Theorem 3.3.7 (The First Theorem [2] pp. 143–146). *Let I be a radical ideal of $R = k[x_1, x_2, \dots, x_n]$ with*

$$V(I) = \{p_1, p_2, \dots, p_m\}.$$

Let $g_1, g_2, \dots, g_m$ be interpolation polynomials for $V(I)$. For any $g \in R$, define the linear map

$$m_g : R/I \rightarrow R/I, \quad m_g([f]) = [fg] \text{ for all } f \in R.$$

Then $\{[g_1], [g_2], \dots, [g_m]\}$ is a basis of eigenvectors for m_g with corresponding eigenvalues

$$\{g(p_1), g(p_2), \dots, g(p_m)\}.$$

Proof. We first show that $[g_i]$ is an eigenvector of $m_g : R/I \rightarrow R/I$ with eigenvalue $g(p_i)$ for each $i = 1, 2, \dots, m$.

For any $j \in \{1, 2, \dots, m\}$, we compute

$$(gg_i - g(p_i)g_i)(p_j) = g(p_j)g_i(p_j) - g(p_i)g_i(p_j).$$

If $j \neq i$, then $g_i(p_j) = 0$, and hence the above expression is 0. If $j = i$, then $g_i(p_i) = 1$, and the expression also equals 0. Therefore,

$$(gg_i - g(p_i)g_i)(p_j) = 0 \quad \text{for all } j,$$

so the polynomial $gg_i - g(p_i)g_i$ vanishes on $V(I)$.

Hence,

$$gg_i - g(p_i)g_i \in I(V(I)).$$

Since I is radical, by the Strong Nullstellensatz we have

$$I(V(I)) = I,$$

and therefore

$$gg_i - g(p_i)g_i \in I.$$

Passing to the quotient gives

$$[gg_i] = g(p_i)[g_i],$$

that is,

$$m_g([g_i]) = g(p_i)[g_i].$$

Thus $[g_i]$ is an eigenvector of m_g with eigenvalue $g(p_i)$.

Next, consider the map

$$\phi : R/I \rightarrow k^m, \quad \phi([f]) = (f(p_1), \dots, f(p_m)).$$

As shown earlier, ϕ is an isomorphism, and for each i ,

$$\phi([g_i]) = e_i,$$

where e_i is the i -th standard basis vector of k^m . Hence the elements $[g_1], \dots, [g_m]$ are linearly independent.

Since there are m of them and $\dim_k(R/I) = m$, they form a basis of R/I . As each $[g_i]$ is an eigenvector of m_g , this basis is an eigenvector basis with eigenvalues

$$g(p_1), g(p_2), \dots, g(p_m).$$

□

Theorem 3.3.8 (Stickelberger [12] pp. 248–253]. *Let I be a zero-dimensional radical ideal of the polynomial ring $R = k[x_1, x_2, \dots, x_n]$. Then*

$$V(I) = \{(a_1, a_2, \dots, a_n) : m_{x_i}([g]) = a_i[g] \text{ for some } [g] \neq 0, \text{ for all } i\}.$$

Proof. Suppose that $V(I) = \{p_1, p_2, \dots, p_m\}$, where $p_1 = (a_1, a_2, \dots, a_n)$. Let $g_1, g_2, \dots, g_m$ be the interpolation polynomials corresponding to $p_1, p_2, \dots, p_m$.

We first show that $(a_1, \dots, a_n)$ arises from a common eigenvector. For each i , consider the polynomial

$$x_i g_1 - a_i g_1.$$

Evaluating at the points of $V(I)$, we obtain $(x_i g_1 - a_i g_1)(p_1) = a_i \cdot 1 - a_i \cdot 1 = 0$, and for $j \neq 1$, $(x_i g_1 - a_i g_1)(p_j) = x_i(p_j) \cdot 0 - a_i \cdot 0 = 0$. Hence $x_i g_1 - a_i g_1$ vanishes on $V(I)$, so

$$x_i g_1 - a_i g_1 \in I(V(I)) = I,$$

since I is radical. Passing to the quotient, we get

$$m_{x_i}([g_1]) = a_i[g_1] \quad \text{for all } i,$$

so p_1 arises as a tuple of eigenvalues corresponding to a common eigenvector.

Conversely, suppose that $b = (b_1, b_2, \dots, b_n)$ and there exists a non-zero class $[g] \in R/I$ such that

$$m_{x_i}([g]) = b_i[g] \quad \text{for all } i.$$

Then

$$[x_i g] = b_i[g],$$

so

$$(x_i - b_i)g \in I \quad \text{for all } i.$$

Since $[g] \neq 0$, we have $g \notin I$, hence there exists $p \in V(I)$ such that $g(p) \neq 0$.

Evaluating $(x_i - b_i)g$ at such a point p , we obtain

$$(x_i(p) - b_i)g(p) = 0.$$

Since $g(p) \neq 0$, it follows that

$$x_i(p) = b_i \quad \text{for all } i.$$

Thus $p = (b_1, b_2, \dots, b_n) \in V(I)$.

Therefore,

$$V(I) = \{(a_1, \dots, a_n) : m_{x_i}([g]) = a_i[g] \text{ for some } [g] \neq 0, \forall i\}.$$

□

Theorem 3.3.9 (The Second Eigenvector Theorem [2] pp. 143–146]. *Let I be a radical ideal of $S = k[x_1, x_2, \dots, x_n]$ with $V(I) = \{p_1, p_2, \dots, p_m\}$. Let g be a linear form in S such that $g(p_1), g(p_2), \dots, g(p_m)$ are distinct. Then $\{C_1, C_2, \dots, C_m\}$ is a basis of eigenvectors of M_g^t such that*

$$M_g^t C_i = g(p_i) C_i \quad \text{for } i = 1, 2, \dots, m.$$

Proof. Let $\{[x^{\alpha_1}], [x^{\alpha_2}], \dots, [x^{\alpha_m}]\}$ be a basis of R/I . Consider the multiplication map

$$m_g : R/I \rightarrow R/I, \quad m_g([f]) = [gf].$$

For each basis element, we can write

$$m_g([x^{\alpha_j}]) = [gx^{\alpha_j}] = \sum_{i=1}^m c_{ij} [x^{\alpha_i}].$$

Thus the matrix representation of m_g with respect to this basis is $M_g = (c_{ij})$.

Now evaluate the above identity at a point $p_r \in V(I)$. Since evaluation is compatible with the quotient structure, we obtain

$$g(p_r) p_r^{\alpha_j} = \sum_{i=1}^m c_{ij} p_r^{\alpha_i}.$$

Interpreting this relation in vector form, we see that the vector

$$C_r = (p_r^{\alpha_1}, p_r^{\alpha_2}, \dots, p_r^{\alpha_m})^t$$

satisfies

$$M_g^t C_r = g(p_r) C_r \quad \text{for } r = 1, 2, \dots, m.$$

Hence each C_r is an eigenvector of M_g^t with eigenvalue $g(p_r)$.

Since the values $g(p_1), g(p_2), \dots, g(p_m)$ are assumed to be distinct, the corresponding eigenvectors $C_1, C_2, \dots, C_m$ are linearly independent. As there are m of them in an m -dimensional space, they form a basis of eigenvectors of M_g^t . $\square$

Chapter 4

Minimal Polynomials of Elements in Field Extensions

In this chapter, we study minimal polynomials of algebraic elements in field extensions from a computational perspective. The main idea is to realize the minimal polynomial as the kernel of a suitable polynomial map, and then compute it using elimination theory and Gröbner basis techniques. This approach allows us to translate questions about field extensions into problems of ideal elimination in polynomial rings, making the computation of minimal polynomials algorithmic.

4.1 Polynomial Maps

In this section, we study k -algebra homomorphisms between polynomial rings.

Definition 4.1.1. A k -algebra homomorphism is a ring homomorphism

$$\phi : k[y_1, y_2, \dots, y_m] \longrightarrow k[x_1, x_2, \dots, x_n]$$

which is also linear over k , i.e., it preserves addition, multiplication, and scalar multiplication by elements of k .

Such a homomorphism is completely determined by the images of the generators. If

$$\phi(y_i) = f_i \in k[x_1, x_2, \dots, x_n], \quad 1 \leq i \leq m,$$

then ϕ acts on any polynomial $h \in k[y_1, \dots, y_m]$ by substitution.

Let

$$h = \sum_{\nu} c_{\nu} y_1^{\nu_1} \cdots y_m^{\nu_m}, \quad c_{\nu} \in k.$$

Then

$$\phi(h) = \sum_{\nu} c_{\nu} f_1^{\nu_1} \cdots f_m^{\nu_m} = h(f_1, f_2, \dots, f_m).$$

Thus, ϕ is a substitution homomorphism induced by $y_i \mapsto f_i$.

Example 4.1.2. Let $\phi : k[y_1, y_2] \rightarrow k[x]$ be defined by

$$\phi(y_1) = x^2, \quad \phi(y_2) = x + 1.$$

Then for $h(y_1, y_2) = y_1 + y_2^2$, we get

$$\phi(h) = x^2 + (x + 1)^2.$$

The kernel of ϕ is an ideal in $k[y_1, \dots, y_m]$, and its image is the k -subalgebra

$$\text{im}(\phi) = k[f_1, f_2, \dots, f_m] \subseteq k[x_1, \dots, x_n].$$

By the First Isomorphism Theorem for k -algebras,

$$k[y_1, \dots, y_m] / \ker(\phi) \cong k[f_1, f_2, \dots, f_m].$$

Our main objective in this section is to use elimination theory and Gröbner bases to study $\ker(\phi)$ and $\text{im}(\phi)$. In particular, we aim to:

- compute a Gröbner basis for $\ker(\phi)$ using elimination,
- determine whether a given polynomial lies in $\text{im}(\phi)$,
- and later apply these ideas to compute minimal polynomials in field extensions.

We will now develop the elimination framework required for this study.

Lemma 4.1.3. *Let $a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n$ be elements of a commutative ring R . Then*

$$a_1 a_2 \cdots a_n - b_1 b_2 \cdots b_n \in \langle a_1 - b_1, a_2 - b_2, \dots, a_n - b_n \rangle.$$

Proof. We prove the statement by induction on n .

For $n = 1$, the statement is immediate since

$$a_1 - b_1 \in \langle a_1 - b_1 \rangle.$$

Assume that the result holds for all positive integers less than n , where $n \geq 2$. Since R is commutative, we can write

$$a_1 a_2 \cdots a_n - b_1 b_2 \cdots b_n = a_n(a_1 a_2 \cdots a_{n-1} - b_1 b_2 \cdots b_{n-1}) + b_1 b_2 \cdots b_{n-1}(a_n - b_n). \quad (4.3)$$

By the induction hypothesis,

$$a_1 a_2 \cdots a_{n-1} - b_1 b_2 \cdots b_{n-1} \in \langle a_1 - b_1, a_2 - b_2, \dots, a_{n-1} - b_{n-1} \rangle.$$

Multiplying by a_n preserves membership in the ideal, hence

$$a_n(a_1 a_2 \cdots a_{n-1} - b_1 b_2 \cdots b_{n-1}) \in \langle a_1 - b_1, \dots, a_{n-1} - b_{n-1} \rangle.$$

Similarly, since $a_n - b_n \in \langle a_n - b_n \rangle$, we have

$$b_1 b_2 \cdots b_{n-1}(a_n - b_n) \in \langle a_n - b_n \rangle.$$

Combining these two observations and using equation (4.3), we conclude that

$$a_1 a_2 \cdots a_n - b_1 b_2 \cdots b_n \in \langle a_1 - b_1, a_2 - b_2, \dots, a_n - b_n \rangle.$$

Thus, the statement holds for n , and by induction it holds for all $n \in \mathbb{N}$. $\square$

Theorem 4.1.4 ([1] Theorem 1, Section 3, p. 133). *Let $\phi : k[y_1, y_2, \dots, y_m] \longrightarrow k[x_1, x_2, \dots, x_n]$ be the k -algebra homomorphism defined by $\phi(y_i) = f_i \in k[x_1, \dots, x_n]$, $1 \leq i \leq m$. Let $I = \langle y_1 - f_1, y_2 - f_2, \dots, y_m - f_m \rangle \subset k[y_1, \dots, y_m, x_1, \dots, x_n]$. Then*

$$\ker(\phi) = I \cap k[y_1, y_2, \dots, y_m].$$

Proof. First, let $g \in I \cap k[y_1, y_2, \dots, y_m]$. Then $g \in I$, so we can write

$$g(y_1, \dots, y_m) = \sum_{i=1}^m (y_i - f_i(x_1, \dots, x_n)) h_i(y_1, \dots, y_m, x_1, \dots, x_n),$$

where $h_i \in k[y_1, \dots, y_m, x_1, \dots, x_n]$. Applying ϕ , and using the fact that $\phi(y_i) = f_i$, we obtain

$$\phi(g) = g(f_1, \dots, f_m) = 0.$$

Hence $g \in \ker(\phi)$.

Conversely, let $g \in \ker(\phi)$. Then $\phi(g) = 0$, i.e.,

$$g(f_1, \dots, f_m) = 0.$$

We write

$$g(y_1, \dots, y_m) = \sum_{\nu} c_{\nu} y_1^{\nu_1} \cdots y_m^{\nu_m},$$

where $c_{\nu} \in k$, $\nu = (\nu_1, \dots, \nu_m) \in \mathbb{Z}_{\geq 0}^m$, and only finitely many c_{ν} are non-zero. Then

$$g - \phi(g) = \sum_{\nu} c_{\nu} (y_1^{\nu_1} \cdots y_m^{\nu_m} - f_1^{\nu_1} \cdots f_m^{\nu_m}).$$

By the previous lemma, each difference

$$y_1^{\nu_1} \cdots y_m^{\nu_m} - f_1^{\nu_1} \cdots f_m^{\nu_m}$$

lies in the ideal

$$\langle y_1 - f_1, y_2 - f_2, \dots, y_m - f_m \rangle = I.$$

Hence each term in the above sum belongs to I , and therefore $g \in I$.

Since g depends only on $y_1, \dots, y_m$, we conclude that

$$g \in I \cap k[y_1, y_2, \dots, y_m].$$

Thus,

$$\ker(\phi) = I \cap k[y_1, y_2, \dots, y_m].$$

□

This theorem allows us to compute the kernel of ϕ using elimination theory and Gröbner bases in a larger polynomial ring.

Definition 4.1.5. A k -algebra A is called an *affine k -algebra* if it is isomorphic, as a k -algebra, to

$$k[x_1, x_2, \dots, x_n]/I$$

for some ideal $I \subset k[x_1, x_2, \dots, x_n]$.

Clearly, the polynomial ring $k[x_1, \dots, x_n]$ is an affine k -algebra. Moreover, if $f_1, \dots, f_m \in k[x_1, \dots, x_n]$, then, as observed earlier, the image $k[f_1, \dots, f_m]$ of the homomorphism

$$\phi : k[y_1, \dots, y_m] \longrightarrow k[x_1, \dots, x_n], \quad y_i \mapsto f_i,$$

is isomorphic to $k[y_1, \dots, y_m]/\ker(\phi)$, and hence it is an affine k -algebra.

We now study k -algebra homomorphisms between affine k -algebras. Let $J \subset k[y_1, \dots, y_m]$ and $I \subset k[x_1, \dots, x_n]$ be ideals, and consider a k -algebra homomorphism

$$\psi : k[y_1, \dots, y_m]/J \longrightarrow k[x_1, \dots, x_n]/I, \quad y_i + J \mapsto f_i + I.$$

This map is well-defined if and only if every polynomial relation defining J is preserved under substitution, that is, if

$$J = \langle g_1, \dots, g_t \rangle \Rightarrow g_i(f_1, \dots, f_m) \in I \text{ for all } i = 1, \dots, t.$$

This condition is a direct analogue of Theorem ??, and it naturally extends the kernel-elimination correspondence to homomorphisms between affine k -algebras.

Theorem 4.1.6 ([1 Theorem 2, Section 3, p. 138]). *Let K be the ideal of $k[y_1, \dots, y_m, x_1, \dots, x_n]$ generated by the ideal I together with the polynomials $y_i - f_i$, $1 \leq i \leq m$, that is, $K = \langle I, y_1 - f_1, \dots, y_m - f_m \rangle$. Then $\ker(\psi) = (K \cap k[y_1, \dots, y_m]) \bmod J$. That is, if $K \cap k[y_1, \dots, y_m] = \langle f_1, \dots, f_p \rangle$, then*

$$\ker(\psi) = \langle f_1 + J, \dots, f_p + J \rangle.$$

Proof. Let $f \in K \cap k[y_1, y_2, \dots, y_m]$. Then $f \in K$, and hence we can write

$$f(y_1, \dots, y_m) = \sum_{i=1}^m (y_i - f_i(x_1, \dots, x_n)) h_i + w,$$

where $h_i \in k[y_1, \dots, y_m, x_1, \dots, x_n]$ and

$$w = \sum_{\nu} u_{\nu}(y_1, \dots, y_m, x_1, \dots, x_n) v_{\nu}(x_1, \dots, x_n), \quad v_{\nu} \in I.$$

Applying the homomorphism ψ and using $\psi(y_i) = f_i$, we obtain $\psi(f + J) = f(f_1, \dots, f_m) + I$. Substituting the decomposition above gives

$$f(f_1, \dots, f_m) + I = w(f_1, \dots, f_m, x_1, \dots, x_n) + I.$$

Since each $v_{\nu} \in I$, we have $w(f_1, \dots, f_m, x_1, \dots, x_n) \in I$, hence $\psi(f + J) = 0$. Therefore $f + J \in \ker(\psi)$.

Conversely, let $f \in k[y_1, \dots, y_m]$ satisfy $\psi(f + J) = 0$. Then $f(f_1, \dots, f_m) \in I$. Write $f = \sum_{\nu} c_{\nu} y_1^{\nu_1} \dots y_m^{\nu_m}$. Then

$$f(y_1, \dots, y_m) - f(f_1, \dots, f_m) = \sum_{\nu} c_{\nu} (y_1^{\nu_1} \dots y_m^{\nu_m} - f_1^{\nu_1} \dots f_m^{\nu_m}).$$

By Lemma 4.1.3, each term lies in $\langle y_1 - f_1, \dots, y_m - f_m \rangle$. Hence $f \in K$. Since f involves only y -variables, we conclude $f \in K \cap k[y_1, \dots, y_m]$.

Thus,

$$\ker(\psi) = (K \cap k[y_1, \dots, y_m]) \bmod J.$$

This completes the proof. $\square$

This establishes a direct connection between k -algebra homomorphisms and elimination theory. In particular, kernels of polynomial maps can be computed as elimination ideals in a larger polynomial ring. This viewpoint will be crucial in the subsequent sections, where we apply these ideas to study algebraic elements in field extensions and compute minimal polynomials using Gröbner basis techniques.

4.2 The Minimal Polynomials of Elements

Let $k \hookrightarrow K$ be a field extension. An element $\alpha \in K$ is said to be algebraic over k if there exists a non-zero polynomial $p(x) \in k[x]$ such that $p(\alpha) = 0$. In this case, the *minimal polynomial* of α over k is defined as the unique monic polynomial of least degree in $k[x]$ having α as a root.

Alternatively, consider the k -algebra homomorphism

$$\phi : k[x] \longrightarrow k[\alpha], \quad x \mapsto \alpha.$$

Then $\ker(\phi)$ is a principal ideal generated by the minimal polynomial $p(x)$ of α , that is, $\ker(\phi) = \langle p(x) \rangle$. Moreover, the map ϕ is surjective. Hence, by the First Isomorphism Theorem for k -algebras, we obtain

$$k[x]/\langle p(x) \rangle \cong k[\alpha].$$

Proposition 4.2.1. *The minimal polynomial $p(x)$ of α over k is irreducible in $k[x]$.*

Proof. Suppose that $p(x)$ is not irreducible over k . Then there exist non-constant polynomials $p_1(x), p_2(x) \in k[x]$ such that

$$p(x) = p_1(x)p_2(x), \quad 1 \leq \deg(p_1) < \deg(p), \quad 1 \leq \deg(p_2) < \deg(p).$$

Evaluating at α , we obtain

$$p(\alpha) = 0 = p_1(\alpha)p_2(\alpha),$$

which implies $p_1(\alpha) = 0$ or $p_2(\alpha) = 0$. This contradicts the minimality of the degree of $p(x)$ as the minimal polynomial of α . Hence, $p(x)$ is irreducible over k .

Since $k[x]$ is a principal ideal domain, the ideal $\langle p(x) \rangle$ is maximal. Therefore,

$$k[x]/\langle p(x) \rangle$$

is a field. Using the isomorphism $k[x]/\langle p(x) \rangle \cong k[\alpha]$, we conclude that $k[\alpha]$ is a field whenever α is algebraic over k . $\square$

Definition 4.2.2. Let α be algebraic over a field k . The field generated by α over k , denoted by $k(\alpha)$, is defined as the smallest field containing both k and α . Equivalently,

$$k(\alpha) = \left\{ \frac{f(\alpha)}{g(\alpha)} \mid f(x), g(x) \in k[x], g(\alpha) \neq 0 \right\}.$$

Clearly, we have $k[\alpha] \subseteq k(\alpha)$.

If α is algebraic over k , then $k[\alpha]$ is already a field (as shown earlier). Hence every element of $k(\alpha)$ lies in $k[\alpha]$, giving $k(\alpha) \subseteq k[\alpha]$. Since we always have $k[\alpha] \subseteq k(\alpha)$, it follows that

$$k(\alpha) = k[\alpha].$$

Consequently, if $p(x)$ is the minimal polynomial of α over k , then

$$k[x]/\langle p(x) \rangle \cong k[\alpha] = k(\alpha).$$

We now come to the main computational objective of this section. Let $K = k(\alpha)$, where α is algebraic over k with minimal polynomial $p(x)$. Since we have the identification $K \cong k[x]/\langle p(x) \rangle$, computations in K can be carried out inside the affine k -algebra $k[x]/\langle p(x) \rangle$.

Our goal is to determine, for any element $\beta \in K$, its minimal polynomial over k . This will be achieved by translating the problem into computations in the quotient algebra and applying the techniques of elimination theory and Gröbner bases developed in the previous sections.

Theorem 4.2.3. [4] *Let $k \subseteq K$ be a field extension and let $\alpha \in K$ be algebraic over k . Let $p(x) \in k[x]$ be the minimal polynomial of α over k . Let $0 \neq \beta \in k(\alpha)$ be given by*

$$\beta = \frac{a_0 + a_1\alpha + \cdots + a_n\alpha^n}{b_0 + b_1\alpha + \cdots + b_m\alpha^m}, \quad a_i, b_j \in k.$$

Let $f(x) = a_0 + a_1x + \cdots + a_nx^n$ and $g(x) = b_0 + b_1x + \cdots + b_mx^m$. Consider the ideal

$$J = \langle p(x), yg(x) - f(x) \rangle \subset k[x, y].$$

Then the minimal polynomial of β over k is the monic generator of the elimination ideal $J \cap k[y]$.

Proof. Note that $k[x]/\langle p(x) \rangle \cong k(\alpha)$, and let $\psi : k[x]/\langle p(x) \rangle \rightarrow k(\alpha)$ be the corresponding field isomorphism. Also, $g(\alpha) \neq 0$ since $\beta = \frac{f(\alpha)}{g(\alpha)} \in k(\alpha)$. Since $p(x)$ is irreducible over k and $p \nmid g$, we have $\gcd(p, g) = 1$. Hence there exists $l(x) \in k[x]$ such that

$$gl \equiv 1 \pmod{p}, \quad \text{i.e. } gl - 1 \in \langle p \rangle.$$

Thus we may write $gl = tp + 1$ for some $t(x) \in k[x]$. Let $h = fl$. Then

$$h(\alpha) = f(\alpha)l(\alpha) = \beta g(\alpha)l(\alpha) = \beta(gl)(\alpha) = \beta(t(\alpha)p(\alpha) + 1) = \beta,$$

since $p(\alpha) = 0$. Hence $h(\alpha) = \beta$.

Consider the k -algebra homomorphism

$$k[y] \xrightarrow{\pi} k[x]/\langle p \rangle \xrightarrow{\psi} k(\alpha)$$

where $\pi(y) = h(x) + \langle p \rangle$ and $\psi(x) = \alpha$. Define $\phi = \psi \circ \pi$. Then

$$\phi(y) = \psi(\pi(y)) = \psi(h(x) + \langle p \rangle) = h(\alpha) = \beta.$$

Thus, $q(y) \in \ker(\phi)$ if and only if $q(\beta) = 0$. Since $\ker(\phi)$ is an ideal of $k[y]$ and $k[y]$ is a PID, it is generated by a single polynomial. Hence, to determine the minimal polynomial of β , it suffices to compute the generator of $\ker(\phi)$.

Clearly,

$$\ker(\phi) = \langle p(x), y - h(x) \rangle \cap k[y].$$

It remains to show that

$$\langle p(x), y - h(x) \rangle = \langle p(x), gy - f \rangle.$$

Since $gl \equiv 1 \pmod{p}$, we have

$$y - h = y - fl \equiv l(gy - f) \pmod{p},$$

so $y - h \in \langle p, gy - f \rangle$, hence

$$\langle p, y - h \rangle \subseteq \langle p, gy - f \rangle.$$

Conversely,

$$gy - f \equiv g(y - h) \pmod{p},$$

so $gy - f \in \langle p, y - h \rangle$, giving

$$\langle p, gy - f \rangle \subseteq \langle p, y - h \rangle.$$

Therefore,

$$\langle p(x), y - h(x) \rangle = \langle p(x), gy - f \rangle.$$

Hence,

$$\ker(\phi) = \langle p(x), gy - f \rangle \cap k[y] = J \cap k[y].$$

Since $k[y]$ is a PID, $\ker(\phi)$ is generated by a single polynomial $q(y)$. This polynomial q is the minimal polynomial of β over k . $\square$

Example 4.2.4. ([4]) Consider the extension field $\mathbb{Q}(\alpha)$ of $\mathbb{Q}$, where α is a root of the irreducible polynomial $x^5 - x - 2$. Now consider the elements $\beta = \frac{1-\alpha-2\alpha^3}{\alpha} \in \mathbb{Q}(\alpha)$. We wish to find the minimal polynomial of β . We consider the ideal $J = \langle x^5 - x - 2, xy + 2x^3 + x - 1 \rangle$ in $\mathbb{Q}[x, y]$. We compute the reduced Gröbner basis G for J with respect to the lex term ordering with $x > y$ using SINGULAR.

```

> ring A = 0, (x,y), lp;
    characteristic : 0
    number of vars : 2

    block 1 : ordering lp
      : names x y
    block 2 : ordering C

> poly p1 = x5-x-2;
> poly p2 = xy+2x3+x-1;
> ideal J = p1,p2;
> J;
    J[1]=x5-x-2
    J[2]=8x3+xy+x-1

> option(redSB);
> ideal G = std(J);
> G;
    G[1]=y5+11/2y4-116y3-395y2+8807y+143737
    G[2]=x-156671811691/67121160191y4
      +60318217/67121160191y3
      +324472719/67121160191y2
      -3870713717/67121160191y
      -11118796005/67121160191

```

Therefore the minimal polynomial of β is

$$y^5 + \frac{11}{2}y^4 - 116y^3 - 395y^2 + 8807y + 143737.$$

Example 4.2.5. Consider the cyclotomic extension field $\mathbb{Q}(\alpha)$ of $\mathbb{Q}$, where $\alpha = \zeta_7$ is a primitive 7th root of unity, defined as a root of the irreducible cyclotomic polynomial

$$x^6 + x^5 + x^4 + x^3 + x^2 + x + 1.$$

Now consider the rational element

$$\beta = \frac{\alpha^4 - 3\alpha^3 + \alpha - 1}{\alpha^2 - \alpha + 6} \in \mathbb{Q}(\alpha).$$

We wish to find the minimal polynomial of β over $\mathbb{Q}$. We consider the ideal

$$J = \langle x^6 + x^5 + x^4 + x^3 + x^2 + x + 1, x^2y - xy + 6y - x^4 + 3x^3 - x + 1 \rangle$$

in $\mathbb{Q}[x, y]$. We compute the reduced Gröbner basis G for J with respect to the lexicographic term ordering with $x > y$ using SINGULAR.

```

> ring A = 0, (x,y), lp;
> A;
// coefficients: QQ considered as a field
// number of vars : 2
//      block 1 : ordering lp
//      : names x y
//      block 2 : ordering C
> poly p1 = x^6+x^5+x^4+x^3+x^2+x+1;
> poly p2 = x^2*y-x*y+6*y-x^4+3*x^3-x+1;
> ideal J = p1,p2;
> option(redSB);
> ideal G = std(J);
> G;
G[1]=46831y6-124014y5-300548y4+98657y3
      +23153944y2-88002359y+402310721
G[2]=344569518335196596974942678584128x
      +142990518142305603207103676141y5
      +352007798488198698663752651283y4
      -2233680896210531572257485647593y3
      -1259657072548268382615203561558y2
      +4286056096660879288855554879314y
      +70439420407832116990789975488237

```

Therefore the minimal polynomial of β over $\mathbb{Q}$ is

$$46831y^6 - 124014y^5 - 300548y^4 + 98657y^3 + 23153944y^2 - 88002359y + 402310721.$$

In the sequel, we extend the method developed in the previous section to compute minimal polynomials in a more general setting of multivariate field extensions of the form $K = k(\alpha_1, \alpha_2, \dots, \alpha_n)$. The main idea is to reduce computations in the field extension

to elimination problems in a suitable polynomial ring over k , thereby enabling the use of Gröbner basis techniques.

To make this precise, we introduce the following lifting notation. For each $i = 2, \dots, n$ and for any polynomial $p \in k(\alpha_1, \alpha_2, \dots, \alpha_{i-1})[x_i]$, we choose a polynomial $\tilde{p} \in k[x_1, x_2, \dots, x_n]$ such that

$$\tilde{p}(\alpha_1, \alpha_2, \dots, \alpha_{i-1}, x_i) = p.$$

Although such a lift $\tilde{p}$ is not uniquely determined, all constructions in which it is used are independent of the particular choice. This allows us to carry out all computations within polynomial rings over k while encoding elements arising from the field extension.

Theorem 4.2.6. [4] *Let $K = k(\alpha_1, \alpha_2, \dots, \alpha_n)$ be an algebraic extension of k . For each $i = 1, 2, \dots, n$, let $p_i \in k(\alpha_1, \alpha_2, \dots, \alpha_{i-1})[x_i]$ be the minimal polynomial of α_i over $k(\alpha_1, \alpha_2, \dots, \alpha_{i-1})$.*

Let $\beta \in k(\alpha_1, \alpha_2, \dots, \alpha_n)$ be given by

$$\beta = \frac{f(\alpha_1, \alpha_2, \dots, \alpha_n)}{g(\alpha_1, \alpha_2, \dots, \alpha_n)},$$

where $f, g \in k[x_1, x_2, \dots, x_n]$.

Consider the ideal $J = \langle p_1, p_2, \dots, p_n, gy - f \rangle \subset k[x_1, x_2, \dots, x_n, y]$. Then the minimal polynomial of β over k is the monic generator of the elimination ideal $J \cap k[y]$.

Proof. We first show that

$$k[x_1, x_2, \dots, x_n] / \langle p_1, p_2, \dots, p_n \rangle \cong k(\alpha_1, \alpha_2, \dots, \alpha_n).$$

Define the k -algebra homomorphism

$$\varphi_n : k[x_1, x_2, \dots, x_n] \rightarrow k[\alpha_1, \alpha_2, \dots, \alpha_n], \quad x_i \mapsto \alpha_i.$$

Clearly, φ_n is surjective. Hence, by the First Isomorphism Theorem,

$$k[x_1, \dots, x_n] / \ker(\varphi_n) \cong k[\alpha_1, \dots, \alpha_n].$$

We claim that

$$\ker(\varphi_n) = \langle p_1, p_2, \dots, p_n \rangle.$$

Proof of the claim: We proceed by induction on n . The case $n = 1$ is immediate. First, observe that $p_i \in \ker(\varphi_n)$ for each i , since

$$p_i(x_1, \dots, x_i) \mapsto p_i(\alpha_1, \dots, \alpha_i) = 0.$$

Conversely, let $f \in k[x_1, \dots, x_n]$ satisfy $f(\alpha_1, \dots, \alpha_n) = 0$. Define

$$h(x_n) = f(\alpha_1, \alpha_2, \dots, \alpha_{n-1}, x_n) \in k(\alpha_1, \dots, \alpha_{n-1})[x_n].$$

Then

$$h(\alpha_n) = 0,$$

so by definition of the minimal polynomial p_n , we have $p_n \mid h$. Hence,

$$h = p_n \ell_n, \quad \ell_n \in k(\alpha_1, \dots, \alpha_{n-1})[x_n].$$

Now consider

$$f - p_n \ell_n \in k[x_1, \dots, x_n].$$

Writing

$$f - p_n \ell_n = \sum_{\gamma} g_{\gamma}(x_1, \dots, x_{n-1}) x_n^{\gamma},$$

and evaluating at $(\alpha_1, \dots, \alpha_{n-1}, x_n)$ gives

$$(f - p_n \ell_n)(\alpha_1, \dots, \alpha_{n-1}, x_n) = 0.$$

Thus each coefficient satisfies

$$g_{\gamma}(\alpha_1, \dots, \alpha_{n-1}) = 0.$$

Hence,

$$g_{\gamma} \in \ker(\varphi_{n-1}),$$

where

$$\varphi_{n-1} : k[x_1, \dots, x_{n-1}] \rightarrow k(\alpha_1, \dots, \alpha_{n-1}).$$

By the induction hypothesis,

$$g_{\gamma} \in \langle p_1, p_2, \dots, p_{n-1} \rangle.$$

Therefore,

$$f - p_n \ell_n \in \langle p_1, p_2, \dots, p_{n-1} \rangle,$$

and hence

$$f \in \langle p_1, p_2, \dots, p_n \rangle.$$

Thus,

$$\ker(\varphi_n) = \langle p_1, p_2, \dots, p_n \rangle.$$

This completes the proof of the claim. The remainder of the proof follows the same argument as in the previous theorem. $\square$

Example 4.2.7 ([4]). Consider the field extension $\mathbb{Q} \subset \mathbb{Q}(\sqrt{2}, \sqrt[3]{5})$. The minimal polynomial of $\sqrt{2}$ over $\mathbb{Q}$ is $p_1 = x_1^2 - 2 \in \mathbb{Q}[x_1]$, and the minimal polynomial of $\sqrt[3]{5}$ over $\mathbb{Q}(\sqrt{2})$ is $p_2 = x_2^3 - 5 \in \mathbb{Q}(\sqrt{2})[x_2]$. We wish to find the minimal polynomial of $\sqrt{2} + \sqrt[3]{5}$. We compute the reduced Gröbner basis for the ideal

$$J = \langle p_1, p_2, y - (x_1 + x_2) \rangle = \langle x_1^2 - 2, x_2^3 - 5, y - (x_1 + x_2) \rangle \subset \mathbb{Q}[x_1, x_2, y]$$

with respect to the lex term ordering with $x_1 > x_2 > y$ using SINGULAR.

```

> ring R = 0,(x(1),x(2),y),lp;
characteristic : 0
number of vars : 3

block 1 : ordering lp
          names x(1) x(2) y
block 2 : ordering C

> poly p1 = x(1)^2-2;
> poly p2 = x(2)^3-5;
> poly p3 = y-x(1)-x(2);

> ideal J = p1,p2,p3;
> option(redSB);
> ideal G = std(J);
> G = simplify(G,1);
> G;

G[1]=y6-6y4-10y3+12y2-60y+17
G[2]=x(2)+48/1187*y5+45/1187*y4-320/1187*y3
      +780/1187*y2-735/1187*y+1820/1187
G[3]=x(1)-48/1187*y5-45/1187*y4+320/1187*y3
      +780/1187*y2+735/1187*y+1820/1187

```

$$G = \begin{cases} x_1 = \frac{48}{1187}y^5 - \frac{45}{1187}y^4 + \frac{320}{1187}y^3 + \frac{780}{1187}y^2 - \frac{735}{1187}y + \frac{1820}{1187}, \\ x_2 = -\frac{48}{1187}y^5 + \frac{45}{1187}y^4 - \frac{320}{1187}y^3 + \frac{780}{1187}y^2 + \frac{735}{1187}y - \frac{1820}{1187}, \\ y^6 - 6y^4 - 10y^3 + 12y^2 - 60y + 17 \end{cases}$$

Therefore the minimal polynomial of $\sqrt{2} + \sqrt[3]{5}$ over $\mathbb{Q}$ is

$$y^6 - 6y^4 - 10y^3 + 12y^2 - 60y + 17.$$

We also see that $\sqrt{2} + \sqrt[3]{5}$ has degree 6 over $\mathbb{Q}$ and since

$$[\mathbb{Q}(\sqrt{2}, \sqrt[3]{5}) : \mathbb{Q}] = 6,$$

we obtain

$$\mathbb{Q}(\sqrt{2} + \sqrt[3]{5}) = \mathbb{Q}(\sqrt{2}, \sqrt[3]{5}).$$

Example 4.2.8 ([4]). The minimal polynomial of $\sqrt[4]{2}$ over $\mathbb{Q}$ is $p_1 = x_1^4 - 2 \in \mathbb{Q}[x_1]$. The splitting field of the polynomial is $\mathbb{Q}(\sqrt[4]{2}, i)$, where $i^2 = -1$. We consider the element $\beta = 1 + \frac{i}{\sqrt[4]{2}}$. We wish to find the minimal polynomial of β over $\mathbb{Q}$. First, the minimal polynomial of i over $\mathbb{Q}(\sqrt[4]{2})$ is $p_2 = x_2^2 + 1$. Then $\bar{p}_1 = p_1$ and $\bar{p}_2 = p_2$. We consider the ideal

$$J = \langle x_1^4 - 2, x_2^2 + 1, x_1y - (x_1 + x_2) \rangle \subset \mathbb{Q}[x_1, x_2, y].$$

The reduced Gröbner basis G for J with respect to the lex term ordering with $x_1 > x_2 > y$ is computed using SINGULAR.

```

> ring R = 0,(x(1),x(2),y), lp;
> R;
characteristic      : 0
number of vars      : 3

block 1             : ordering lp
                     : names   x(1) x(2) y
block 2             : ordering C

> poly p1 = x(1)^4-2;
> poly p2 = x(2)^2+1;
> poly p3 = x(1)*y-x(1)-x(2);

> ideal J = p1,p2,p3;
> option(redSB);
> ideal G = std(J);
> G = simplify(G,1);
> G;
G[1]=y^4-4*y^3+6*y^2-4*y+1/2
G[2]=x(2)-2+1
G[3]=x(1)-2*x(2)*y^3+6*x(2)*y^2
      -6*x(2)*y+2*x(2)

```

$$G = \left\{ \begin{array}{l} x_1 - 2x_2y^3 + 6x_2y^2 - 6x_2y + 2x_2, \\ x_2^2 + 1, \\ y^4 - 4y^3 + 6y^2 - 4y + \frac{1}{2} \end{array} \right\}.$$

Therefore the minimal polynomial of β is

$$y^4 - 4y^3 + 6y^2 - 4y + \frac{1}{2}.$$

Alter : The minimal polynomial of $\sqrt[4]{2}$ over $\mathbb{Q}$ is $p_1 = x_1^4 - 2 \in \mathbb{Q}[x_1]$. We know that $\mathbb{Q}(\sqrt[4]{2}, i) = \mathbb{Q}(\sqrt[4]{2}, i\sqrt[4]{2}) = \mathbb{Q}(\alpha_1, \alpha_2)$, where $\alpha_1 = \sqrt[4]{2}$ and $\alpha_2 = i\sqrt[4]{2}$. The minimal polynomial of $i\sqrt[4]{2}$ over $\mathbb{Q}(\sqrt[4]{2})$ is $p_2 = x_2^2 + 2$. Thus $\bar{p}_1 = x_1^4 - 2$ and $\bar{p}_2 = x_2^2 + x_1^2$. We again want to compute the minimal polynomial of $\beta = 1 + \frac{i}{\sqrt[4]{2}}$.

Hence,

$$\beta = \frac{i + \sqrt[4]{2}}{\sqrt[4]{2}} = \frac{(\sqrt[4]{2})^2 + i\sqrt[4]{2}}{(\sqrt[4]{2})^2} = \frac{\alpha_1^2 + \alpha_2}{\alpha_1^2} = \frac{f(\alpha_1, \alpha_2)}{g(\alpha_1, \alpha_2)},$$

where $f(x_1, x_2) = x_2^2 + x_2$ and $g(x_1, x_2) = x_1^2$. Therefore, $gy - f = x_1^2y - (x_1^2 + x_2)$.

We consider the ideal $I = \langle x_1^4 - 2, x_2^2 + x_1^2, x_1^2y - (x_1^2 + x_2) \rangle \subset \mathbb{Q}[x_1, x_2, y]$. The reduced Gröbner basis H for I with respect to the lex term ordering with $x_1 > x_2 > y$ is computed by SINGULAR.

```

> ring R = 0,(x1,x2,y),lp;
characteristic : 0
number of vars : 3

block 1 : ordering lp
: names x1 x2 y
block 2 : ordering C

> poly p1 = x1^4-2;
> poly p2 = x2^2+x1^2;
> poly p3 = x1^2*y-(x1^2+x2);

> ideal I = p1,p2,p3;
> option(redSB);
> ideal H = std(I);
> H = simplify(H,1);
> H;

H[1]=y^4-4y^3+6y^2-4y+1/2
H[2]=x2+2y^3-6y^2+6y-2
H[3]=x1^2+2y^2-4y+2

```

Hence,

$$H = \left\{ \begin{array}{l} x_1^2 + 2y^2 - 4y + 2, \\ x_2 + 2y^3 - 6y^2 + 6y - 2, \\ y^4 - 4y^3 + 6y^2 - 4y + \frac{1}{2} \end{array} \right\}.$$

Therefore the minimal polynomial of β is

$$y^4 - 4y^3 + 6y^2 - 4y + \frac{1}{2}.$$

Since $[\mathbb{Q}(\beta) : \mathbb{Q}] = 4$, and since $[\mathbb{Q}(\sqrt[4]{2}, i) : \mathbb{Q}] = 8$, we see that $\mathbb{Q}(\beta)$ is a subfield of $\mathbb{Q}(\sqrt[4]{2}, i)$, but is not equal to it.

Example 4.2.9. Consider the multiquadratic extension field $\mathbb{Q}(\alpha_1, \alpha_2, \alpha_3)$ of $\mathbb{Q}$, where $\alpha_1 = \sqrt{2}$, $\alpha_2 = \sqrt{3}$, and $\alpha_3 = \sqrt{5}$ are defined as the roots of the irreducible polynomials

$$x_1^2 - 2, \quad x_2^2 - 3, \quad x_3^2 - 5,$$

respectively. Now consider the fully symmetric cubic rational element

$$\beta = \frac{\alpha_1^3 + \alpha_2^3 + \alpha_3^3}{\alpha_1 + \alpha_2 + \alpha_3} \in \mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{5}).$$

We wish to find the minimal polynomial of β over $\mathbb{Q}$. We consider the ideal

$$I = \langle x_1^2 - 2, x_2^2 - 3, x_3^2 - 5, x_1y + x_2y + x_3y - x_1^3 - x_2^3 - x_3^3 \rangle$$

in $\mathbb{Q}[x_1, x_2, x_3, y]$. We compute the reduced Gröbner basis H for I with respect to the lexicographic term ordering with

$$x_1 > x_2 > x_3 > y$$

using SINGULAR.

```
> ring R = 0 , (x1,x2,x3,y), lp;
> R;
// coefficients: QQ considered as a field
// number of vars : 4
//      block 1 : ordering lp
//      : names x1 x2 x3 y
//      block 2 : ordering C
> poly p1 = x1^2-2;
> poly p2 = x2^2-3;
> poly p3 = x3^2-5;
> poly p4 = x1*y+x2*y+x3*y-x1^3-x2^3-x3^3;
> ideal I = p1, p2, p3, p4;
> option(redSB);
> ideal H = std(I);
> H;
H[1]=2*y^4-20*y^3+26*y^2+240*y-603
H[2]=x3^2-5
H[3]=15*x2-4*x3*y^3+22*x3*y^2+44*x3*y-267*x3
H[4]=35*x1+4*x3*y^3-18*x3*y^2-54*x3*y+218*x3
```

Therefore the minimal polynomial of β over $\mathbb{Q}$ is

$$2y^4 - 20y^3 + 26y^2 + 240y - 603.$$

Chapter 5

Coloring of Graphs

In this chapter, we present an application of Gröbner basis theory to graph coloring problems. Graph coloring is a classical topic in combinatorics with numerous applications in scheduling, optimization, and theoretical computer science. We show how the problem of determining whether a graph admits a proper coloring can be translated into a system of polynomial equations, thereby allowing algebraic methods to be applied to combinatorial questions. In particular, Gröbner bases provide an effective framework for analyzing the solvability of these polynomial systems and hence for studying graph colorability.

5.1 Graphs and Graph Coloring

We begin this chapter by recalling some basic definitions and terminology from graph theory that will be used throughout. These concepts are standard and can be found in Chapter 1 of Bondy and Murty [13].

Definition 5.1.1. A *graph* is an ordered pair $G = (V, E)$, where V is a non-empty finite set whose elements are called the *vertices* of G , and E is a set of unordered pairs of distinct vertices, called the *edges* of G .

Thus, if $\{u, v\} \in E$, then the vertices u and v are connected by an edge in the graph.

Example 5.1.2. Consider the graph

$$G = (V, E),$$

where

$$V = \{v_1, v_2, v_3, v_4, v_5\},$$

and

$$E = \{\{v_1, v_2\}, \{v_1, v_3\}, \{v_1, v_4\}, \{v_2, v_3\}, \{v_2, v_5\}, \{v_3, v_4\}, \{v_3, v_5\}, \{v_4, v_5\}\}.$$

Then G is a graph with five vertices and eight edges, illustrated below.

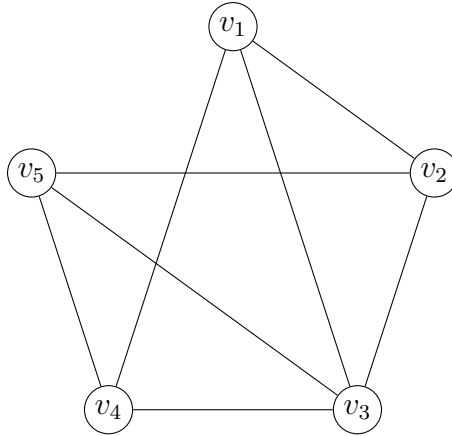


Figure 5.1: A graph with five vertices and eight edges.

Definition 5.1.3. A graph $G = (V, E)$ is called a *simple graph* if it has no loops and no multiple edges; that is, no edge joins a vertex to itself, and no two distinct edges connect the same pair of vertices.

Two vertices $u, v \in V$ are said to be *adjacent* if $\{u, v\} \in E$. In this case, u and v are called *adjacent vertices*.

Example 5.1.4. The graph in the previous example is a simple graph, since it contains no loops and no multiple edges.

Moreover, in that graph:

- v_1 and v_2 are adjacent because $\{v_1, v_2\} \in E$;
- v_3 and v_5 are adjacent because $\{v_3, v_5\} \in E$;
- v_2 and v_4 are not adjacent because $\{v_2, v_4\} \notin E$.

Definition 5.1.5. Let $G = (V, E)$ be a graph and let C be a set of colors. A *graph coloring* of G is a map

$$c : V \rightarrow C,$$

which assigns to each vertex of G a color in C .

A graph coloring is called a *proper graph coloring* if

$$c(u) \neq c(v) \quad \text{whenever } \{u, v\} \in E.$$

That is, adjacent vertices must receive distinct colors.

Example 5.1.6. Consider the graph from Example 5.1.2 and let

$$C = \{\text{Red}, \text{Blue}, \text{Green}\}.$$

Define a coloring $c : V \rightarrow C$ by

$$c(v_1) = \text{Red}, \quad c(v_2) = \text{Blue}, \quad c(v_3) = \text{Green}, \quad c(v_4) = \text{Blue}, \quad c(v_5) = \text{Red}.$$

If every pair of adjacent vertices receives distinct colors under c , then c is a proper graph coloring of G .

5.2 3-Colorability of Graphs

One of the central problems in graph coloring is to determine whether a given graph admits a proper coloring using only three colors. A graph possessing such a coloring is called *3-colorable*. In this section, we study the notion of 3-colorability and prepare for its algebraic characterization.

Let $\omega = e^{2\pi i/3} \in \mathbb{C}$ be a primitive cube root of unity. Then the polynomial $x^3 - 1$ has the three distinct roots $1, \omega, \omega^2$, and hence is separable over $\mathbb{Q}$. We use these three cube roots of unity to represent the three available colors, namely $1, \omega$, and ω^2 .

Let $G = (V, E)$ be a simple graph with vertex set $V = \{x_1, x_2, \dots, x_n\}$. To assign one of the three colors to each vertex x_i , we require that the variable x_i take values among the cube roots of unity. This condition is encoded by

$$x_i^3 - 1 = 0, \quad 1 \leq i \leq n.$$

Now let x_i and x_j be adjacent vertices. Since adjacent vertices must receive distinct colors, we require $x_i \neq x_j$. As both x_i and x_j satisfy $x^3 - 1 = 0$, we have $x_i^3 - x_j^3 = 0$, and hence

$$(x_i - x_j)(x_i^2 + x_i x_j + x_j^2) = 0.$$

Because $x_i \neq x_j$, it follows that

$$x_i^2 + x_i x_j + x_j^2 = 0.$$

Thus, for every edge $\{x_i, x_j\} \in E$, we impose the polynomial equation

$$x_i^2 + x_i x_j + x_j^2 = 0.$$

Accordingly, define the ideal

$$I = \langle x_i^3 - 1 : 1 \leq i \leq n, x_i^2 + x_i x_j + x_j^2 : \{x_i, x_j\} \in E \rangle \subseteq \mathbb{C}[x_1, x_2, \dots, x_n].$$

Then the variety $V(I) \subseteq \mathbb{C}^n$ consists precisely of all assignments of cube roots of unity to the vertices satisfying the above coloring constraints. Consequently, each point of $V(I)$ corresponds to a proper 3-coloring of the graph G .

We now seek to determine when the variety $V(I)$ is non-empty, equivalently, when the graph G admits a proper 3-coloring.

Theorem 5.2.1. *Let $G = (V, E)$ be a simple graph with vertex set $V = \{x_1, x_2, \dots, x_n\}$, and let $I = \langle x_i^3 - 1 : 1 \leq i \leq n, x_i^2 + x_i x_j + x_j^2 : \{x_i, x_j\} \in E \rangle \subseteq \mathbb{C}[x_1, x_2, \dots, x_n]$. Then G is 3-colorable if and only if $V(I) \neq \emptyset$.*

Proof. Suppose first that $V(I) \neq \emptyset$. Then there exists $t = (t_1, t_2, \dots, t_n) \in \mathbb{C}^n$ such that $t \in V(I)$. By definition of $V(I)$, the point t satisfies all generators of the ideal I . In particular,

$$t_i^3 - 1 = 0, \quad 1 \leq i \leq n.$$

Hence each t_i is a cube root of unity, and therefore $t_i \in \{1, \omega, \omega^2\}$ for all $i = 1, 2, \dots, n$.

Assign to each vertex x_i the color represented by t_i . This gives a coloring of the vertices of G using the three colors $1, \omega, \omega^2$.

Now let $\{x_i, x_j\} \in E$. Since $t \in V(I)$, we have

$$t_i^2 + t_i t_j + t_j^2 = 0.$$

If $t_i = t_j$, then $t_i^2 + t_i t_j + t_j^2 = 3t_i^2 \neq 0$, a contradiction. Hence $t_i \neq t_j$, and so adjacent vertices receive distinct colors. Therefore the coloring is proper, and thus G is 3-colorable.

Conversely, suppose that G is 3-colorable. Represent the three colors by $1, \omega, \omega^2$, and let $(v_1, v_2, \dots, v_n) \in \mathbb{C}^n$ be the corresponding coloring, where v_i is the color assigned to the vertex x_i .

Since each v_i is a cube root of unity, we have $v_i^3 - 1 = 0$ for all $1 \leq i \leq n$. Further, if $\{x_i, x_j\} \in E$, then $v_i \neq v_j$ because the coloring is proper. Since v_i and v_j are distinct cube roots of unity, it follows that

$$v_i^2 + v_i v_j + v_j^2 = 0.$$

Thus $(v_1, v_2, \dots, v_n)$ satisfies all generators of I , and hence $(v_1, v_2, \dots, v_n) \in V(I)$. Therefore $V(I) \neq \emptyset$.

Hence G is 3-colorable if and only if $V(I) \neq \emptyset$. □

Theorem [5.2.1](#) reduces the problem of determining whether a graph is 3-colorable to deciding whether the variety $V(I)$ is non-empty. To make this criterion computationally effective, we use the following standard Gröbner basis characterization for emptiness of affine varieties.

Lemma 5.2.2. *Let $k = \bar{k}$ be an algebraically closed field, and let $J \subseteq k[x_1, x_2, \dots, x_n]$ be an ideal. If G is the reduced Gröbner basis of J , then*

$$V(J) = \emptyset \quad \text{if and only if} \quad 1 \in G.$$

Proof. By Corollary [3.1.5](#), $V(J) = \emptyset \iff J = k[x_1, x_2, \dots, x_n]$. Since $J = k[x_1, x_2, \dots, x_n]$ if and only if $1 \in J$, we obtain

$$V(J) = \emptyset \iff 1 \in J.$$

Let G be the reduced Gröbner basis of J . If $1 \in G$, then clearly $1 \in J$ since G generates J .

Conversely, suppose $1 \in J$. Then the remainder upon division of 1 by G must be zero, because every element of J reduces to zero modulo a Gröbner basis of J . Since 1 is a nonzero constant polynomial, this is possible only if some element of G has leading term dividing 1, which can happen only when $1 \in G$.

Therefore $1 \in J$ if and only if $1 \in G$, and hence

$$V(J) = \emptyset \iff 1 \in G.$$

□

Combining Theorem [5.2.1](#) and Lemma [5.2.2](#), we obtain an effective algebraic criterion for deciding whether a graph is 3-colorable. Given a simple graph $G = (V, E)$, one first constructs the associated ideal

$$I = \langle x_i^3 - 1 : 1 \leq i \leq n, x_i^2 + x_i x_j + x_j^2 : \{x_i, x_j\} \in E \rangle \subseteq \mathbb{C}[x_1, x_2, \dots, x_n].$$

One then computes a reduced Gröbner basis G of the ideal I . The graph G is 3-colorable if and only if $1 \notin G$. Equivalently, the graph is not 3-colorable if and only if the reduced Gröbner basis of I contains the polynomial 1.

Example 5.2.3. ([4]) Consider the graph $G = (V, E)$ shown below, where

$$V = \{x_1, x_2, \dots, x_8\},$$

and

$$E = \{\{x_1, x_2\}, \{x_1, x_5\}, \{x_1, x_6\}, \{x_2, x_3\}, \{x_2, x_4\}, \{x_2, x_8\}, \\ \{x_3, x_4\}, \{x_3, x_8\}, \{x_4, x_5\}, \{x_4, x_7\}, \{x_5, x_6\}, \{x_5, x_7\}, \\ \{x_6, x_7\}, \{x_7, x_8\}\}.$$

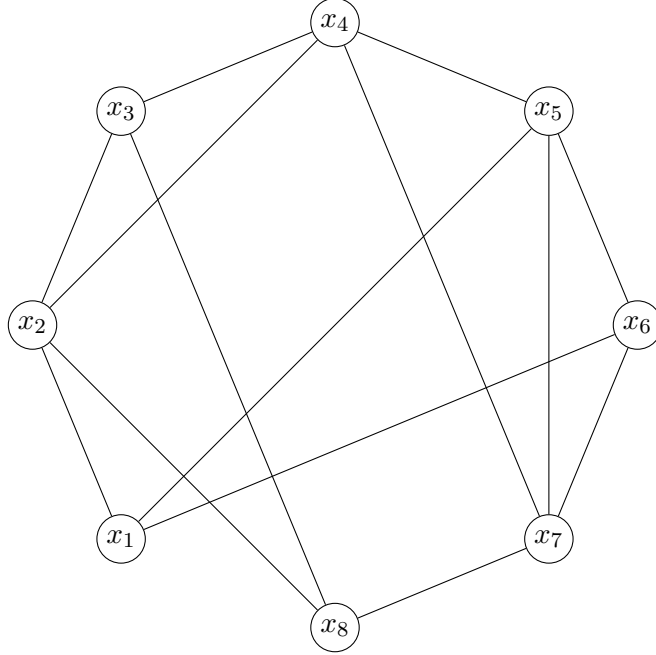


Figure 5.2: Graph whose 3-colorability is analyzed in this example.

Represent the three colors red, green, and blue by the cube roots of unity $1, \omega, \omega^2$, where $\omega = e^{2\pi i/3}$. To encode a 3-coloring algebraically, for each vertex x_i we impose the condition

$$x_i^3 - 1 = 0, \quad 1 \leq i \leq 8,$$

so that each vertex variable takes one of the three cube roots of unity.

Further, for every edge $\{x_i, x_j\} \in E$, we impose the relation

$$x_i^2 + x_i x_j + x_j^2 = 0,$$

which guarantees that adjacent vertices receive distinct colors.

Hence the graph determines the ideal

$$I = \langle x_i^3 - 1 : 1 \leq i \leq 8, x_i^2 + x_i x_j + x_j^2 : \{x_i, x_j\} \in E \rangle \subseteq \mathbb{C}[x_1, x_2, \dots, x_8].$$

Since each variable x_i can assume at most three values, we have $|V(I)| \leq 3^8$. Therefore $V(I)$ is finite, and consequently I is a zero-dimensional ideal.

Let $G = \{g_1, g_2, \dots, g_t\}$ be the reduced Gröbner basis of I with respect to the lexicographic term order

$$x_1 > x_2 > \dots > x_8.$$

Since I is zero-dimensional, the Elimination Theorem implies that the elements of G may be ordered in triangular form so that:

- g_1 involves only the variable x_8 ,
- g_2 involves only the variables x_7, x_8 with $\text{LT}(g_2)$ a power of x_7 ,
- g_3 involves only the variables x_6, x_7, x_8 with $\text{LT}(g_3)$ a power of x_6 ,

and so on.

Therefore, if $1 \notin G$, then $V(I) \neq \emptyset$, and one can recursively solve the triangular system determined by the Gröbner basis to obtain a proper 3-colouring of the graph. Using SINGULAR, we compute the reduced Gröbner basis G for the ideal I , obtaining

$$G = \{x_8^3 - 1, x_7^2 + x_7x_8 + x_8^2, x_6 - x_8, x_5 + x_7 + x_8, x_4 - x_8, x_3 - x_7, x_2 + x_7 + x_8, x_1 - x_7\}.$$

Since $1 \notin G$, we conclude that $V(I) \neq \emptyset$, and hence by the previous theorem the graph is 3-colourable.

We now use the reduced Gröbner basis G to obtain an explicit colouring, since the system of equations represented by G is easy to solve.

The polynomial $x_8^3 - 1 \in G$ involves only the variable x_8 . First, we fix a colour for x_8 . Let $x_8 = 1$. Since $x_4 - x_8 \in G$ and $x_6 - x_8 \in G$, we obtain $x_4 = x_6 = 1$. Next, from $x_7^2 + x_7x_8 + x_8^2 \in G$, we see that x_7 must be a cube root of unity different from 1. Hence we take $x_7 = \omega$. Since $x_1 - x_7 \in G$ and $x_3 - x_7 \in G$, it follows that $x_1 = x_3 = \omega$. From $x_2 + x_7 + x_8 \in G$, we obtain $x_2 = -x_7 - x_8 = -\omega - 1 = \omega^2$. Similarly, from $x_5 + x_7 + x_8 \in G$, we get $x_5 = \omega^2$. Thus, we obtain a valid assignment of colours.

Finally, we interpret the cube roots of unity as colours:

$$1 \equiv \text{red}, \quad \omega \equiv \text{green}, \quad \omega^2 \equiv \text{blue}.$$

Hence, one proper 3-colouring of the graph $G = (V, E)$ is obtained as shown in the following figure:

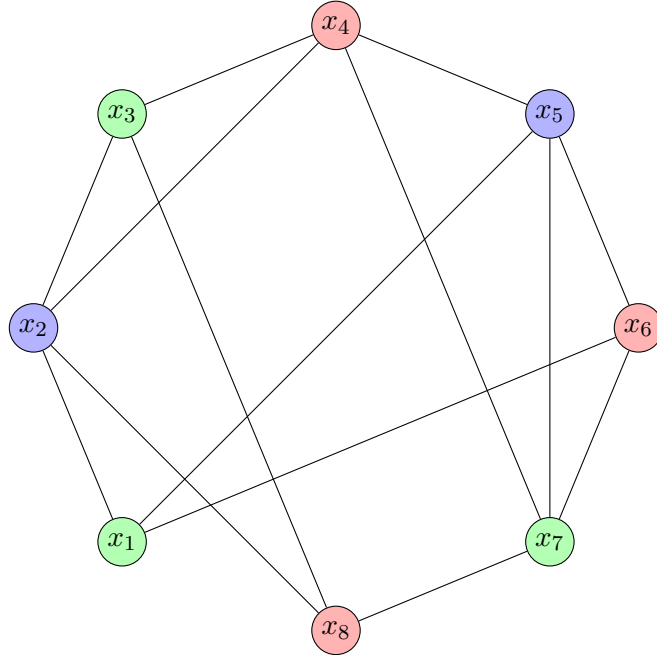


Figure 5.3: A proper 3-colouring of the graph obtained from the Gröbner basis solution.

We observe from the reduced Gröbner basis G in the above example that the 3-colouring of the graph is unique up to a permutation of the colours.

However, if a graph admits multiple distinct 3-colourings, then the corresponding Gröbner basis typically has a more complicated structure.

This phenomenon is illustrated in the following example.

Example 5.2.4. ([4]) Consider the graph $G = (V, E)$ shown below, where $V = \{x_1, x_2, \dots, x_9\}$, and

$$E = \{\{x_1, x_2\}, \{x_1, x_4\}, \{x_1, x_5\}, \{x_1, x_6\}, \{x_2, x_3\}, \{x_2, x_5\}, \{x_2, x_7\}, \\ \{x_3, x_4\}, \{x_3, x_6\}, \{x_3, x_9\}, \{x_5, x_6\}, \{x_6, x_8\}, \{x_7, x_8\}, \{x_8, x_9\}\}.$$

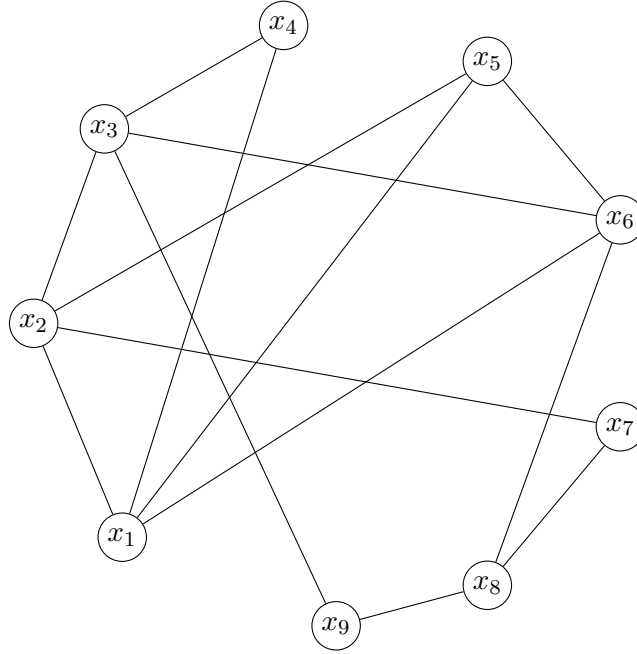


Figure 5.4: Graph whose 3-colorability is analyzed in this example.

We associate to each vertex the relation $x_i^3 - 1 = 0$ for $i = 1, \dots, 9$, and for each edge $(i, j) \in E$ the relation $x_i^2 + x_i x_j + x_j^2 = 0$. Let $I \subseteq \mathbb{C}[x_1, \dots, x_9]$ be the ideal generated by these polynomials.

We compute a Gröbner basis of I with respect to the lexicographic order $x_1 > x_2 > \dots > x_9$.

Using SINGULAR, the reduced Gröbner basis $G = \{g_1, \dots, g_{12}\}$ is obtained as:

$$\begin{aligned} g_1 &= x_9^3 - 1, \\ g_2 &= x_8^2 + x_8 x_9 + x_9^2, \\ g_3 &= (x_7 - x_9)(x_7 + x_8 + x_9), \\ g_4 &= x_6 + x_7 + x_8, \\ g_5 &= (x_5 - x_7)(x_5 - x_8), \\ g_6 &= (x_5 - x_9)(x_4 x_7 + x_4 x_8 + x_7 x_8 + x_4 x_9 + x_7 x_9 + x_8 x_9), \\ g_7 &= (x_4 - x_5)(x_4 + x_7 + x_8), \\ g_8 &= x_3 x_7 + x_3 x_8 - x_7 x_8 + x_3 x_9 + x_9^2, \end{aligned}$$

$$\begin{aligned}
g_9 &= x_3x_4 + x_4x_5 - x_4x_7 + x_5x_7 - x_4x_8 + x_5x_8 - x_3x_9 - x_9^2, \\
g_{10} &= x_3^2 + x_3x_9 + x_9^2, \\
g_{11} &= x_2 + x_7 + x_8, \\
g_{12} &= x_1 + x_5 - x_7 - x_8.
\end{aligned}$$

Since $1 \notin G$, we conclude that $V(I) \neq \emptyset$, and hence the graph is 3-colourable.

We now construct an explicit solution. Let $x_9 = 1$. From g_2 , choose $x_8 = \omega$. From g_3 , take $x_7 = 1$. Then from g_4 , we get $x_6 = \omega^2$. From g_5 , $x_5 = 1$, and from g_7 , $x_4 = 1$. From g_{11} , we obtain $x_2 = \omega^2$, and from g_{12} , $x_1 = \omega$. Finally, from g_8 , we get $x_3 = \omega$.

Thus the solution is $x = (\omega, \omega^2, \omega, 1, 1, \omega^2, 1, \omega, 1)$.

Interpreting $1 \equiv \text{red}$, $\omega \equiv \text{green}$, and $\omega^2 \equiv \text{blue}$, we obtain a proper 3-colouring of the graph G , as shown in the following figure.

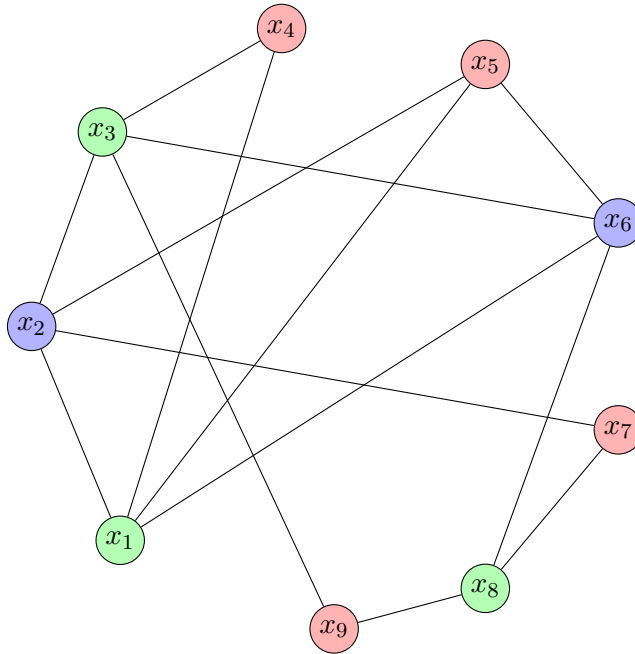


Figure 5.5: A proper 3-colouring of the graph obtained from the Gröbner basis solution.

5.3 k-Colorability of Graphs

In this section, we extend the algebraic approach developed for 3-colorability to the general case of k -colorability of graphs. The aim is to describe how the problem of assigning k colors to the vertices of a graph can be translated into the language of polynomial ideals, and how Gröbner basis techniques can be used to study the existence of such colorings.

Lemma 5.3.1. [3, *Polynomials over fields, Chapter 14*] Let R be a domain and let $f \in R[t]$. Suppose $a \in R$ is a root of f . Then a is a multiple root of f if and only if $(t - a)$ divides both f and f' .

Proof. Let $l \geq 1$ be the multiplicity of the root a . Then we can write

$$f(t) = (t - a)^l g(t),$$

where $g(t) \in R[t]$ and $g(a) \neq 0$.

Differentiating, we obtain

$$f'(t) = l(t-a)^{l-1}g(t) + (t-a)^l g'(t).$$

First suppose that a is a multiple root of f . Then $l \geq 2$. In this case, both terms in the expression of $f'(t)$ contain the factor $(t-a)$, so $(t-a)$ divides $f'(t)$. Since $l \geq 1$, it also follows that $(t-a)$ divides $f(t)$.

Conversely, suppose that $(t-a)$ divides both f and f' . Then $f(a) = 0$ and $f'(a) = 0$. Writing $f(t) = (t-a)^l g(t)$ with $g(a) \neq 0$, the condition $f'(a) = 0$ implies that $l \geq 2$. Hence a is a multiple root of f . $\square$

Theorem 5.3.2. [3, *Polynomials over fields, Chapter 14*] Let F be a field whose characteristic does not divide a positive integer n , and let E be the splitting field of $f(t) = t^n - 1$ over F . Then E contains n distinct n th roots of unity. In particular, the polynomial $f(t)$ is separable over F .

Proof. Let $f(t) = t^n - 1 \in F[t]$. Then its formal derivative is given by $f'(t) = nt^{n-1}$.

Since $\text{char}(F)$ does not divide n , the element n is nonzero in F , and hence $f'(t) \neq 0$. Moreover, any common divisor of f and f' must divide both $t^n - 1$ and t^{n-1} . This forces $\gcd(f, f') = 1$, so f and f' are relatively prime in $F[t]$.

It follows that $f(t)$ has no repeated roots in its splitting field E . Therefore, all roots of f in E are distinct.

Since a splitting field contains all roots of the polynomial, E contains exactly n distinct roots of $t^n - 1$, namely the n th roots of unity. Hence f is separable over F . $\square$

Definition 5.3.3. Let $G = (V, E)$ be a simple graph with $V = \{1, 2, \dots, n\}$. The graph polynomial of G is defined as

$$f_G = \prod_{(i,j) \in E, i < j} (x_i - x_j).$$

Let F be an algebraically closed field whose characteristic does not divide a positive integer k . Then F contains k distinct k th roots of unity.

Consider the polynomial ring $F[x_1, \dots, x_n]$. To encode k -colourability algebraically, we associate to G the following ideals.

First, we ensure that each vertex variable takes values among the k th roots of unity by defining

$$I_{n,k} = \langle x_i^k - 1 : i = 1, 2, \dots, n \rangle.$$

Next, for each edge $(i, j) \in E$, we impose the condition that adjacent vertices receive distinct k th roots of unity. This is encoded by the polynomial

$$x_i^{k-1} + x_i^{k-2}x_j + \dots + x_i x_j^{k-2} + x_j^{k-1}.$$

Thus, we define the graph ideal

$$I_{G,k} = I_{n,k} + \langle x_i^{k-1} + x_i^{k-2}x_j + \dots + x_i x_j^{k-2} + x_j^{k-1} : (i, j) \in E \rangle.$$

This construction provides an algebraic framework for studying k -colourability of graphs using polynomial ideals and Gröbner basis techniques.

Proposition 5.3.4 ([15, Lemma 3.1, Section 3, p. 405]). *The varieties $V(I_{n,k})$, $V(I_{G,k})$, and $V(I_{n,k} + \langle f_G \rangle)$ are in bijection with all k -colourings, proper k -colourings, and improper k -colourings of G , respectively.*

Proof. We prove the three parts separately.

(1) The variety $V(I_{n,k})$ consists of all points $v = (v_1, \dots, v_n) \in F^n$ such that $v_i^k = 1$ for all i . Hence each coordinate v_i is a k th root of unity. Therefore, each such point corresponds naturally to an arbitrary assignment of k colours to the vertices of G , that is, to all k -colourings of G .

(2) Let $v = (v_1, \dots, v_n) \in V(I_{G,k})$. Then $v \in V(I_{n,k})$, so each v_i is a k th root of unity. Moreover, for each edge $\{i, j\} \in E$, we have

$$p_{ij}(x) = x_i^{k-1} + x_i^{k-2}x_j + \dots + x_j^{k-1} \in I_{G,k},$$

and hence $p_{ij}(v) = 0$.

If $v_i = v_j$, then

$$p_{ij}(v) = kv_i^{k-1} = 0,$$

which contradicts the assumption that $\text{char}(F)$ does not divide k . Therefore $v_i \neq v_j$ whenever $\{i, j\} \in E$, and thus v defines a proper k -colouring of G .

Conversely, let $v = (v_1, \dots, v_n)$ be a proper k -colouring of G , so each v_i is a k th root of unity and $v_i \neq v_j$ for every edge $\{i, j\}$. Then $v \in V(I_{n,k})$.

For each edge $\{i, j\} \in E$, we note that

$$p_{ij}(v)(v_i - v_j) = v_i^k - v_j^k = 1 - 1 = 0.$$

Since $v_i \neq v_j$, it follows that $p_{ij}(v) = 0$. Hence $v \in V(I_{G,k})$.

Thus, proper k -colourings of G are in bijection with $V(I_{G,k})$.

(3) Let $v \in V(I_{n,k} + \langle f_G \rangle)$. Then $v \in V(I_{n,k})$, so each coordinate is a k th root of unity, and additionally $f_G(v) = 0$. Since

$$f_G = \prod_{(i,j) \in E} (x_i - x_j),$$

there exists at least one edge $\{i, j\} \in E$ such that $v_i = v_j$. Hence the corresponding colouring is not proper, and so v represents an improper k -colouring.

Conversely, if v is an improper k -colouring, then $v \in V(I_{n,k})$ and there exists an edge $\{i, j\} \in E$ such that $v_i = v_j$. This implies $f_G(v) = 0$, and hence $v \in V(I_{n,k} + \langle f_G \rangle)$. $\square$

Lemma 5.3.5 ([15, Lemma 3.2, Section 3, p. 406]). *The ideals $I_{n,k}$, $I_{G,k}$, and $I_{n,k} + \langle f_G \rangle$ are radical ideals.*

Proof. The proof follows from Lemma 3.2 in [15]. $\square$

Lemma 5.3.6 ([1, Corollary 11, Section 4, p. 204]). *Let k be a field and consider the polynomial ring $k[x_1, \dots, x_n]$. Let V and W be varieties in k^n . Then*

$$I(V) : I(W) = I(V \setminus W).$$

Proof. See [1, Corollary 11, Section 4, p. 204]. $\square$

Lemma 5.3.7 ([15, Lemma 3.3, Section 3, p. 406]). *We have*

$$I_{n,k} : I_{G,k} = I_{n,k} + \langle f_G \rangle.$$

Proof. Let V denote the set of all k -colourings of the graph G , and let W denote the set of all proper k -colourings of G .

By Proposition 5.3.4, we have

$$V = V(I_{n,k}), \quad W = V(I_{G,k}),$$

and the set difference $V \setminus W$ corresponds to the set of improper k -colourings. Again by Proposition 5.3.4, this implies

$$V \setminus W = V(I_{n,k} + \langle f_G \rangle).$$

Since F is algebraically closed, we apply Hilbert's Nullstellensatz and Lemma 5.3.5 to obtain

$$I(V) = I(V(I_{n,k})) = I_{n,k}, \quad I(W) = I(V(I_{G,k})) = I_{G,k},$$

and

$$I(V \setminus W) = I(V(I_{n,k} + \langle f_G \rangle)) = I_{n,k} + \langle f_G \rangle.$$

Now using Lemma 5.3.6, we have the identity

$$I(V) : I(W) = I(V \setminus W).$$

Substituting the above identifications yields

$$I_{n,k} : I_{G,k} = I_{n,k} + \langle f_G \rangle.$$

This completes the proof. □

Definition 5.3.8. Let $G = (V, E)$ be a finite graph. The *chromatic polynomial* of G , denoted by $\chi_G(k)$, is defined as the number of proper k -colourings of G , i.e.,

$$\chi_G(k) = \#\{\text{proper } k\text{-colourings of } G\},$$

for every positive integer k .

Illustration 5.3.9. Using the Finiteness Theorem (see [14]), we can summarize all the results discussed above.

$$\#(k\text{-colourings}) = |V(I_{n,k})| = \dim_F \frac{F[x_1, \dots, x_n]}{I_{n,k}} = k^n,$$

$$\#(\text{proper } k\text{-colourings}) = |V(I_{G,k})| = \dim_F \frac{F[x_1, \dots, x_n]}{I_{G,k}},$$

$$\#(\text{improper } k\text{-colourings}) = |V(I_{n,k} + \langle f_G \rangle)| = \dim_F \frac{F[x_1, \dots, x_n]}{I_{n,k} + \langle f_G \rangle}.$$

We now summarize the algebraic framework developed in this section for studying k -colourability of a graph G . We have associated to G two ideals $I_{n,k}$ and $I_{G,k}$ in the polynomial ring $F[x_1, \dots, x_n]$, together with the graph polynomial f_G . The variety $V(I_{n,k})$ corresponds to all k -colourings of G , while $V(I_{G,k})$ corresponds to proper k -colourings. Using Hilbert's Nullstellensatz and the Finiteness Theorem, these combinatorial sets are translated into algebraic objects such as quotient vector spaces and ideal-theoretic conditions. In particular, the structure of these ideals completely determines whether the graph admits a proper k -colouring or not. This leads to the following equivalent algebraic characterizations of non- k -colourability.

Theorem 5.3.10 ([15, Theorem 1.1, Section 1, p. 401]). *The following statements are equivalent:*

- (i) *The graph G is not k -colourable.*
- (ii) $\dim_F \frac{F[x_1, \dots, x_n]}{I_{G,k}} = 0$.
- (iii) $1 \in I_{G,k}$.
- (iv) $f_G \in I_{n,k}$.

Proof. (i) (1) $\Rightarrow$ (2): Suppose that G is not k -colourable. Then $\chi_G(k) = 0$. Hence $\dim_F \frac{F[x_1, \dots, x_n]}{I_{G,k}} = 0$.

(ii) (2) $\Rightarrow$ (3): If the dimension is 0, then the quotient space is trivial. Hence $I_{G,k} = F[x_1, \dots, x_n]$, so $1 \in I_{G,k}$.

(iii) (3) $\Rightarrow$ (4): Assume $1 \in I_{G,k}$. Then $I_{G,k} = F[x_1, \dots, x_n]$. Using Lemma 5.3.6,

$$I_{n,k} : I_{G,k} = I_{n,k} + \langle f_G \rangle,$$

so $f_G \in I_{n,k}$.

(iv) (4) $\Rightarrow$ (1): Suppose $f_G \in I_{n,k}$. Then $I_{n,k} + \langle f_G \rangle = I_{n,k}$, which implies $\chi_G(k) = 0$. Hence G is not k -colourable. □

Example 5.3.11. Consider the graph $G = (V, E)$ shown below, where

$$V = \{x_1, x_2, \dots, x_{10}\},$$

and

$$\begin{aligned} E = \{ & \{x_1, x_2\}, \{x_1, x_4\}, \{x_1, x_5\}, \{x_2, x_3\}, \{x_2, x_4\}, \\ & \{x_3, x_4\}, \{x_3, x_8\}, \{x_4, x_9\}, \{x_5, x_6\}, \{x_5, x_7\}, \{x_6, x_7\}, \\ & \{x_6, x_9\}, \{x_7, x_8\}, \{x_7, x_{10}\}, \{x_8, x_9\}, \{x_8, x_{10}\}, \{x_9, x_{10}\}, \\ & \{x_1, x_{10}\}, \{x_2, x_9\} \}. \end{aligned}$$

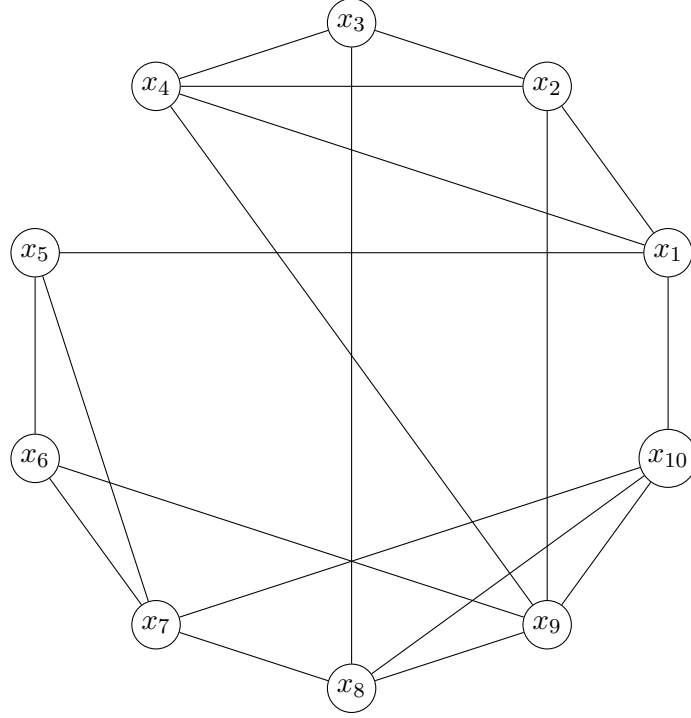


Figure 5.6: Graph whose 4-colorability is analyzed in this example.

Represent the four colors red, green, blue, and yellow by the fourth roots of unity $1, i, -1, -i$. To encode a 4-coloring algebraically, for each vertex x_j we impose the condition

$$x_j^4 - 1 = 0, \quad 1 \leq j \leq 10,$$

so that each vertex variable takes one of the four fourth roots of unity. Further, for every edge $\{x_j, x_k\} \in E$, we impose the relation

$$x_j^3 + x_j^2 x_k + x_j x_k^2 + x_k^3 = 0,$$

which guarantees that adjacent vertices receive distinct colors. Hence the graph determines the ideal

$$I = \langle x_j^4 - 1 : 1 \leq j \leq 10, x_j^3 + x_j^2 x_k + x_j x_k^2 + x_k^3 : \{x_j, x_k\} \in E \rangle \subseteq \mathbb{C}[x_1, x_2, \dots, x_{10}].$$

Using SINGULAR, the reduced Gröbner basis $G = \{g_1, \dots, g_{10}\}$ under the lexicographic order $x_1 > x_2 > \dots > x_{10}$ is obtained as:

$$\begin{aligned} g_1 &= x_{10}^4 - 1, \\ g_2 &= x_9^3 + x_9^2 x_{10} + x_9 x_{10}^2 + x_{10}^3, \\ g_3 &= x_8^2 + x_8 x_9 + x_8 x_{10} + x_9^2 + x_9 x_{10} + x_{10}^2, \\ g_4 &= x_7 + x_8 + x_9 + x_{10}, \\ g_5 &= x_6 - x_8, \\ g_6 &= x_5 - x_9, \\ g_7 &= x_4 - x_{10}, \\ g_8 &= x_3 - x_7, \\ g_9 &= x_2 - x_8, \\ g_{10} &= x_1 - x_7. \end{aligned}$$

Since $1 \notin G$, we conclude that $V(I) \neq \emptyset$, meaning the graph is 4-colourable.

From $g_1 = x_{10}^4 - 1$, we select the primary color root: $x_{10} = 1$. Substituting this into g_2 yields $x_9^3 + x_9^2 + x_9 + 1 = 0$, from which we choose $x_9 = i$. From g_3 , substituting $x_{10} = 1$ and $x_9 = i$ yields $x_8^2 + (1+i)x_8 + i = 0$, which factors into $(x_8 + 1)(x_8 + i) = 0$. We select: $x_8 = -1$.

From g_4 , we calculate $x_7 = -x_8 - x_9 - x_{10} = -(-1) - i - 1 = -i$. The remaining variables collapse linearly from the single-variable polynomials in the basis:

- From g_5 : $x_6 = x_8 = -1$.
- From g_6 : $x_5 = x_9 = i$.
- From g_7 : $x_4 = x_{10} = 1$.
- From g_8 : $x_3 = x_7 = -i$.
- From g_9 : $x_2 = x_8 = -1$.
- From g_{10} : $x_1 = x_7 = -i$.

Mapping our final coordinates back to the color key $1 \equiv \text{red}$, $i \equiv \text{green}$, $-1 \equiv \text{blue}$, and $-i \equiv \text{yellow}$, we construct a flawless proper 4-colouring where no connected vertices share a color:

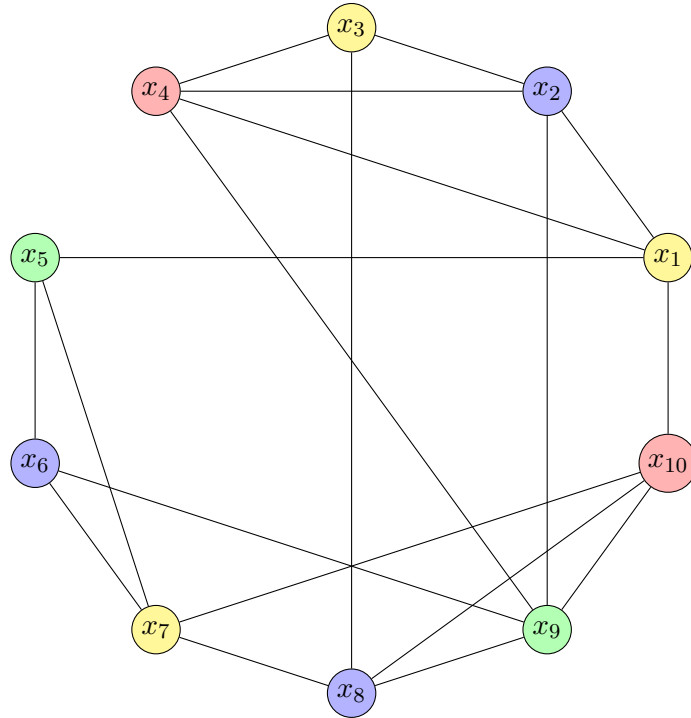


Figure 5.7: A proper 4-colouring of the graph obtained from the Gröbner basis solution

Bibliography

- [1] David A. Cox, John Little, and Donal O'Shea, *Ideals, Varieties, and Algorithms*, 4th Edition, Springer, 2015.
- [2] D. A. Cox, J. B. Little, and D. O'Shea, *Using Algebraic Geometry*, Springer, 2nd edition, 2005.
- [3] D. S. Dummit and R. M. Foote, *Abstract Algebra*, 3rd edition, Wiley India, 2004.
- [4] B. Hajra, *Gröbner Bases and Their Applications*, Master of Science Project Report, Department of Mathematics, Indian Institute of Technology Bombay, 2016.
- [5] D. Eisenbud, *Commutative Algebra with a View Toward Algebraic Geometry*, 3rd corrected printing, Springer, New York-Berlin-Heidelberg, 1999.
- [6] T. Becker and V. Weispfenning, *Gröbner Bases*, Springer, New York-Berlin-Heidelberg, 1993.
- [7] M. Kreuzer and L. Robbiano, *Computational Commutative Algebra*, Vol. 1, Springer, New York, 2000.
- [8] L. E. Dickson, *Finiteness of the Odd Perfect and Primitive Abundant Numbers with n Distinct Prime Factors*, American Journal of Mathematics, Vol. 35, No. 4, 1913, pp. 413–422.
- [9] D. Hilbert, *Über die Theorie der algebraischen Formen*, Mathematische Annalen, Vol. 36, 1890, pp. 473–534.
- [10] D. Hilbert, *Ueber die vollen Invariantensysteme*, Mathematische Annalen, Vol. 42, No. 3, 1893, pp. 313–373.
- [11] B. Buchberger, *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal*, Ph.D. Thesis, University of Innsbruck, Austria, 1965.
- [12] L. Stickelberger, *Ueber eine neue Eigenschaft der Diskriminanten algebraischer Zahlkörper*, in Verhandlungen des ersten Internationalen Mathematiker-Kongresses, Zürich, 1897, pp. 248–253. Leipzig: B. G. Teubner, 1898.
- [13] J. A. Bondy and U. S. R. Murty, *Graph Theory*, Graduate Texts in Mathematics, Vol. 244, Springer, New York, 2008.
- [14] Jugal K. Verma, *Solving Polynomial Equations*, 3 lectures, Miniworkshop on Algebra and Number Theory, Kumaun University, Almora, March 2012.

- [15] C. J. Hillar and T. Windfeldt, *An algebraic characterization of uniquely vertex colorable graphs*, Journal of Combinatorial Theory, Series B, 98(2) (2008), 400–414.
- [16] W. Decker and C. Lossen, *Computing in Algebraic Geometry: A Quick Start using SINGULAR*, Algorithms and Computation in Mathematics, Vol. 16, Springer, Berlin, 2006.

