COMPUTATIONAL ALGEBRA
WITH APPLICATIONS

Rocky Mountain Mathematics
Consortium

Summer Conference
University of Wyoming
June 21–July 2, 1999

Joseph Brennan
North Dakota State University
and
Reinhard Laubenbacher
New Mexico State University
Contents

Chapter 1. Gröbner Basis Theory 5
1. Introduction 5
2. Term Orders 7
3. Gröbner Bases 9
4. The Buchberger Algorithm 16
5. Applications of Gröbner Bases to Ideal Theory 20
6. Gröbner Bases for Binomial Ideals 25
7. Gröbner Bases for Modules 26

Chapter 2. A Brief Introduction to Algebraic Geometry 31
1. Affine Varieties 31
2. The Coordinate Ring of an Affine Variety 36

Chapter 3. Application to the Solution of Polynomial Systems 39
1. Introduction 39
2. Elimination Theory 39
3. Zero-dimensional Systems 43

Chapter 4. Application to Cryptography 45
1. Introduction 45
2. The Polly Cracker System 46
3. A Combinatorial Version of Polly Cracker 46
4. A Generalization 47

Bibliography 49
CHAPTER 1

Gröbner Basis Theory

1. Introduction.

In this chapter we give a survey of the fundamentals of Gröbner basis theory and its use in computational algebra, to the extent needed later on in the conference.

Our main object of study will be the ring of polynomials

\[ f = \sum_\alpha r_\alpha x^\alpha \]

in variables \( x_1, \ldots, x_n \) with coefficients in a field \( k \). Here, \( \alpha \) is an \( n \)-tuple \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \) of nonnegative integers, and \( x^\alpha \) denotes the monomial \( x_1^{\alpha_1}x_2^{\alpha_2}\cdots x_n^{\alpha_n} \). The coefficients \( r_\alpha \) are elements of the field \( k \), and there are only finitely many nonzero terms \( r_\alpha x^\alpha \) in the sum. That is,

\[ k[x_1, \ldots, x_n] = \left\{ \sum_\alpha r_\alpha x^\alpha | \alpha \in \mathbb{N}^n, r_\alpha \in k \right\}. \]

In most cases of interest to us, \( k \) will be either a finite field, or one of the fields \( \mathbb{Q}, \mathbb{R}, \mathbb{C} \). In general we will abbreviate \( k[x_1, \ldots, x_n] \) by \( k[x] \).

For the case of one variable \( x \), we all know the ring \( k[x] \) from high school. The most important property of this ring is that there is a division algorithm for polynomials in one variable. Given nonzero polynomials \( f, g \in k[x] \), with \( \deg(g) \leq \deg(f) \), there exist unique polynomials \( q, r \in k[x] \) such that \( f = qg + r \), and either \( r = 0 \) or \( \deg(r) < \deg(g) \). This algorithm has as one consequence that \( k[x] \) is a principal ideal domain (Exercise 1.1). In particular, the division algorithm allows the solution of the so-called ideal membership problem. Given an ideal \( I \subset k[x] \) and a polynomial \( f \in k[x] \), decide whether \( f \in I \). Namely, \( I \) is principal, generated by a polynomial \( g \), and \( f \in I = (g) \) if and only if \( g \) divides \( f \) with zero remainder.

Once we consider polynomial rings of several variables, things get a bit more complicated. First of all, in general ideals are not generated by just one polynomial anymore (but finitely many), and there are
problems with the division algorithm for several variables as well. The general ideal membership problem is formulated as follows.

**Ideal Membership Problem.** Let

\[ I = \langle f_1, \ldots, f_m \rangle \subseteq k[x_1, \ldots, x_n] \]

be an ideal and \( f \in k[x_1, \ldots, x_n] \). Decide whether \( f \in I \), that is, whether there are polynomials \( g_1, \ldots, g_m \) such that

\[ f = f_1g_1 + \cdots + f_mg_m. \]

This problem could be solved just as in the one-variable case, if we had a similar division algorithm. First of all, we need to specify an ordering on the terms of the polynomials \( f \) and \( g \). (In one variable, we tacitly assumed that the terms were ordered according to highest degree, and the division algorithm starts first with this highest term.) Since the coefficients play no role in this ordering, it is enough to order all monomials.

One common ordering is the lexicographic ordering. Here we order the variables first, for instance \( x_1 > x_2 > \cdots > x_n \). Then two monomials \( x^\alpha \) and \( x^\beta \) are compared by comparing their exponent vectors \( \alpha \) and \( \beta \) lexicographically, that is, \( x^\alpha < x^\beta \) for n-tuples \( \alpha = (\alpha_1, \ldots, \alpha_n) \) and \( \beta = (\beta_1, \ldots, \beta_n) \) if there is \( 1 \leq i \leq n \), such that \( \alpha_j = \beta_j \) for all \( j < i \), and \( \alpha_i < \beta_i \). For instance, for two variables \( x > y \), we have that \( x > y^{10} \), and \( x^2y^3 > x^2y^2 \). We will discuss other orderings in the next section.

There is still a division algorithm in the general case, and it works essentially like in the one-variable case. First choose the leading terms of \( f_1, \ldots, f_m \) and \( f \), that is, the terms that are largest under the chosen term order. Then proceed as in the one-variable case to write \( f = f_1g_1 + \cdots + f_mg_m + r \). That is, first divide \( f \) by \( f_1 \), then the remainder by \( f_2 \), and so on. The complication that arises is that the remainder \( r \) is no longer unique in general and depends on all the choices we have made, such as term order and the ordering of the \( f_i \).

**Example 1.1** Consider the following example for two variables \( x \) and \( y \), using the above lexicographic order with \( x > y \). In \( k[x, y] \), let \( f_1 = x^2 - 1, f_2 = xy + 2, \) and \( f = x^2y + xy + 2x + 2 \). If we carry out our division procedure, we obtain that

\[ f = yf_1 + f_2 + 2x + y. \]

Since \( 2x + y \) cannot be reduced further, we might conclude that \( f \) is not in the ideal \( \langle f_1, f_2 \rangle \). However, if we change the roles of \( f_1 \) and \( f_2 \), then we obtain that

\[ f = (x + 1)f_1 = (x + 1)(xy + 2). \]
It is easy to make examples where a similar phenomenon happens when we change term orders. It thus seems that we have lost the most powerful tool available in the one-variable case. Fortunately, not all is lost. While the division algorithm cannot be improved in general, there is still a way to solve our ideal membership problem. It turns out that if we use certain very special generating sets for our ideals, then it is still true that \( f \in \langle f_1, \ldots, f_m \rangle \) if and only if the remainder \( r \) in the above division algorithm is zero. Such a special generating set for an ideal is called a \textit{Gröbner basis}. This concept provides the work horse for many symbolic algorithms in commutative algebra and algebraic geometry.

Gröbner bases have been used for some time, as early as the beginning of this century. For a brief history see [5, pp. 337–338]. The modern story begins with Wolfgang Gröbner, a professor at the University of Innsbruck, Austria. One of his Ph.D. students, Bruno Buchberger, was given the problem of finding an algorithm to compute these special generating sets. This algorithm is now known as the Buchberger algorithm, and he named these generating sets after his advisor.

2. Term Orders.

As mentioned in the previous section, if we want to solve the ideal membership problem for polynomials in several variables with something like the division algorithm, then we need to use an ordering of the terms, or, equivalently, monomials in \( k[x] \). In the case of one variable there is the obvious ordering, according to degree. In the multivariable case, there are many different term orders. For theoretical purposes we can use any one in most cases, but for practical purposes not all term orders are created equal. For instance, there are considerable differences with regard to computational complexity. We will have more to say about this subject.

Note that a monomial \( x^\alpha \) in \( k[x_1, \ldots, x_n] \) is uniquely determined by the exponent vector \( \alpha \). Thus, to order the monomials in \( k[x] \), it is sufficient to order the elements of \( \mathbb{N}^n \), where \( \mathbb{N} \) denotes the set of nonnegative integers.

**Definition.** A \textit{term order} on \( k[x_1, \ldots, x_n] \) is a relation \( \prec \) on \( \mathbb{N}^n \) that satisfies the following properties:

1. it is a total ordering;
2. if \( \alpha > \beta \), then \( \alpha + \gamma > \beta + \gamma \) for all \( \gamma \in \mathbb{N}^n \);
3. it is a well-ordering, that is, any subset of \( \mathbb{N}^n \) has a smallest element under \( \prec \).
We give here some of the most commonly used term orders as examples. The reader is encouraged to verify for each of them that they indeed satisfy the properties in the above definition.

**Lexicographic Order.** This term order was introduced in the previous section. It is commonly used, and is one of the few term orders available in MAPLE, for example.

**Graded Lex Order.** Let \( \alpha = (\alpha_1, \ldots, \alpha_n), \beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}^n \). Define \( \alpha <_{\text{grlex}} \beta \) if

\[
|\alpha| = \sum_{i=1}^{n} \alpha_i > |\beta| = \sum_{i=1}^{n} \beta_i, \text{or } |\alpha| = |\beta| \text{ and } \alpha >_{\text{lex}} \beta.
\]

That is, \( \text{grlex} \) orders first according to total degree and breaks ties with \( \text{lex} \). In particular, the variables get ordered according to \( \text{lex} \). Other examples are: \( (1, 2, 3, 4) <_{\text{grlex}} (4, 3, 2, 1) \), since their total degrees are equal, and the first is lexicographically smaller than the second; \( (2, 0, 4) <_{\text{grlex}} (1, 2, 4) \), since the total degree of the second is larger, even though lexicographically the relationship would be opposite.

**Graded Reverse Lex Order.**

This order is a variation of \( \text{grlex} \). Define \( \alpha >_{\text{grevlex}} \beta \) if either \( \alpha \) has larger total degree than \( \beta \), or their total degrees are equal and, in \( \alpha - \beta \in \mathbb{N}^n \), the right-most nonzero entry is negative.

Note first, that both \( \text{grlex} \) and \( \text{grevlex} \) order the variables lexicographically. At first glance one might think that \( \text{grevlex} \) is simply the reverse of \( \text{lex} \), but consider \( (5, 1, 1) \) and \( (4, 1, 2) \). Their total degrees are both equal to 7, so \( (5, 1, 1) >_{\text{grevlex}} (4, 1, 2) \). But this is also true in \( \text{grevlex} \). Thus, \( \text{grevlex} \) is not simply the reverse of \( \text{grlex} \).

We finally describe a method of representing a term order by a vector of real numbers. Let

\[
\omega = (\omega_1, \ldots, \omega_n) \in \mathbb{R}^n.
\]

For a polynomial \( f = \sum_i c_i x^{\alpha_i} \in k[x] \) define the initial form \( \text{LT}_\omega(f) \) to be the sum of all terms \( c_i x^{\alpha_i} \) for which the inner product \( \omega \cdot \alpha_i \) is maximal. If \( I \) is an ideal, we define the initial ideal \( \text{LT}_\omega(I) \) of \( I \) to be the ideal generated by all initial forms of polynomials in \( I \). In general, this ideal need not be a monomial ideal (Exercise 1.6).

**Definition.** Let \( \omega \geq 0 \), and let \( \prec \) be an arbitrary term order on \( k[x] \). Define a new term order \( \prec_\omega \) as follows. For \( \alpha, \beta \in \mathbb{N}^n \), let \( \alpha <_\omega \beta \) if either \( \omega \cdot \alpha < \omega \cdot \beta \), or if \( \omega \cdot \alpha = \omega \cdot \beta \) and \( \alpha \prec \beta \).

Then \( \prec_\omega \) is a term order (Exercise 1.7) which is closely related to \( \prec \), as we shall see.
Proposition 2.1. If $I \subset k[x]$, then
\[ \text{LT}_<(\text{LT}_\omega(I)) = \text{LT}_<(I). \]
(For the proof see Exercise 1.8.)

Theorem 2.2. For any term order $<$ and any ideal $I \subset k[x]$, there exists an integer vector $\omega \in \mathbb{N}^n$ such that $\text{LT}_\omega(I) = \text{LT}_<(I)$.

For a proof see [15, pp. 4-5].
If $\omega \in \mathbb{R}^n$ is any vector such that $\text{LT}_\omega(I) = \text{LT}_<(I)$, then $\omega$ is called a term order for $I$, or that $\omega$ represents $<$ for $I$.

For a complete classification of all term orders see [14].

To conclude this section we present a rather remarkable theorem about term orders, which is the starting point of an extremely interesting and fruitful combinatorial theory of Gröbner bases, and which we will not touch on. The reader is referred to [15]. We will prove this result after we develop some more machinery in the next section.

While there are lots of term orders, they can be naturally collected into finitely many classes. For a given ideal $I \subset k[x]$, each term order $<$ defines the monomial ideal generated by all leading terms of polynomials in $I$, the so-called initial ideal $\text{LT}_<(I)$ of $I$ with respect to $<.$

Theorem 2.3. Each ideal of $k[x]$ has only finitely many initial ideals.


We now come to the heart of the computational algebra “machine,” the tool that forms the basis of the vast majority of algorithms in computational algebra and algebraic geometry. To motivate the definition of a Gröbner basis of an ideal in a polynomial ring, consider again the example in the previous section.

Example 1.1 continued. Let $I \subset k[x, y]$ be generated by $f_1 = x^2 - 1$ and $f_2 = xy + 2$. We asked whether $f = x^2y + xy + 2x + 2$ was an element of $I$, that is, was a linear combination of $f_1$ and $f_2$. Dividing first by $f_1$ and then by $f_2$, we obtained that
\[ f = yf_1 + f_2 + 2x + y, \]
using lex order with $x > y$. We did find, however, that by carrying out the division in a different order, $f$ was indeed in $I$. That implies that the remainder $2x + y$ must also lie in $I$. But its leading term is not divisible by the leading term of either $f_1$ or $f_2$. That is what stopped us in our tracks. The ideal $I$ contains elements whose leading terms are not divisible by the leading term of any of the generators of $I$. 

If we had a (finite) generating set for $I$, so that each element of $I$ had a leading term that was divisible by the leading term of one of the generators, then we could get a reliable answer to the ideal membership question. Whenever the division algorithm produces a remainder whose leading term is not divisible by the leading term of a generator, we have indeed arrived at a polynomial that is not in $I$, and we can quit. The polynomial in question will not be in $I$. This is precisely the idea behind the concept of a Gröbner basis for $I$. We begin by quite naively requiring that our generating set have this property.

**Definition.** Let $I \subset k[x_1, \ldots, x_n] = k[x]$ be an ideal. Let $\langle \text{LT}(I) \rangle$ be the ideal of $k[x]$ generated by all leading terms of polynomials in $I$. A finite subset $\{g_1, \ldots, g_m\}$ of $I$ is a Gröbner basis (or standard basis) of $I$ if $\text{LT}(g_1), \ldots, \text{LT}(g_m)$ generate the ideal $\langle \text{LT}(I) \rangle$, that is, if the leading term of every polynomial in $I$ is divisible by the leading term of some $g_i$.

**Example 1.1 continued.** The polynomials $f_1$ and $f_2$ above do not form a Gröbner basis of the ideal $I$ they generate, as we have seen. We claim that $\{g_1 = y^2 - 4, g_2 = 2x + y\}$ is a Gröbner basis for $I$ with respect to $\text{lex}$, with $x > y$. For this we need to show that the leading term of every polynomial in $I$ is divisible by the leading term of $g_1$ or $g_2$, that is, by $y^2$ or by $x$. (We can ignore constants, of course.) Let

$$h = h_1(x^2 - 1) + h_2(xy + 2) \in I.$$ 

If the leading term of $h$ contains an $x$, then it is divisible by the leading term of $g_2$. So suppose that the leading term of $h$ is a power of $y$. This implies that $h$ is a polynomial in $y$. We need to show that it has degree at least two.

By way of contradiction, assume that $h = y + a$. If $a = 0$, then $I$ also contains 2, hence $I = k[x, y]$. This implies that $f_1$ and $f_2$ generate the unit ideal. To see that this is not possible, note that $I$ is contained in the ideal $\langle x-1, xy+2 \rangle = \langle x-1, y+2 \rangle$. So it is sufficient to convince ourselves that this larger ideal cannot be all of $k[x, y]$. But this is clear, since $k[x, y]/\langle x-1, y+2 \rangle \cong k$.

So assume that $a \neq 0$. Then we also get that $x + \frac{2}{a} \in I$, and, consequently, $xy + \frac{2}{a}y \in I$. Subtracting $xy + 2$ from this, clearing denominators and dividing by 2, we obtain that $y - a \in I$, together with $y + a$. Thus, $2a \in I$, which is a unit. But we showed earlier that $I$ was not the unit ideal. Hence $I$ does not contain any polynomials in $y$ of degree one. Thus, we have shown that $\{g_1, g_2\}$ is a Gröbner basis for $I$. 
The next proposition and its corollary show that Gröbner bases would indeed save the day for us.

**Proposition 3.1.** Let $G = \{g_1, \ldots, g_m\}$ be a Gröbner basis for an ideal $I \subseteq k[x]$, and let $f \in k[x]$. Then there exists a unique $r \in k[x]$ such that no term of $r$ is divisible by the leading term of any $g_i$, and there is a $g \in I$ such that $f = g + r$.

**Proof.** The division algorithm will certainly give us an $r$ that satisfies the first condition, and it will also produce a $g$ for us with the required property. The whole crux, as we have seen, lies in the uniqueness of $r$. So suppose that $f = g_1 + r_1 = g_2 + r_2$, with $g_i$ and $r_i$ satisfying the conditions of the proposition. Then $r_1 - r_2 = g_2 - g_1 \in I$. If $r_1 \neq r_2$, then

$$LT(r_1 - r_2) \in LT(I) = \langle LT(g_1), \ldots, LT(g_m) \rangle.$$ 

Hence, $LT(r_1 - r_2)$ is divisible by the leading term of some $g_i$, which is impossible since no term of either $r_1$ or $r_2$ is divisible by the leading term of any $g_i$. Hence $r_1 = r_2$, and the proof of the proposition is complete.

**Corollary 3.2.** Let $G$ be a Gröbner basis for an ideal $I \subseteq k[x]$, and $f \in k[x]$. Then $f \in I$ if and only if the remainder of $f$ under division by the elements of $G$ is zero.

If we could now prove that every ideal has a Gröbner basis, and that a Gröbner basis is in fact a generating set for the ideal, then we would be in business, provided we can also find an algorithm to compute such a basis. In this section we solve the first problem, showing that every ideal indeed has a Gröbner basis. The main reason why this is true is the so-called Hilbert Basis Theorem, named after David Hilbert (1862–1943), who discovered it in the context of invariant theory. This theorem asserts that every ideal of the polynomial ring $k[x]$ is finitely generated. We first prove this theorem for ideals generated by monomials, which is known as Dickson’s Lemma.

Let

$$I = \langle x^\alpha | \alpha \in A \rangle$$

be an ideal generated by monomials $x^\alpha$, with $\alpha \in A \subseteq \mathbb{N}^n$. We begin with the observation that a monomial $x^\beta$ lies in $I$ if and only if $x^\beta$ is divisible by some $x^\alpha$. That is,

$$\{x^\beta | x^\beta \in I\} = \{x^\beta | \beta = \alpha + \gamma, \alpha \in A, \gamma \in \mathbb{N}^n \}.$$

This observation, which is straightforward to prove, allows us to draw pictures of monomial ideals, that are commonly used and are quite illustrative. As an example, consider the ideal $I = \langle x^3y, xy^3, x^2y^2 \rangle \subset k[x,y]$. The next proposition and its corollary show that Gröbner bases would indeed save the day for us.
We plot the exponent vectors of the elements in $I$. They are all the grid points that lie in the shaded area of the diagram:

For more than two variables, it becomes a bit more complicated, of course, to draw such pictures. Lego blocks help for three variables.

**Theorem 3.3.** (Dickson’s Lemma) Let

$$I = \langle x^\alpha | \alpha \in A \rangle \subset k[x_1, \ldots, x_n] = k[x]$$

be a monomial ideal. Then there exists a finite subset

$$\{x^{\alpha(1)}, \ldots, x^{\alpha(m)}\} \subset \{x^\alpha | \alpha \in A\}$$

that generates $I$.

**Proof.** We follow the proof in [3, p. 69]. The proof proceeds by induction on the number $n$ of variables. If $n = 1$, then we can simply choose a generator $x^\alpha$ with $\alpha \in \mathbb{N}$ minimal. So assume that $n > 1$ and the theorem is true for $n - 1$.

We rename the last variable $x_n$ as $y$, and write monomials in $k[x]$ as $x^\alpha y^s$, with $\alpha \in \mathbb{N}^{n-1}$. Let $J \subset k[x_1, \ldots, x_{n-1}]$ be the monomial ideal generated by all $x^\alpha$ such that $x^\alpha y^s \in I$ for some $s$. Then $J$ is a monomial ideal, so by induction hypothesis we can find $\alpha(1), \ldots, \alpha(t)$, which generate $J$. Let

$$x^{\alpha(1)} y^{m_1}, \ldots, x^{\alpha(t)} y^{m_t}$$

be monomials in $I$, and let $m$ be the maximum of the $m_i$. Consider the ideals

$$J_i = \langle x^\alpha \in k[x_1, \ldots, x_{n-1}] | x^\alpha y^i \in I \rangle$$

of $k[x_1, \ldots, x_{n-1}]$, for $i = 0, \ldots, n - 1$. By induction, each of these ideals is finitely generated, with generating set

$$x^{\alpha_{i(1)}}, \ldots, x^{\alpha_{i(t_i)}}.$$
We claim that $I$ is generated by the set $S$ of monomials

\[ x^{a(1)}y^m, \ldots, x^{a(t)}y^m; \]

\[ x^{a_0(1)}, \ldots, x^{a_0(s)}; \]

\[ x^{a_1(1)}y, \ldots, x^{a_1(t)}y; \]

\[ \cdots \]

\[ x^{a_{m-1}(1)}y^{m-1}, \ldots, x^{a_{m-1}(t_{m-1})}y^{m-1}. \]

First of all, every monomial in $I$ is divisible by one of these monomials. Let $x^a y^r \in I$. If $r \geq m$, then $x^a y^r$ is divisible by some $x^{a(i)}y^m$, because of the construction of $J$. If $r \leq m-1$, then the monomial is divisible by one of the generators coming from $J_r$. This implies that the monomials in the finite set $S$ generate $I$.

It remains to show that we can choose the finite generating set from the collection of original generators for $I$. But this follows from the fact that each of the monomials in $S$ is divisible by one of the original generators. Conversely, each of the original generators $x^a, a \in A$, of $I$ are divisible by a monomial in $S$. This completes the proof of Dickson's Lemma.

We now prove the same result for general ideals in polynomial rings.

**Theorem 3.4. (Hilbert Basis Theorem)** Every ideal $I \subset k[x]$ is finitely generated.

**Proof.** If $I = 0$, there is nothing to be done. Choose a term order, and let $\langle LT(I) \rangle$ be the initial ideal of $I$. Since it is a monomial ideal, Dickson's Lemma tells us that $\langle LT(I) \rangle$ is actually generated by the leading terms of finitely many polynomials $g_1, \ldots, g_m \in I$. It turns out that these polynomials actually generate $I$. Namely, for $f \in I$, use the division algorithm to obtain an expression

\[ f = h_1g_1 + \cdots + h_mg_m + r, \]

where $r \in I$ is not divisible by the leading term of any $g_i$. But then $r$ must be zero, since the leading term of $r$ is in $\langle LT(I) \rangle$, which is generated by the leading terms of the $g_i$. This completes the proof.

**Corollary 3.5.** Every ideal $I$ in $k[x]$ has a Gröbner basis. Furthermore, any Gröbner basis for $I$ is also a generating set for $I$.

**Proof.** The finite generating set in the proof of the Hilbert Basis Theorem is a Gröbner basis by construction.

For a given ideal and a given term order, there are of course many different Gröbner bases. For instance, given one, we obtain another one by adding a finite number of polynomials. There are, however, some distinguished ones.
Definition. Let $I \subset k[x]$. A Gröbner basis $G = \{g_1, \ldots, g_m\}$ for $I$ is minimal if all leading coefficients are equal to 1, and if no $LT(g_i)$ is divisible by another $LT(g_j)$ for $i \neq j$. And $G$ is called reduced if all leading coefficients are equal to 1, and no term of $g_i$ is divisible by $LT(g_j)$ for $i \neq j$.

It is not hard to see that reduced Gröbner bases are unique (Exercise 1.9).

We now mention a condition on $k[x]$ which is equivalent to every ideal being finitely generated. While this is an aside in a sense, it allows us to introduce useful terminology.

Definition. A ring $R$ is called Noetherian if every ascending chain

$$I_1 \subset I_2 \subset I_3 \subset \cdots$$

of ideals of $R$ eventually becomes stationary, that is, if there exists an $N \geq 1$ such that $I_N = I_{N+1} = I_{N+2} = \cdots$.

This definition is inspired by the work of Emmy Noether (1882–1935), who identified rings with this property as an important class. Her aim was to classify all rings for which the Fundamental Theorem of Arithmetic for ideals holds true, that is, rings for which every ideal is uniquely a product of prime ideals. The Noetherian condition assures the existence of such a product decomposition. (More conditions are needed for uniqueness. Rings for which this factorization theorem holds are known as Dedekind domains.)

Corollary 3.6. The polynomial ring $k[x]$ is Noetherian.

Proof. Given an ascending chain of ideals of $k[x]$ as in the definition, consider the ideal

$$I = \bigcup_{i=1}^{\infty} I_i \subset k[x].$$

By the Hilbert Basis Theorem, $I$ is finitely generated, say $I = \langle g_1, \ldots, g_m \rangle$. Each $g_i$ is contained in some $I_j$, hence all the $g_i$ are contained in some $I_N$. Consequently, the chain becomes stationary from that point on. This completes the proof.

One can prove, conversely, that in a Noetherian ring every ideal is finitely generated.

To close this section we prove Theorem 2.3. Let $I \subset k[x]$ be an ideal, and let $<$ be a term order on $k[x]$. Call those monomials which are not leading monomials of polynomials in $I$ standard monomials.
**Proposition 3.7.** The (images of the) standard monomials form a \( k \)-vector space basis for \( k[x]/I \).

**Proof.** Let \( G \) be the reduced Gröbner basis for \( I = \langle f_1, \ldots, f_m \rangle \). Let \( f \in k[x] \) be a representative for an element of the quotient \( k[x]/I \). Then the division algorithm rewrites \( f \) uniquely as a \( k \)-linear combination of standard monomials.

**Proof of Theorem 2.3.** (See [15, p. 1] and also [1, p. 50].) Let \( I = \langle f_1, \ldots, f_r \rangle \subset k[x] \) be an ideal. We need to show that \( I \) has only finitely many initial ideals. Arguing by contradiction, suppose that there are infinitely many distinct initial ideals of \( I \). There are only finitely many terms appearing in \( f_1, \ldots, f_r \), so for each \( i \) there must be a term \( m_i \) of \( f_i \) which appears as leading term in infinitely many initial ideals of \( I \). Let \( S \) be the infinite set of initial ideals which contain \( m_i \) for each \( i \). Suppose that there is an initial ideal which has \( m_1, \ldots, m_r \) as a basis. Then the \( f_i \) form a Gröbner basis for \( I \) for any term order with this initial ideal. Let \( M \) be another initial ideal which contains the \( m_i \). Then \( \{f_i\} \) is also a Gröbner basis with respect to any term order giving this initial ideal (Exercise 1.10). Hence \( m_1, \ldots, m_r \) is a basis for all initial ideals in \( S \), so that \( S \) contains only one element. Since we assumed that \( S \) was infinite, this cannot happen.

Hence there exists \( h_1 \in I \) such that \( \text{LT}(h_1) \) is not divisible by any \( m_i \). Now add \( h_1 \) to \( f_1, \ldots, f_r \) and \( \text{LT}(h_1) \) to \( m_1, \ldots, m_r \) and repeat the argument. Thus, there exist infinitely many initial ideals which contain \( m_1, \ldots, m_r, h_1 \). Continuing in this way, we obtain an infinite strictly increasing sequence of monomial ideals

\[
\langle m_1, \ldots, m_r \rangle \subset \langle m_1, \ldots, m_r, h_1 \rangle \subset \langle m_1, \ldots, m_r, h_1, h_2 \rangle \subset \cdots.
\]

Since \( k[x] \) is Noetherian, this is a contradiction. The proof is complete.

**Definition.** Let \( I \subset k[x] \). A finite subset \( G \) of \( I \) is called a **universal Gröbner basis** for \( I \) if it is a Gröbner basis for \( I \) with respect to any term order on \( k[x] \).

**Corollary 3.8.** Every ideal has a universal Gröbner basis.

**Proof.** Theorem 2.3 implies that there are only finitely many distinct reduced Gröbner bases for \( I \) (Exercise 1.11). Then their union is finite and is a universal Gröbner basis for \( I \). This completes the proof.

We are left with the problem of actually finding Gröbner bases of ideals (see Exercise 1.12 for an example), and with the related problem of verifying that a given set of polynomials is a Gröbner basis. This will be addressed in the next section.
4. The Buchberger Algorithm.

It was B. Buchberger who gave an algorithmic criterion for a set of polynomials to be a Gröbner basis, and this criterion forms the basis for practically every known algorithm to compute a Gröbner basis from a given set of generators for an ideal. The motivation for it lies precisely in the deficiency of general generating sets that prevents them from being Gröbner bases. If \( I = \{ f_1, \ldots, f_n \} \subset k[x] \), then \( I \) may contain polynomials whose leading term is not divisible by any leading term of the \( f_i \). One way this can happen is if there are linear combinations \( P_i \alpha_i f_i \) in which the leading terms cancel. This is what happened in Example 1.1. We will see that the main culprit are linear combinations of two generators. We need some terminology.

**Definition.** Let \( f, g \in k[x] \), with leading monomials \( x^\alpha \) and \( x^\beta \), respectively. Let \( \gamma = (\gamma_1, \ldots, \gamma_n) \), with \( \gamma_i = \max(\alpha_i, \beta_i) \). Call \( x^\gamma \) the least common multiple of \( x^\alpha \) and \( x^\beta \), denoted by \( \text{LCM}(\text{LM}(f), \text{LM}(g)) \). Define the \( S \)-polynomial of \( f \) and \( g \) to be

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

As an example, consider our old friends \( f_1 = x^2 - 1 \) and \( f_2 = xy + 2 \) in \( k[x, y] \). Using \( \text{lex} \) order with \( x > y \) we find that

\[
S(f_1, f_2) = \frac{x^2y}{x^2}(x^2 - 1) - \frac{x^2y}{xy}(xy + 2) = y - 2x.
\]

For good measure, let us compute another \( S \)-polynomial, namely

\[
S(y + 2x, f_2) = \frac{1}{2}y^2 - 2.
\]

Up to constants we obtain precisely the Gröbner basis for \( (f_1, f_2) \) found in Example 1.1. Obviously, \( S \)-polynomials are set up precisely to produce the kind of cancellation that results in leading terms of polynomials in the ideal, that are not divisible by any leading term of the generators. The next lemma shows that they are responsible for all cancellations of this kind.

**Lemma 4.1.** Suppose that \( f_1, \ldots, f_r \in k[x] \) with \( \text{LM}(f_i) = m \). Fix a term order on \( k[x] \), and let \( f = \sum c_i f_i \), with \( c_i \in k \). If \( \text{LM}(f) < m \), then \( f \) is a \( k \)-linear combination of \( S \)-polynomials of the \( f_i \).

**Proof.** (We follow [1, p. 41].) Let \( f_i = a_i m + \text{lower terms} \), with \( a_i \in k \). Then the hypothesis implies that \( \sum c_i a_i = 0 \). Observe that

\[
S(f_i, f_j) = \frac{1}{a_i} f_i - \frac{1}{a_j} f_j,
\]
since the $f_i$ all have the same leading monomial. We now rewrite $f$ as
\[
\begin{align*}
f &= c_1 f_1 + \cdots + c_r f_r \\
&= c_1 a_1 \left( \frac{1}{a_1} f_1 \right) + \cdots + c_r f_r \left( \frac{1}{a_r} f_r \right) \\
&= c_1 a_1 \left( \frac{1}{a_1} f_1 - \frac{1}{a_2} f_2 \right) + (c_1 a_1 + c_2 a_2) \left( \frac{1}{a_2} f_2 - \frac{1}{a_3} f_3 \right) + \cdots \\
&\quad + (c_1 a_1 + \cdots + c_{r-1} a_{r-1}) \left( \frac{1}{a_{r-1}} f_{r-1} - \frac{1}{a_r} f_r \right) \\
&\quad + (c_1 a_1 + \cdots + c_r a_r) \frac{1}{a_r} f_r \\
&= c_1 a_1 S(f_1, f_2) + (c_1 a_1 + c_2 a_2) S(f_2, f_3) + \cdots \\
&\quad + (c_1 a_1 + \cdots + c_{r-1} a_{r-1}) S(f_{r-1}, f_r),
\end{align*}
\]
since $\sum c_i a_i = 0$. This completes the proof.

To simplify things we introduce the following notation. Let $G = \{g_1, \ldots, g_r\} \subset k[x]$ and $f \in k[x]$. Dividing $f$ by the elements of $G$ in the order indicated results in a remainder $r$. We denote this fact by
\[
f \rightarrow^G r.
\]

**Theorem 4.2. (Buchberger’s Criterion)** Let $I \subset k[x]$, and consider a fixed term order $<$ for $k[x]$. A generating set $G = \{g_1, \ldots, g_r\}$ of $I$ is a Gröbner basis for $I$ with respect to $<$ if and only if $S(f_i, f_j) \rightarrow^G 0$ for all $i \neq j$.

**Proof.** (We follow [1, pp. 41–42].) If $G$ is a Gröbner basis, then $f \rightarrow^G 0$ for all $f \in I$, including $S(g_i, g_j)$.

Conversely, assume that all $S$-polynomials of the $g_i$ reduce to 0. Let $f = \sum_i h_i g_i \in I$. We need to show that $\text{LT}(f)$ is divisible by $\text{LT}(g_j)$ for some $j$. In order to do so we want to choose the $h_i$ such that
\[
m = \max_i (\text{LT}(h_i) \text{LT}(g_i))
\]
is minimal with respect to the chosen term ordering. If $m = \text{LT}(f)$, then we are done. Otherwise, $m < \text{LT}(f)$, we will derive a contradiction by finding a representation of $f$ with smaller $m$.

Let
\[
X = \{i \mid \text{LT}(h_i) \text{LT}(g_i) = m\}.
\]
For $i \in X$, write $h_i = c_i m_i + \text{lower terms}$. Let $g = \sum_{i \in X} c_i m_i g_i$. Then $\text{LT}(m_i g_i) = m$ for all $i \in X$, but $\text{LT}(g) < m$. By the previous lemma it follows that $g$ is a linear combination of $S$-polynomials:
\[
g = \sum_{i,j \in X, i \neq j} d_{ij} S(m_i g_i, m_j g_j).
\]
We now compute these $S$-polynomials. Since the least common multiple of all pairs of leading terms of the $m_i g_i$ is $m$, we obtain from the
definition of \( S \)-polynomials that
\[
S(m_i g_i, m_j g_j) = \frac{m}{m_{ij}} S(g_i, g_j),
\]
where \( m_{ij} = \text{LCM} \left( \text{LT}(g_i), \text{LT}(g_j) \right) \). But by hypothesis we have that \( S(g_i, g_j) \rightarrow^G 0 \) for all \( i \neq j \). Hence \( S(m_i g_i, m_j g_j) \rightarrow^G 0 \) also. In other words, we obtain a linear combination
\[
S(m_i g_i, m_j g_j) = \sum_s \varepsilon_{ij,s} g_s,
\]
and
\[
\max_s \left( \text{LT}(h_{ij,s}) \text{LT}(g_s) \right) = \text{LT}(S(m_i g_i, m_j g_j)) < \max \left( \text{LT}(m_i g_i), \text{LT}(m_j g_j) \right) = m.
\]
By back-substituting these expressions into \( g \), and then \( g \) into \( f \), we obtain that
\[
f = \sum_i h_i' g_i,
\]
with
\[
\max_i \left( \text{LT}(h_i') \text{LT}(g_i) \right) < m,
\]
which is a contradiction. This completes the proof.

We illustrate this theorem with our usual example. Earlier we worked hard to show that \( \langle g_1 = y^2 - 4, g_2 = 2x + y \rangle \) is a Gröbner basis for the ideal \( \langle x^2 - 1, xy + 2 \rangle \). With Buchberger’s Criterion this is a piece of cake, since all we need to check is that
\[
S(g_1, g_2) = \frac{2xy^2}{y^2}(y^2 - 4) - \frac{2xy^2}{2x}(2x + y) = -y(y^2 - 4) - 2(2x + y),
\]
that is, \( S(g_1, g_2) \rightarrow^G 0 \).

It is now easy to see that one can take Buchberger’s Criterion and turn it around into an algorithm to compute a Gröbner basis for a given ideal. We give here the basic idea of the algorithm; for practical implementations there are many possible improvements one can and should make. Some of these can be found in [1, 2, 3] and the references given there.

**The Buchberger Algorithm.**

**Input:** Polynomials \( f_1, \ldots, f_r \) in \( k[x_1, \ldots, x_n] \), and a term order \(<\).

**Output:** A Gröbner basis \( G \) with respect to \(<\) for the ideal generated by the \( f_i \).

1. Set \( G = \{f_1, \ldots, f_r\} \).
2. Compute all $S$-polynomials $S(g_i, g_j)$ for $g_i, g_j \in G$, and reduce them modulo $G$.

3. If all $S$-polynomials are zero, output $G$. Otherwise, add all nonzero reductions of $S$-polynomials to $G$ and go back to Step (2).

To make the output of the algorithm unique, one can now compute the reduced Gröbner basis by eliminating redundant polynomials and reduce the remaining ones, so that no term of $g_i$ is divisible by the leading term of any $g_j, j \neq i$. We present an example to illustrate this process.

**Example.** Let

$I = \langle f_1 = x^2 + y^2 + z^2 - 4, f_2 = x^2 + 2y^2 - 5, f_3 = xz - 1 \rangle \subset \mathbb{Q}[x, y, z],$

and let $<$ be lex order with $x > y > z$. To begin with, set $G = \{f_1, f_2, f_3\}$. We compute $S$-polynomials

\[
\begin{align*}
S(f_1, f_2) &= 3y^2 + z^2 - 9, \\
S(f_1, f_3) &= -x + zy^2 + z^3 - 4z, \\
S(f_2, f_3) &= -x + 2zy^2 - 5z.
\end{align*}
\]

None of these is reducible modulo $G$, so we add them and begin again with the new $G = \{f_1, \ldots, f_6\}$.

At the next step we need to compute further $S$-polynomials, such as

\[
S(f_1, f_3) = -y^4 - y^2z^2 + 4y^2 - x^2z^2 - x^2 \rightarrow G - x^2z^2 - x^2 - 2z^2 + xz + 3.
\]

Continuing in this fashion until all $S$-polynomials reduce to zero we obtain the Gröbner basis

\[
G = \{x + 2z^3 - 3z, y^2 - z^2 - 1, 2z^4 - 3z^2 + 1\}.
\]

It is clear that the algorithm indeed produces a Gröbner basis for the ideal generated by the $f_i$. What is not so clear is that it terminates after finitely many steps. After all, at each step we add potentially lots more polynomials to our generating set, which in turn produces many more $S$-polynomials. Once again we are saved by the Hilbert Basis Theorem. Namely, let $\text{LT}(G_0)$ denote the monomial ideal of leading terms of $\langle f_1, \ldots, f_r \rangle$. During the next step in the algorithm we add leading terms of $S$-polynomials to this ideal to obtain $\text{LT}(G_1)$, strictly larger than $\text{LT}(G_0)$. Thus, at each step we produce a monomial ideal strictly larger than the previous one. In this way we construct a strictly increasing sequence of ideals in $k[x]$, which has to terminate by
the Hilbert Basis Theorem. That is, at some point, the leading terms of the elements in \( G \) generate the ideal of leading terms of \( (G) \). But this is precisely the criterion for being a Gröbner basis.

To summarize the results of this rather long section, we have seen that the concept of a Gröbner basis for an ideal is precisely what we need to get a division algorithm for a polynomial ring in several variables that produces a unique remainder, and we can use it to solve the ideal membership problem. Another, rather harmless looking, result we obtained is that the standard monomials, that is, those monomials that are not leading terms of polynomials in the ideal \( I \) in question, form a \( k \)-vector space basis for the quotient \( k[x]/I \). Furthermore, each element of the quotient has a unique representative that is a \( k \)-linear combination of standard monomials. But that allows us to make computations in the quotient ring by lifting equivalence classes to their canonical representatives in \( k[x] \), doing the computation up there, and then projecting back down. (One needs to be just a bit careful in doing this.) This represents a huge extension of the class of rings in which we can compute. We will say more about this topic in the coming sections.

5. Applications of Gröbner Bases to Ideal Theory.

In this section we collect some fundamental algorithms which allow us to work constructively with ideals. They all have the Buchberger algorithm as an essential component. Specifically, we will compute the intersection of an ideal with a polynomial ring in fewer variables, intersections of ideals, ideal quotients, and operations in quotient rings \( k[x]/I \).

To begin with, it is worth pointing out that we can use Gröbner bases to tell whether two ideals are equal. As we have seen, this is not always a trivial matter. Given two sets of polynomials \( \{f_1, \ldots, f_r\} \) and \( \{h_1, \ldots, h_s\} \), we can determine whether they generate the same ideal by computing a reduced Gröbner basis for both sets. The two ideals are equal if and only if the resulting Gröbner bases are equal. In particular, we can tell whether a set of polynomials generates the whole ring. This is the case precisely if the reduced Gröbner basis consists of the single element 1.

We first consider the intersection of an ideal and a subring of \( k[x] \). The next result will be extremely useful when we apply Gröbner basis methods to the solution of systems of polynomial equations. It forms the basis for eliminating variables.

**Theorem 5.1. (Elimination Theorem)** Let \( I \subset k[x_1, \ldots, x_n] \), and let \( G \) be a Gröbner basis for \( I \) with respect to \( \text{lex} \), where \( x_1 > x_2 > \ldots > x_n \). Then...
\[ \cdots > x_n. \] Then for every \( 0 \leq r \leq n \), the set
\[ G_r = G \cap k[x_{r+1}, \ldots, x_n] \]
is a Gröbner basis for the ideal
\[ I_r = I \cap k[x_{r+1}, \ldots, x_n]. \]
(The ideal \( I_r \) is called the \( r \)-th elimination ideal.)

**Proof.** Fix \( k \) and let \( G = \{ g_1, \ldots, g_m \} \). We may assume that the
\( g_i \) are ordered so that exactly the first \( t \) elements lie in \( k[x_{r+1}, \ldots, x_n] \), that is,
\[ G_r = \{ g_1, \ldots, g_t \}. \]
Certainly, \( G_r \subseteq I_r \). Let \( f \in I_r \), that is, \( f \) does not contain any of the
variables \( x_1, \ldots, x_r \). Using the division algorithm with the ordering of
the variables as above to divide \( f \) by the elements of \( G \), we see that \( f \)
is a linear combination of \( g_1, \ldots, g_t \), that is, \( I_r = \langle g_1, \ldots, g_t \rangle \).

To show that \( G_r \) is a Gröbner basis we need to show that all \( S \)-polynomials reduce to zero. But they do upstairs in \( k[x_{r+1}, \ldots, x_n] \), and
all steps of the division algorithm take place downstairs in \( k[x_{r+1}, \ldots, x_n] \).
This completes the proof.

This is a rather amazing result, since it provides an easy solution
to an apparently very difficult problem. There are no free lunches,
however. Computation of elimination ideals is notoriously expensive.

**Example.** The example after the Buchberger Algorithm was in
fact an elimination ideal computation for
\[ I = \langle f_1 = x^2 + y^2 + z^2 - 4, f_2 = x^2 + 2y^2 - 5, f_3 = xz - 1 \rangle \subset \mathbb{Q}[x, y, z]. \]
Using \( \text{lex} \) order with \( x > y > z \) we obtain
\[ I \cap \mathbb{Q}[y, z] = \langle y^2 - z^2 - 1 \rangle, \]
and
\[ I \cap \mathbb{Q}[z] = \langle 2z^4 - 2z^2 + 1 \rangle. \]

Now we consider the intersection
\[ I \cap J = \{ f \in k[x] \mid f \in I \quad \text{and} \quad f \in J \} \]
of two ideals \( I, J \subseteq k[x] \). It is straightforward to check that this is
again an ideal. We will now give an algorithm to compute generators
for \( I \cap J \) from generators for \( I \) and \( J \), by reducing it to the problem of
computing an elimination ideal, solved above. We invite the reader to
try the following example by hand: let
\[ I = \langle (x+y)^2(z^2-1)(x+y)^4 \rangle \]
and

\[ J = \langle (x + y) (z - 1) (2x + yz) \rangle. \]

For the next theorem, recall that if \( I = \langle \{ f_i \} \rangle, J = \langle \{ g_j \} \rangle \) are ideals of \( k[x] \), then

\[ I + J = \{ f + g | f \in I, g \in J \} \]

is also an ideal of \( k[x] \), with generators \( \{ f_i, g_j \} \).

**Theorem 5.2.** Let \( I, J \subset k[x] \), and let \( t \) be a variable. Then

\[ I \cap J = (tI + (1 - t)J) \cap k[x], \]

where

\[ tI + (1 - t)J = \langle t \cdot f | f \in I \rangle + \langle (1 - t) \cdot g | g \in J \rangle \]

is an ideal of \( k[x, t] \).

**Proof.** If \( f \in I \cap J \), then

\[ f = tf + (1 - t)f \in tI + (1 - t)J. \]

Conversely, suppose \( f = g(x, t) + h(x, t) \), with \( g(x, t) \in tI, h(x, t) \in (1 - t)J \), and \( f \) does not involve the variable \( t \). Setting \( t = 0 \) we obtain that \( f = h(x, 0) \in J \). Similarly, for \( t = 1 \), we have that \( f = g(x, 1) \in I \). This completes the proof.

**Example.** Let \( I = \langle xy + x^3, y^2 + x \rangle \) and \( J = \langle y^3 + y, y^3 - xy \rangle \) be ideals in \( \mathbb{Q}[x, y] \). We form the ideal

\[ K = \langle xyt + x^3 t, y^2 t + xt, y^3 + y - y^3 t - yt, y^3 t - xyt - y^3 - xy \rangle \subset \mathbb{Q}[x, y, t]. \]

We now compute a Gröbner basis of \( K \) with \( lex \) order and \( t > x > y \), and obtain

\[ \{ xt + y^4 + y^2, yt - y^3 - y, xy + y^3, y^5 + y^3 \}. \]

Then

\[ I \cap J = K \cap \mathbb{Q}[x, y] = \langle xy + y^3, y^5 + y^3 \rangle. \]

We can use this result to construct an algorithm to compute the greatest common divisor of two polynomials as follows. Observe first that, for any two polynomials \( f, g \in k[x] \),

\[ \text{LCM}(f, g) \text{GCD}(f, g) = fg. \]

Thus, we can compute the greatest common divisor of \( f \) and \( g \) if we can compute their least common multiple.
Proposition 5.3. Let \( f, g \in k[x] \). Then
\[
\langle f \rangle \cap \langle g \rangle = \langle \text{LCM}(f, g) \rangle.
\]
Hence,
\[
\text{GCD}(f, g) = \frac{fg}{\text{LCM}(f, g)}.
\]

The proof is left to the reader as an exercise.

Another frequently used operation on two ideals is the formation of their ideal quotient.

**Definition.** Let \( I, J \subset k[x] \). Their ideal quotient \( I : J \) is the ideal
\[
I : J = \{ f \in k[x] \mid fg \in I \text{ for all } g \in J \}.
\]

The reader is invited to verify that \( I : J \) is indeed an ideal, and that it contains \( I \).

As a simple example, the reader can verify that
\[
\langle x^2, xy \rangle : \langle y \rangle = \langle x \rangle.
\]

We now derive an algorithm to compute general ideal quotients (following [3, Ch 4]).

Proposition 5.4. Let \( I, I_i, J, J_i, K \subset k[x] \), be a finite collection of ideals. Then
1. \[
\left( \bigcap_i I_i \right) : J = \bigcap_i (I_i : J);
\]
2. \[
I : \left( \bigcap_i J_i \right) = \bigcap_i (I : J_i);
\]
3. \[
( I : J ) : K = I : JK.
\]

The proof is straightforward and is left to the reader.

Theorem 5.5. Let \( I \subset k[x] \), and \( g \in k[x] \). If \( \{ h_i \} \) is a basis for \( I \cap \langle g \rangle \), then \( \{ h_i / g \} \) is a basis for \( I : \langle g \rangle \).

**Proof.** Clearly, the elements \( h_i / g \) are contained in \( I : \langle g \rangle \). Conversely, let \( f \in I : \langle g \rangle \), then
\[
g f \in I \cap \langle g \rangle = \langle h_i \rangle.
\]
Hence $gf = \sum_i r_i h_i$ for some $r_i \in k[x]$. Each $h_i$ is contained in $\langle g \rangle$, so that $h_i/g$ makes sense. We obtain that
\[ f = \sum_i r_i (h_i/g), \]
which completes the proof of the theorem.

**Algorithm to Compute $I : J$.**

**Input:** ideals $I = \langle f_1, \ldots, f_r \rangle$ and $J = \langle g_1, \ldots, g_s \rangle$

**Output:** generators for $I : J$

1. Compute $I : \langle g_i \rangle$ for each $i$, by computing generators for $I \cap \langle g_i \rangle$ and dividing them by $g_i$, according to the above theorem.
2. Observe that
\[ I : J = I : \left(\sum_i \langle g_i \rangle\right) = \bigcap_i \left(I : \langle g_i \rangle\right). \]

Compute the intersection inductively.

Finally, we summarize how to carry out computations in a quotient ring $k[x]/I$ by lifting up to $k[x]$ using normal forms of polynomials.

**Proposition 5.6.** Let $G$ be a Gröbner basis for an ideal $I \subset k[x]$ with respect to any term order. Let $[f] \in k[x]/I$, and let $\overline{f} = \overline{f} = \overline{f} \in k[x]$ be the remainder of $f$ under division by $G$. (That is, $\overline{f}$ is a $k$-linear combination of monomials which are not leading terms of polynomials in $I$, i.e. standard monomials.) Call $\overline{f}$ the canonical representative of $[f]$ in $k[x]$. Then
\[ \overline{f} + \overline{g} \]
is the canonical representative of $[f] + [g]$;
\[ \overline{f} \cdot \overline{g} \]
is the canonical representative of $[f] \cdot [g]$.

The only observation to make for the proof is that $\overline{f} \cdot \overline{g}$ cannot already be the canonical representative of $[f] \cdot [g]$, since it might contain nonstandard monomials.

To summarize, in this section we have given algorithms for some fundamental operations on ideals in a polynomial ring, and for computations in quotient rings $k[x]/I$. Later we will need algorithms for somewhat more sophisticated operations such as the radical of an ideal and primary decomposition of ideals. These are considerably more complicated, and will be treated later on in these notes.
6. Gröbner Bases for Binomial Ideals

As we have seen in earlier sections, monomial ideals play an important role in the computational theory of general ideals. Furthermore, it is easy to carry out computations in them. Most importantly, the class of monomial ideals is closed under a number of operations on ideals, that we will study later on, such as the formation of radicals and primary decomposition. Recent work by Sturmfels and others has shown also that an understanding of all aspects of monomial ideals is crucial in answering some of the fundamental questions in computational ideal theory. The next natural class of ideals to consider is that of binomial ideals, that is, ideals that are generated by differences of two terms. It turns out that one can develop a good theory for them, and that they are an interesting class for some applications. In this section we will describe the Gröbner basis theory and some basic ideal operations for binomial ideals. We follow the treatment in the influential article by Eisenbud and Sturmfels [7], and the reader is encouraged to study that paper in more detail.

Let \( I \subset k[x] \) be an ideal generated by polynomials of the form

\[ ax^a - bx^b, \ a, b \in k. \]

Such polynomials are called binomials, and the ideal they generate is a binomial ideal.

**Proposition 6.1.** Let \( I \) be a binomial ideal. Then any reduced Gröbner basis \( G \) of \( I \) consists of binomials. Furthermore, the normal form \( \overline{a}^G \) of any term \( a \) is again a term.

**Proof.** If we start with a binomial generating set for \( I \), then it is straightforward to see that any new element produced by the Buchberger algorithm is again a binomial. Furthermore, each step of the division algorithm modulo a set of binomials takes a term to another term.

This proposition immediately allows us to test whether a given ideal is binomial.

**Corollary 6.2.** Let \(< \) be a term order on \( k[x] \). An ideal \( I \) is binomial if and only if some (equivalently, every) reduced Gröbner basis for \( I \) consists of binomials.

**Corollary 6.3.** If \( I \subset k[x_1, \ldots, x_n] \) is a binomial ideal, then every elimination ideal \( I \cap k[x_r, \ldots, x_n] \) is again binomial.

Lest the reader get the impression that all operations on binomial ideals again produce binomial ideals, we mention that, for instance,
the ideal quotient of a binomial ideal by another binomial ideal, even a principal one, is not in general binomial (Exercise 1.15). We will return to these issues later on.

It now seems natural to ask what nice properties the class of trinomial ideals might have, that is, ideals generated by the sum of three terms. One can show, however, that every polynomial ideal can be represented as a trinomial ideal, so there is no hope of being able to say anything substantial about them.

7. Gröbner Bases for Modules.

For some of the applications we will discuss later we need to consider certain types of modules over the polynomial ring $k[x]$, in particular submodules of free modules and their quotient modules. Free modules are just the analog of vector spaces over a field. One difference between fields and more general rings is that rings admit more than just free modules. To simplify notation, we denote $k[x]$ by $A$ in this section. For $m \geq 1$, the Cartesian product

$$A^m = \{(a_1, \ldots, a_m)^t \mid a_i \in A\}$$

is called the free $A$-module of rank $m$. It is an abelian group under coordinate-wise addition, and $A$ acts on it by coordinate-wise multiplication. In this sense, modules are just like vector spaces. The reason we write the elements of $A^m$ as column vectors is so that composition of maps follows the usual convention. In general, an $A$-module is just an abelian group with an $A$-action on it. It is finitely generated, if there exists a finite subset, so that every element is an $A$-linear combination of elements in that subset. A submodule of $A^m$ is a subgroup which is closed under the action of $A$, just as in the vector space case. Quotient modules are also defined just like for vector spaces. The main difference here is that subspaces and quotient spaces do not need to be isomorphic to a free module again. For $A^1 = A$ submodules are just ideals.

In the last section we showed that $A$ is a Noetherian ring. One property that makes Noetherian rings good to work with is that all their finitely generated modules are Noetherian also, that is, every submodule is finitely generated. For our purposes all we need to know is that submodules of $A^m$ are finitely generated. Not surprisingly, the proof of this fact relies heavily on the Hilbert Basis Theorem, and can be found in, e.g., [1, pp. 115ff].

It is possible to develop a theory of Gröbner bases for free $A$-modules and their submodules, quite analogous to the theory we have
constructed here. Such theories can be found in [1, 5]. The free module $A^m$ plays the role of $k[x]$, and submodules play the role of ideals. If $m = 1$, we recover exactly what we have developed so far. In the rest of this section we outline this theory for modules. All we really need to do is to decide what should take the place of the monomials in the polynomial ring case. If it doesn’t matter what the rank of the free module is we will simply denote it by $F$.

Let

$$e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$$

be the canonical $i$-th basis vector for $A^m$, where the 1 is in the $i$-th position. The set $\{e_1, \ldots, e_m\}$ is a basis for $A^m$ in the sense that every element can be written as an $A$-linear combination of these vectors.

**Definition.** Let $F$ be a free $A$-module. A **monomial** in $F$ is an element of the form $m \cdot e_i$, where $m$ is a monomial in $k[x]$, and $e_i$ is a canonical basis vector. A **term** is a monomial multiplied by a constant.

With these definitions it is clear that every element of $F$ can be written uniquely as a sum of terms.

**Definition.** Let $F$ be a free $A$-module. A **term order** on $F$ is a total order $>$ on the terms of $F$ such that, if $m_1, m_2$ are terms of $F$, and $1 \neq m \in A$ is a monomial, then $m_1 > m_2$ implies that

$$mm_1 > mm_2 > m_2.$$ 

To get an idea of how to manufacture term orders for free modules, we give a different way of looking at them. Consider the polynomial ring in $n + m$ variables

$$k[x_1, \ldots, x_n, e_1, \ldots, e_m],$$

and in it the ideal $\langle e_ie_j \rangle$. Let $R$ be the quotient. In $R$ let $I$ be the ideal generated by the images in $R$ of the $e_i$. Then it is straightforward to check that $I \cong A^m$ as $A$-modules. Lifting back to the polynomial ring $k[x, e]$, we can represent $A^m$ by the ideal generated by the $e_i$, with the understanding that elements in $A^m$ correspond to polynomials which are linear in the $e_i$. This is also a way to carry out computations in, e.g., MAPLE. So, one quick way of getting term orders for $A^m$ is to order the two sets of variables $\{x_i\}$ and $\{e_j\}$ separately, choose a term order for each of them, and decide whether $x_i > e_j$ or $x_i < e_j$ for all $i, j$.

For instance, one could make all the $x_i$ larger than all the $e_j$, and choose **lex** order for the $x_i$ and the $e_i$. That is, given two monomials

$$a = m(x)e_i, \quad b = m'(x)e_j,$$
we compare them by first comparing \( m(x) \) to \( m'(x) \), using the term order on \( k[x] \), and break ties with the position they are in. As examples,
\[
x y e_1 < x^2 e_2, \quad x^3 y e_3 < x^3 y e_1.
\]
If we make the \( e_i \) bigger than the \( x_j \), then position is considered first, so the first inequality in the above example would be turned around. These two term orderings are in fact the most commonly used in making computations in free modules.

Now, that we have term orders in place, we can talk about the leading term of an element in the free module \( F \), and we can make all the same definitions as in the polynomial ring case. It turns out that everything, including the Buchberger algorithm, carries over practically verbatim, and we get a Gröbner basis theory for modules. The details will be explored in the exercises.

**Exercises.**

**Exercise 1.1.** Use the division algorithm to show that \( k[x] \) is a principal ideal domain.

**Exercise 1.2.** Show that the only term order on \( k[x] \) is the degree order, given by \( 1 < x < x^2 < \cdots \).

**Exercise 1.3.** Find an example of an ideal
\[
I = (f_1, \ldots, f_m) \subset k[x],
\]
\( f \in I \), and term orders \( <_1 \) and \( <_2 \) such that dividing \( f \) by \( f_1, \ldots, f_m \) leaves a nonzero remainder when using \( <_1 \) and zero remainder when using \( <_2 \).

**Exercise 1.4.** Verify that \( lex, grlex, grevlex \) are indeed term orders.

**Exercise 1.5.** Show that in \( k[x, y] \) \( grlex \) and \( grevlex \) define the same term order.

**Exercise 1.6.** Let \( I \subset k[x, y] \) be the ideal generated by the polynomial
\[
x^5 y + x^4 y^4 + x^4 + x^2 y^5 + xy^2 + x^6 + y.
\]
Find the initial ideals for \( \omega = (1, 1), (1, 2), (0, 0) \). Find all initial ideals of \( I \). Determine those that are monomial ideals.

**Exercise 1.7.** Show that the ordering \( <_\omega \) is indeed a term ordering.

**Exercise 1.8.** Prove Proposition 2.1.
EXERCISE 1.9. Show that there is a unique reduced Gröbner basis for an ideal with respect to a given term order.

EXERCISE 1.10. Let $I \subset k[x]$ be an ideal, and let $<_1$ and $<_2$ be term orders. Let $\{g_1, \ldots, g_r\}$ be a Gröbner basis of $I$ with respect to $<_1$. Suppose that $\text{LT}(g_i)_{<_1} = \text{LT}(g_i)_{<_2}$ for all $i$. Show that $\{g_1, \ldots, g_r\}$ is also a Gröbner basis of $I$ with respect to $<_2$.

EXERCISE 1.11. Show that an ideal $I \subset k[x]$ can only have finitely many distinct reduced Gröbner bases with respect to all possible term orders.

EXERCISE 1.12. Compute a universal Gröbner basis for the ideal $\langle x - y^2, xy - x \rangle \subset Q[x, y]$.

EXERCISE 1.13. Compute a Gröbner basis for the ideal

$I = \langle xyz - 1, x^2 + z^3, x + y + z - 1, y^2 - z^3 \rangle \subset Q[x, y, z],$

using $\text{lex}$, $\text{grlex}$, and $\text{grevlex}$. Change the ordering of the variables to see how it affects the computations.

EXERCISE 1.14. Show that the Hilbert Basis Theorem is equivalent to the statement that every ideal has a finite Gröbner basis.

EXERCISE 1.15. Find an example of two binomial ideals whose intersection is not a binomial ideal.

EXERCISE 1.16. Define the concept of a Gröbner basis for submodules of free modules over a polynomial ring.

EXERCISE 1.17. Show that every submodule of a free module (of finite rank) over a polynomial ring has a Gröbner basis.

EXERCISE 1.18. Formulate a Buchberger algorithm for Gröbner bases for modules.

EXERCISE 1.19. Let $A = k[x, y]$. Compute a Gröbner basis for the submodule of $A^3$ generated by the elements

$$(xy, 0, y^3), (x^2, xy^2, x), (x, 0, y),$$

with respect to two different term orders. Do one of these calculations by hand, as well as with Macaulay, and find a way to do the other one using MAPLE.
A Brief Introduction to Algebraic Geometry

1. Affine Varieties

Polynomial algebra is intimately connected with algebraic geometry, one of the most important, active, and oldest subjects. It is concerned with solution sets of systems of polynomial equations. During the second half of this century algebraic geometry has undergone an explosive development, beginning with the monumental and groundbreaking work of Alexander Grothendieck and his school. Its tremendous power has come at the expense of a huge machinery of very abstract tools, that are far removed from solution sets of polynomial systems. Consequently, it is not easily accessible. Here we are only concerned with very concrete aspects closely tied to polynomial rings and their ideals, namely the theory of affine algebraic varieties. Using the language of geometry, we can better understand some aspects of Gröbner basis theory, and it is also needed in some of the applications. For a deeper introduction to the subject, the reader can consult, e.g., the excellent book by J. Harris [11]. A more detailed introduction to affine varieties is contained in [3].

Definitions. Let $k$ be a field.

1. Let $k^n$ be the cartesian product of $n$ copies of $k$. The set of points in $k^n$ is called affine $n$-space, denoted $\mathbb{A}^n$.
2. Let $f_1, \ldots, f_s \in k[x_1, \ldots, x_n]$ be polynomials. Let

$$V(f_1, \ldots, f_s) = \{a = (a_1, \ldots, a_n) \in \mathbb{A}^n | f_i(a) = 0 \text{ for all } i = 1, \ldots, s\}$$

be the set of solutions of the system $f_i(x) = 0; i = 1, \ldots, s$. Then $V(f_1, \ldots, f_s)$ is called the affine variety defined by $f_1, \ldots, f_s$.

Examples.

1. An important collection of examples of affine varieties are curves in the plane and in 3-space. Let $k = \mathbb{R}$. Then any polynomial function $y = f(x)$ describes an affine variety $V(f(x) - y)$ in real affine 2-space $\mathbb{R}^2$, namely the graph of the function. As another example, the variety $V(xy - 1, x^2 - y)$ consists of the intersection point $(1, 1) \in \mathbb{R}^2$ of the parabola $x^2 = y$ and the hyperbola $y = 1/x$. Finally, the variety
\( V(x^2 + 1) \) is empty, since the equation \( x^2 + 1 = 0 \) has no real solutions. This example shows that the field \( k \) plays an important role. One frequently assumes that the underlying field \( k \) is algebraically closed, which excludes varieties without points, as we shall see (except for trivial cases, such as when \( f \) is a nonzero constant, in which case \( V(f) \) is always empty.)

2. In affine 3-space more interesting examples appear. We still get curves. One of the most useful examples of a curve in 3-space is the so-called twisted cubic, obtained as \( V(y - x^2, z - x^3) \). The reader is well-advised to study this curve carefully, since it is frequently used as an example for a variety of phenomena.

In addition to curves we now also obtain surfaces, for instance surfaces of revolution like \( V(z - x^2 - y^2) \), which is a paraboloid. Other surfaces are the coordinate planes, such as the xy-plane \( V(z) \), as well as more complicated surfaces like \( V(x^2 - y^2 z^2 + z^3) \). Furthermore, we obtain “mixtures” of curves and planes, such as \( V(xz, yz) \subset \mathbb{R}^3 \), which is the union of the \((x, y)\)-plane and the z-axis. One important invariant of a variety is its dimension. The last example shows that it is rather subtle to define the dimension of a variety in general. See [3, Ch. 9]. The reader is encouraged to use the graphing package of MAPLE to visualize these and other varieties in affine 3-space.

For higher dimensions and for fields other than \( \mathbb{R} \) it is more difficult or impossible to obtain geometric information about a variety. This was one of the major stumbling blocks for algebraic geometers around the turn of the century. Proofs tended to rely more and more on intuition, and consequently contained mistakes quite frequently. With the discovery that one could use the newly developed axiomatic abstract algebra to study geometric objects algebraic geometry gained a whole new set of powerful tools to make its arguments rigorous. An excellent account of these developments can be found in the biography of Oscar Zariski [13], one of the key contributors to this expansion of algebraic geometry.

Let \( V = V(f_1, \ldots, f_s) \) be an affine variety. Consider the ideal \( I \subset k[x_1, \ldots, x_n] \) generated by the \( f_i \). Then it is clear that every polynomial in \( I \) vanishes on the points of \( V \). Furthermore, it is clear that if we choose a different set \( \{g_1, \ldots, g_r\} \) of generators for \( I \), then

\[ V(f_1, \ldots, f_s) = V(g_1, \ldots, g_r). \]

We call \( I = I(V) \) the ideal of \( V \). In this way we obtain a mapping from the collection of affine varieties in \( \mathbb{A}^n \) to the collection of ideals in \( k[x_1, \ldots, x_n] \). It is this connection between varieties and ideals of polynomial rings that allows us to compute with varieties.
There is a way to construct a mapping in the other direction. Let \( I \subset k[x] \) be an ideal. Define
\[
\mathbf{V}(I) = \{ a \in A^n | f(a) = 0 \text{ for all } f \in I \}.
\]
Since \( I \) is finitely generated by the Hilbert Basis Theorem, \( \mathbf{V}(I) \) is actually an affine variety. The natural question now arises what this correspondence has to do with the one above. It is clear that if we start with a variety \( V \), then \( \mathbf{V}(I(V)) = V \). First of all, it is easy to see that it is not a one-to-one correspondence. For instance, \( \mathbf{V}(\langle x \rangle) = \mathbf{V}(\langle x^2 \rangle) \), that is, different ideals can define the same variety. Furthermore, any polynomial that has no roots in the field \( k \) will define the empty variety. So, at the very least, if we want to obtain a one-to-one correspondence, then we need to assume that the field is algebraically closed and we need to exclude examples like \( \langle x \rangle \) and \( \langle x^2 \rangle \). This is done by restricting ourselves to so-called radical ideals.

**Definition.** The radical \( \text{rad}(I) \) of an ideal \( I \subset k[x] \) is the ideal
\[
\text{rad}(I) = \{ f \in k[x] | f^m \in I \text{ for some } m \geq 1 \}.
\]
An ideal \( I \) is a radical ideal if \( I = \text{rad}(I) \).

The exact relationship between ideals and varieties is spelled out in the celebrated Hilbert Nullstellensatz. We first deal with the problem of the empty variety.

**Theorem 1.1.** (Weak Hilbert Nullstellensatz) Let \( k \) be an algebraically closed field, and \( I \subset k[x] \) an ideal. If \( I \neq k[x] \), then \( \mathbf{V}(I) \neq \emptyset \).

**Proof.** (taken from [10, p.66f]) We proceed by induction on the number \( n \) of variables. If \( n = 1 \), then \( k[x] \) is a PID, so \( I = \langle f \rangle \) is principal. Then \( \mathbf{V}(I) \) is just the set of roots of the univariate polynomial \( f \). If \( f \) has no roots, then \( f \) is a nonzero constant by the Fundamental Theorem of Algebra, since \( k \) is algebraically closed. Hence \( I = k[x] \).

Now assume \( n > 1 \). Let \( I = \langle f_1, \ldots, f_s \rangle \), and no \( f_i \) is a constant. First assume that \( I \cap k[x_i] = \langle 0 \rangle \) for some \( i \), which we may choose to be \( 1 \). We want to reduce the number of variables by projecting \( I \) onto \( k[x_2, \ldots, x_n] \). This we accomplish by substituting a carefully chosen element of \( k \) for \( x_1 \) in the generators of \( I \). First invert the variable \( x_1 \) and extend \( I \) to \( \tilde{I} \subset k(x_1)[x_2, \ldots, x_n] \), where \( k(x_1) \) denotes the field of rational functions in \( x_1 \). First observe that \( \tilde{I} \) is still a proper ideal. Otherwise, let \( 1 = \sum_j h_j f_j \), with \( h_j \in k(x_1)[x_2, \ldots, x_n] \). Multiplying by the least common denominator, we obtain an \( h \in k(x_1) \) such that
\[
h(x_1) = (h(x_1)h_1(x_1)) f_1 + \cdots + (h(x_1)h_s(x_1)) f_s \in I.
\]
But this contradicts the assumption that \( I \cap k[x_1] = \langle 0 \rangle \).
Let $G = \{g_i\}$ be a (finite) reduced Gröbner basis for $\tilde{I}$. There are only finitely many coefficients from $k(x_1)$ involved in the computation of $G$, which altogether have only finitely many zeros in $k$. Let $a \in k$ be different from any of those zeros. Let

$$G_a = \{g_i(a, x_2, \ldots, x_n)\} \subset k[x_2, \ldots, x_n].$$

Furthermore, $G_a$ is also a reduced Gröbner basis for the ideal that it generates, since the computation in $k(x_1)[x_2, \ldots, x_n]$ projects down to $k[x_2, \ldots, x_n]$ by virtue of the choice of $a$.

Finally, $\langle G_a \rangle \neq k[x_2, \ldots, x_n]$, otherwise $G_a = \{1\}$. But that would imply that the $g_i$ are polynomials involving $x_1$ only, hence are constants in $k(x_1)$. This is a contradiction, since $\tilde{I}$ is a proper ideal. Note also that $f_i(a, x_2, \ldots, x_n) \in \langle G_a \rangle$ for all $i$. We can now use the induction hypothesis to conclude that $V(\langle G_a \rangle) \neq \emptyset$. Hence all elements of $\langle G_a \rangle$, including the $f_i(a, x_2, \ldots, x_n)$, have a common zero $(a_2, \ldots, a_n) \in \mathbb{A}^{n-1}$. This implies that the $f_i$ in $k[x_1, \ldots, x_n]$ have a common zero $(a, a_2, \ldots, a_n) \in \mathbb{A}^n$. This completes the proof under the assumption that $I \cap k[x_i] = \langle 0 \rangle$ for some $i$.

So assume that $I \cap k[x_1] = \langle f_1 \rangle \neq \langle 0 \rangle$. Since $\deg(f_1) > 0$ and $k$ is algebraically closed, $f_1$ has a factor of the form $x_1 - a_1$ with $a_1 \in k$, and $f_1 = (x_1 - a_1) \cdot g$. Then

$$I \subset I + \langle x_1 - a_1 \rangle \neq \langle 1 \rangle.$$ 

The inequality can be seen as follows. Suppose that

$$1 = f + h(x_1 - a_1) \in I + \langle x_1 - a_1 \rangle.$$

Then

$$g = gf + hf_1 \in I,$$

which is a contradiction, since $\deg(g) < \deg(f_1)$. Now consider

$$\tilde{I} \subset k[x_1, \ldots, x_n]/\langle x_1 - a_1 \rangle \cong k[x_2, \ldots, x_n].$$

Then $\tilde{I}$ is a proper ideal, so we can apply the induction hypothesis, and obtain a common zero $(a_2, \ldots, a_n)$ of the polynomials in $\tilde{I}$. Then $(a_1, a_2, \ldots, a_n)$ is a common zero of all polynomials in $I$. This completes the proof.

**Corollary 1.2.** Let $k$ be algebraically closed. The maximal ideals in $k[x_1, \ldots, x_n]$ are precisely the ideals of the form

$$\langle x_1 - a_1, \ldots, x_n - a_n \rangle,$$

for $(a_1, \ldots, a_n) \in k^n$. 

Proof. To begin with, it is clear that ideals of this form are maximal. Furthermore, it is clear that the ideal $I((a_1, \ldots, a_n))$ of polynomials that vanish on $(a_1, \ldots, a_n)$ is precisely the ideal $\langle x_1-a_1, \ldots, x_n-a_n \rangle$.

Now let $M$ be a maximal ideal. By the previous theorem $V(M) \neq \emptyset$, so it contains a point $(a_1, \ldots, a_n)$. Then $M \subset I(a_1, \ldots, a_n) = \langle x_1-a_1, \ldots, x_n-a_n \rangle$, hence they are equal. This completes the proof.

We can now establish the precise relationship between ideals and varieties.

Theorem 1.3. (Strong Hilbert Nullstellensatz) Let $k$ be algebraically closed and $I \subset k[x]$ an ideal. Then $I(V(I)) = \text{rad}(I)$.

Proof. It is straightforward to verify that

\[ \text{rad}(I) \subset I(V(I)). \]

To show the other inclusion, suppose that $f \in I(V(I))$. We now use a little trick. Let $y$ be a new variable, and consider the ideal $J \subset k[x,y]$ generated by $I$ and the polynomial $fy - 1$. If there is $(a_1, \ldots, a_n,a) \in V(J) \subset k^{n+1}$, then $(a_1, \ldots, a_n) \in V(I)$, since $I \subset J$. Hence $f(a_1, \ldots, a_n) = 0$, and so $f(a_1, \ldots, a_n)c - 1 = -1$. This implies that $V(J) = \emptyset$, so that $J = k[x,y]$ by the Weak Nullstellensatz. Thus, there exist $g_i \in k[x,y]$ and $f_i \in I$ such that

\[ 1 = g_1f_1 + \cdots + g_mf_m + g_{m+1}(fy - 1). \]

Viewing this equation as an identity in $y$, it remains true if we substitute $y = 1/f$, to obtain

\[ 1 = g_1(x,1/f)f_1(x) + \cdots + g_m(x,1/f)f_m(x) + g_{m+1}(x,1/f)\left(f(x)/f(x) - 1\right) \]

\[ = \sum_{i=1}^m g_i(x,1/f)f_i(x). \]

After clearing denominators, we can express some power $f^r$ as a linear combination of elements in $I$ with coefficients in $k[x]$, which implies that $f^r \in I$. Thus, $f \in \text{rad}(I)$. This completes the proof.

Corollary 1.4. There is a one-to-one correspondence between radical ideals in $k[x_1, \ldots, x_n]$, that is, ideals who are equal to their radical, and affine varieties in $\mathbb{A}^n$.

It seems rather unfortunate that this correspondence does not extend to all ideals. One point of view one could take, and Grothendieck did, is that the class of affine varieties is simply not rich enough and needs to be extended in order to make a one-to-one correspondence with the class of all ideals possible. This leads to the concept of an affine scheme. We give a brief outline of the construction and leave details to the reader.
Let $k$ be an algebraically closed field, and let $I \subseteq k[x]$ be an ideal. We now construct a “geometric” object associated to $I$ which will play the role of $V(I)$ in the desired correspondence between ideals and geometric objects. Consider the quotient ring $k[x]/I$. Define the spectrum $\text{Spec}(k[x]/I)$ of $k[x]/I$ to be the set of all prime ideals in $k[x]/I$. We define a topology on this set by taking as a basis of closed sets all sets of the form

$$Z(J) = \{ P \in \text{Spec}(k[x]/I) | J \subset P \}$$

for all ideals $J \subseteq k[x]/I$. The resulting topology is called the Zariski topology on $\text{Spec}(k[x]/I)$. (The same construction can of course be carried out for any commutative ring.) Topological spaces of this type are called affine schemes.

The Zariski topology is actually a rather coarse topology, as can be seen by playing around with it a bit. For instance, $\text{Spec}(\mathbb{Z})$ is very instructive. Furthermore, if we choose $J$ to be a maximal ideal, the corresponding closed set $Z(J)$ consists of $J$ alone. Hence maximal ideals correspond to closed points. In fact, they are exactly the one-point sets in $\text{Spec}(k[x]/I)$ that are closed.

The relationship to affine varieties is as follows. Recall that for an algebraically closed field $k$, the maximal ideals of $k[x_1, \ldots, x_n]$ are precisely of the form $(x_1 - a_1, \ldots, x_n - a_n)$ for points $(a_1, \ldots, a_n) \in \mathbb{A}^n$. Let $V(I)$ be the variety of an ideal $I$. If $M = (x_1 - a_1, \ldots, x_n - a_n)$ is a maximal ideal containing $I$, then it is clear that $(a_1, \ldots, a_n) \in V(I)$. Conversely, if $(a_1, \ldots, a_n) \in V(I)$, then $I$ is contained in the corresponding maximal ideal. Thus, we get a one-to-one correspondence between the points of $V(I)$ and the closed points of $\text{Spec}(k[x]/I)$, or, equivalently, the maximal ideals in $k[x]/I$.

If we now consider affine schemes instead of affine varieties, then we indeed get a one-to-one correspondence between all ideals in $k[x]$ and affine schemes. Furthermore, affine schemes retain all the geometric information contained in affine varieties. The main object of study in modern algebraic geometry is more general still, namely general schemes, which are objects that “locally” look like affine schemes. A great introduction to schemes is [6].

2. The Coordinate Ring of an Affine Variety

An important object attached to an affine variety is the collection of polynomial functions from the variety to the field.

**Definition.** Let $V$ be an affine variety. The ring of polynomial functions from $V$ to $k$ is called the coordinate ring of $V$, denoted by $k[V]$. 
Each such function can be represented by a polynomial, but different polynomials can represent the same function on the variety. For instance, if $V = V(x^2y - 1)$ in the affine plane, then the polynomial function $xy : V \to k$ and the function given by $xy + x^2y - 1$ are equal as functions on $V$. As can be easily seen, two polynomials describe the same function on $V(I)$ if and only if they differ by a polynomial in $I$. Thus we obtain an isomorphism

$$k[V(I)] \cong k[x]/I.$$

We have not yet mentioned functions between affine varieties. As in all areas of mathematics, the study of a certain type of object involves in an essential way the study of functions relating those objects.

**Definition.** Let $V \subset \mathbb{A}^n$, $W \subset \mathbb{A}^m$ be affine varieties. A transformation

$$\phi : V \to W$$

is a set mapping from $V$ to $W$ given as follows. For $(a_1, \ldots, a_n) \in V$ let

$$\phi(a_1, \ldots, a_n) = (\phi_1(a_1, \ldots, a_n), \ldots, \phi_m(a_1, \ldots, a_n)),$$

such that each $\phi_i \in k[x_1, \ldots, x_n]$, that is, $\phi$ is composed of coordinate-wise polynomial functions.

The association of the coordinate ring to a variety now extends to transformations of varieties. Every transformation

$$\phi : V \to W$$

induces a ring homomorphism

$$f : k[W] \to k[V].$$

Note that the direction of the mapping changes. This homomorphism is induced as follows. Let

$$w : W \to k$$

be an element of $k[W]$. Then

$$w \circ \phi : V \to W \to k$$

is an element of $k[V]$. Let $f(w) = w \circ \phi$. (For the categorically inclined, this assignment defines a contravariant functor from the category of varieties and transformations to the category of commutative rings and ring homomorphisms.) In the exercises the reader is invited to show that, for a fixed variety $V$, this correspondence induces a correspondence between ideals of the coordinate ring $k[V]$ and subvarieties of $V$, that is, subsets of $V$, which are affine varieties themselves. Furthermore, every homomorphism of coordinate rings which fixes field
coefficients induces a transformation of varieties. As a corollary we get that two varieties are isomorphic (with the obvious definition of isomorphism) if and only if the corresponding coordinate rings are isomorphic (Exercise 2.2). This shows that its coordinate ring is an excellent algebraic model for a variety.

**Exercises.**

**Exercise 2.1.** Show that the union and intersection of affine varieties are again affine varieties.

**Exercise 2.2.** Show that two varieties are isomorphic, that is there exist transformations between them that are inverse to each other, if and only if the corresponding coordinate rings are isomorphic.

**Exercise 2.3.** Let $f \in k[x]$, and $I \subset k[x]$. Show that $f \in \sqrt{I}$ if and only if $(I, 1 - ft) = k[x, t]$. 
CHAPTER 3

Application to the Solution of Polynomial Systems

1. Introduction

The contribution to the solution of polynomial systems is probably the most important application of Gröbner bases, and one of the most active research areas. The most promising approach to solving polynomial systems seems to be a combination of symbolic and numerical methods, with the symbolic part serving as a preprocessor to optimize the numerical methods. Our basic object of study in this chapter is a system of polynomial equations

\[
\begin{align*}
  f_1(x_1, \ldots, x_n) &= 0, \\
  \vdots \\
  f_m(x_1, \ldots, x_n) &= 0,
\end{align*}
\]

with \( f_1, \ldots, f_m \in k[x_1, \ldots, x_n] \). If the \( f_i \) are linear, then one can solve the system by Gauss-Jordan elimination. Here a solution of the system means a basis of the vector space of solutions. For a general system the solution set forms an affine variety, and it is not clear what “solving the system” should actually mean. An important class of systems are those for which this variety has only finitely many points. In this case, “solving the system” just means to list those finitely many points. Such systems are called zero-dimensional, and we will treat them in some detail. In the higher dimensional case, it is sometimes possible to find a particularly good way to describe the solution set, for instance by finding a parametrization.

Since Gauss-Jordan elimination works so well in the linear case, a major objective in the general case is to find a method as close to it as possible. Such a method exists in principle, and is referred to as elimination theory.

2. Elimination Theory

The first method for solving polynomial systems that we discuss is based on the Elimination Theorem we proved earlier. It will allow us to mimic Gauss-Jordan elimination. We first recall its statement.
Theorem 2.1. (Elimination Theorem) Let $I \subset k[x_1, \ldots, x_n]$, and let $<$ be lex order with $x_1 > x_2 > \cdots > x_n$. Let $G$ be a Gröbner basis for $I$ with respect to $<$. Let

$$I_r = I \cap k[x_{r+1}, \ldots, x_n]$$

be the $r$-th elimination ideal of $I$, for $1 \leq r \leq n-1$. Then

$$G_r = G \cap k[x_{r+1}, \ldots, x_n]$$

is a Gröbner basis for $I_r$.

This result shows that a Gröbner basis for $I$ with respect to lexicographic order successively eliminates more and more variables. This suggests the following strategy for solving a polynomial system. Compute a Gröbner basis for the ideal generated by the left-hand sides of the equations with respect to lex order. Then choose the polynomials in $G$ with the smallest number of variables and try to solve them first. Having found partial solutions, attempt to back-substitute in the other equations and extend the partial solutions to solutions of the full system. The following example illustrates this strategy.

Example. Consider the following system:

$$x^2 + y^2 + z^2 - 4 = 0,$$
$$x^2 + 2y^2 - 5 = 0,$$
$$xz - 1 = 0.$$

taken from [3, Ch. 3.1]. Computing a lexicographic Gröbner basis $G$ for the ideal in $\mathbb{Q}[x, y, z]$ generated by the left-hand sides of the equations, with $x > y > z$, gives

$$G = \{2z^3 - 3z + x, -1 + y^2 - z^2, 1 + 2z^4 - 3z^2\}.$$ 

Thus we need to solve the equivalent system

$$2z^3 - 3z + x = 0,$$
$$-1 + y^2 - z^2 = 0,$$
$$1 + 2z^4 - 3z^2 = 0.$$

The last polynomial depends only on $z$ and factors as

$$1 + 2z^4 - 3z^2 = (z - 1)(z + 1)(2z^2 - 1).$$

Hence we obtain the four roots $\pm 1, \pm 1/\sqrt{2}$. Backsubstituting we find that the system has eight solutions

$$(1, \pm \sqrt{2}, 1), (-1, \pm \sqrt{2}, -1), (\sqrt{2}, \pm \sqrt{6}/2, 1/\sqrt{2}), (-\sqrt{2}, \pm \sqrt{6}/2, -1/\sqrt{2}).$$

Here we have chosen a system that is particularly nice for several reasons. Firstly, it is easy to find the roots of the last polynomial. And,
secondly, all the partial solutions extend to solutions of the whole system. Neither of these phenomena happens in general.

We discuss the second one first. Consider the system

\[ xy - 1 = 0, \]
\[ xz - 1 = 0. \]

Computing the first elimination ideal of \( I = \langle xy - 1, xz - 1 \rangle \) we obtain

\[ I_1 = I \cap k[y, z] = \langle y - z \rangle. \]

The equation \( y - z = 0 \) has solution set \( \{(a, a) | a \in k\} \). If we now try to extend these partial solutions to solutions of the initial system, we get that \((a, a)\) extends to \((1/a, a, a)\) for all \(a \neq 0\). Hence, the partial solution \((0, 0)\) does NOT extend to a global solution. The following theorem gives a criterion for partial solutions to extend.

**Theorem 2.2.** *(Extension Theorem, [3, Ch. 3.1, Theorem 3.1])** Let \( k \) be an algebraically closed field, and let \( I = \langle f_1, \ldots, f_s \rangle \subset k[x_1, \ldots, x_n] \). Let \( I_1 \) be the first elimination ideal of \( I \). For each \( 1 \leq i \leq s \) write \( f_i \) in the form

\[ f_i = g_i(x_2, \ldots, x_n)x_1^{N_i} + \text{terms in which } x_1 \text{ has degree } < N_i, \]

where \( N_i \geq 0 \), and \( g_i \in k[x_2, \ldots, x_n] \) is nonzero. Suppose that \((a_2, \ldots, a_n) \in V(I_1)\) is a partial solution. If \((a_2, \ldots, a_n) \notin V(g_1, \ldots, g_s)\), then there exists \( a_1 \in k \) such that \((a_1, \ldots, a_n) \in V(I)\).

The proof of this theorem is quite involved and can be found in [3]. Some observations are in order. First of all, note that \( k \) is assumed to be algebraically closed. To see that this assumption is needed consider the system

\[ x^2 - y = 0, \]
\[ x^2 - z = 0. \]

over \( \mathbb{R} \). Eliminating \( x \), we obtain the equation \( y - z = 0 \), as before. If we now choose a partial solution \((a, a)\) with \( a < 0 \), then it is clear that this solution does not extend. It does extend, however, if we let our field be \( \mathbb{C} \).

We have also seen that the condition \((a_2, \ldots, a_n) \notin V(g_1, \ldots, g_s)\) is needed. Finally, note that the ideal \( \langle g_1, \ldots, g_s \rangle \) depends on the choice of \( f_i \). So we might want to choose the \( f_i \) so that \( V(g_1, \ldots, g_s) \) is as small as possible. Details of all these observations are discussed in [3].

The other issue that needs to be discussed is problems related to finding numerical approximations of roots of univariate polynomials. In most cases, it is impossible to give an algebraic or precise numerical...
value for the roots of such a polynomial. The reader will find it very easy to construct examples. Most randomly constructed polynomials will tend to be irreducible. A frequently used numerical method for root approximation is Newton's Method. First choose some initial approximation $x_0$ to a root of $f(x) = 0$. Then construct a sequence of numbers

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)},$$

where $f'$ is the usual derivative of $f$ in the sense of calculus. In many cases this sequence converges very quickly to a root of $f$, and one can choose as good an approximation as needed.

If we now want to solve a system by elimination and approximate the roots of one of its polynomials, then the following problem can occur. When we substitute the approximate values into the other equations, the resulting system is also only an approximation of the original one. But many systems are very sensitive to small perturbations of the coefficients. Here is one of the more pathological examples. Let

$$f(x) = (x + 1)(x + 2) \cdots (x + 20) = x^{20} + 210x^{19} + \cdots + 20!$$

This polynomial has the roots $x = -1, -2, \ldots, -20$. The reader is now invited to perturb one of the coefficients just a little bit, preferably those of higher powers of $x$. The result will likely be a polynomial that has a completely different set of roots, in particular it can have complex ones. Also, the number of roots may change, from lots of them to none, for instance. More details about numerical issues related to elimination theory can be found in [4, Ch. 2.1].

There does exist an alternative to using numerical approximations. Suppose we have a system of polynomial equations which has all rational coefficients, and we are interested in finding all real roots of the system. For applications this is a very important situation. Suppose further that we have computed a Gröbner basis for the corresponding ideal with respect to lcx order and want to apply elimination theory to solve the system. Rather than back substituting an approximation for a root we can extend the ground field over which we are working by formally adjoining a root to $\mathbb{Q}$.

As an elementary example, if our Gröbner basis contains the polynomial $x^3 - 2$, then rather than calculating with a numerical approximation of $\sqrt[3]{2}$, we can simply substitute a variable $\alpha$ for $x_1$ and carry out the subsequent calculations using $\alpha$ and $\alpha^2$ and the relation $\alpha^3 = 2$. That is, we consider our polynomial to live in the extension field

$$\mathbb{Q}(\sqrt[3]{2}) = \mathbb{Q}(\alpha).$$
For details see [4].

3. Zero-dimensional Systems

There are special tools available for so-called zero-dimensional systems, that is, systems which have finitely many solutions. First we give a characterization of systems with finitely many solutions in terms of the quotient ring of \( k[\mathbf{x}] \) with respect to the associated ideal. We assume again that the field \( k \) is algebraically closed.

**Theorem 3.1.** Let \( f_1 = \cdots = f_s = 0 \) with \( f_i \in k[\mathbf{x}] \), and let \( I = (f_1, \ldots, f_s) \). Let \( G = \{g_1, \ldots, g_r\} \) be the reduced Gröbner basis of \( I \) with respect to lex order and the variable ordering \( x_1 < \cdots < x_n \). The following are equivalent.

1. The variety \( \mathbf{V}(I) \) contains only finitely many points;
2. for each \( i = 1, \ldots, n \) there exists \( j \) such that \( \text{LT}(g_j) = x_i^{m_i} \) for some \( m_i \in \mathbb{N} \);
3. the dimension of the \( k \)-vector space \( k[\mathbf{x}]/I \) is finite.

**Proof.** (1) \( \Rightarrow \) (2). Suppose that the variety of \( I \) is finite. If it is empty, then \( I \) is the unit ideal by the Weak Nullstellensatz. Then \( G = \{1\} \), and (2) is trivially satisfied. So we can assume that \( \mathbf{V}(I) \neq \emptyset \). Fix an \( i \in \{1, \ldots, n\} \). We will show that \( I \) contains a polynomial \( f \) whose leading term is a power of \( x_i \). This will then imply (2). Let \( a_1, \ldots, a_t \) be the distinct \( i \)th coordinates of the points in \( \mathbf{V}(I) \). Then

\[
\begin{align*}
  f &= (x_i - a_1) \cdots (x_i - a_t) \in k[x_i] \subset k[\mathbf{x}]
\end{align*}
\]

vanishes at all points of \( \mathbf{V}(I) \), hence

\[
  f \in I(\mathbf{V}(I)) = \text{rad}(I),
\]

by the Strong Nullstellensatz. Thus, some power of \( f \) lies in \( I \), which has as leading term some power of \( x_i \).

(2) \( \Rightarrow \) (3). We have seen that the set of standard monomials form a vector space basis for \( k[\mathbf{x}]/I \). (2) implies that there are only finitely many standard monomials.

(3) \( \Rightarrow \) (1). It is enough to show that for each \( i = 1, \ldots, n \) there are only finitely many distinct values for the \( i \)th coordinate of the points in \( \mathbf{V}(I) \). By assumption, the set of monomials \( 1, x_i, x_i^2, \ldots \) are \( k \)-linearly dependent modulo \( I \). This implies that \( I \) contains a polynomial \( \sum c_j x_i^j \), which can have only finitely many roots. Since the \( i \)th coordinate of each point in \( \mathbf{V}(I) \) is a root of this polynomial, this proves (1).

An ideal \( I \) satisfying any of the equivalent conditions of this theorem is called a zero-dimensional ideal.
COROLLARY 3.2. Let $I$ be a zero-dimensional ideal and $G$ the reduced Gröbner basis of $I$ with respect to lex order and the monomial ordering $x_1 < \cdots < x_n$. Then we can order the elements $g_1, \ldots, g_r$ of $G$ so that $g_1$ contains only the variable $x_1$, $g_2$ contains only the variables $x_1$ and $x_2$ and its leading term is a power of $x_2$, etc.

Thus, for zero-dimensional systems elimination theory provides us with a method to find solutions provided we can compute a lex order Gröbner basis and can solve the numerical issues that appear in root approximations.

One can in fact show a bit more, namely that the number of zeros of the system is exactly equal to the vector space dimension of $k[x]/I$, counted with appropriate multiplicities.

One of the very big drawbacks of elimination theory is that it requires the computation of a Gröbner basis with respect to lex order. This is often extremely expensive and the resulting basis is not very nice, that is, has many polynomials in it, of high degrees. Some computational savings comes from an algorithm that allows the conversion of a Gröbner basis with respect to one term order into one with respect to any other term order. Details can be found in [4, Chapter 2.3].

Another method for dealing with real root finding for zero-dimensional systems is described in [2, Chapter 8.8]. It uses so-called Sturm sequences.

**Exercises.** Let $R$ denote the polynomial ring $k[x]$.

**EXERCISE 3.1.** Show that if $f \in R$ is a non-zero-divisor modulo $I$, then $((I \cdot t, (1-t) \cdot f) \cap R \cdot f^{-1}) = I$.

**EXERCISE 3.2.** Show that $((I \cdot t, (1-t) \cdot f) \cap R) \cdot f^{-1} = I : (f)$.  

**EXERCISE 3.3.** Show that $(I, 1- ft) \cap R = \bigcup_{n \geq 1} (I: f^n)$.  

**EXERCISE 3.4.** Show that if $J = (g_1, \ldots, g_n)$ then $I : J = \bigcap_{i=1}^n (I : g_i)$.  

**EXERCISE 3.5.** Show that if $J = (g_1, \ldots, g_n)$ then  

$I : J = (I \cdot R[t] : (g_1 + g_2 \cdot t + g_3 \cdot t^2 + \cdots g_n \cdot t^{n-1})) \cup R$.

**EXERCISE 3.6.** Show for $I, J$ ideals of $R$ that  

$I = J$ if and only if $I : J = (1)$.  

CHAPTER 4

Application to Cryptography

1. Introduction

One of the most important aspects of cryptography is the design of so-called public-key cryptosystems. The most famous of these is the RSA system, used extensively for internet commerce and other forms of information transmission. In “classical” cryptography there is one key which is used for both encryption and decryption of messages. For many situations this setup is quite inconvenient. For instance, encrypting credit card information during internet purchases cannot feasibly rely on a crypto scheme, which requires a prior key exchange. What is needed is an encryption scheme that comes with two keys, a public one for encryption and a secret one for decryption. The first such scheme, the RSA system, was discovered in 1976. The basic idea is to use a function for encryption whose inverse function is very hard to compute without additional information. This additional information is the secret key. RSA uses exponentiation modulo a large carefully chosen exponent as the encryption function. The public key contains as one piece of information the product of two large prime numbers. In order to decrypt a message it is necessary to know the two prime factors. Since known algorithms for factoring integers are computationally very expensive, RSA is believed to be secure for the time being.

Like factoring of integers, any other “hard” problem can in principle be used as the basis for a public-key cryptosystem. One important source of such problems is of course combinatorics. Problems relevant for our purposes are the ideal membership problem and the problem of solving nonlinear polynomial systems. In this chapter we describe a type of public-key system which is combinatorial, but whose security is essentially based on the difficulty of these two problems. At present it is unknown how secure systems of this type actually are. Our exposition follows [12, Chapter 5], which the reader is encouraged to consult for more details.
2. The Polly Cracker System

This system is a general public-key cryptosystem based on the ideal membership problem. It has instantiations as a variety of combinatorial systems. Let $k$ be a finite field, and let $\{x_1, \ldots, x_n\}$ be a set of variables. It is customary in this subject to send messages between A(lice) and B(ob). A message in this case consists of an element $m$ in $k$, which Bob wants to send to Alice, using her public encryption key. She then decrypts the message using her secret key $S$, which consists of an arbitrarily chosen vector $y \in k^n$. Alice’s public key is a finite set $P = \{q_j\}$ of polynomials in $k[x]$ such that

$$q_j(y) = 0$$

for all $j$. Bob uses this public key $P$ as follows. He generates an element

$$p = \sum h_j q_j \in \langle \{q_j\} \rangle \subset k[x]$$

and send Alice the polynomial

$$c = p + m.$$  

Alice then decrypts the ciphertext polynomial $c$ by evaluating at $y$:

$$c(y) = p(y) + m = m.$$  

Note that Bob only needs to choose random polynomials $h_j$ in order to generate $p$. Likewise, it is easy for Alice to generate her public and private keys. She simply chooses a random $y$ and random polynomials $\tilde{q}_j$, and sets

$$q_j = \tilde{q}_j - \tilde{q}_j(y).$$

Of course, in order to increase the security of her system, these keys must be chosen very carefully. An important feature of this scheme is that the encryption is probabilistic rather than deterministic, since Bob has many choices for the $h_j$, and so the same plain text message may be encoded in many ways into a cyphertext message.

The problem faced by an eavesdropper is to reconstruct $m$ from the public key $P = \{q_j\}$ and the intercepted polynomial $c$. What he knows is that there is a field element, namely $m$, such that $c - m$ is in the ideal generated by the $q_j$. Since the field is finite, he must therefore solve a finite number of ideal membership problems. Alternatively, he must find a common zero of the $\{q_j\}$.

3. A Combinatorial Version of Polly Cracker

We describe here several instantiations of Polly Cracker as combinatorial problems.
1. **Graph 3-Coloring**. The public key for this cryptosystem is a graph $G$ with vertex set $V$ and edge set $E$. The private key is derived from a 3-coloring of $G$, that is, an assignment of one of three colors to each vertex so that no two adjacent vertices are assigned the same color.

Define a collection $B = B(G)$ of polynomials in the variables

$$\{x_{v,i} | v \in V, 1 \leq i \leq 3\}$$

as follows. Let $B = B_1 \cup B_2 \cup B_3$, where

$$B_1 = \{x_{v,1} + x_{v,2} + x_{v,3} - 1 | v \in V\};$$

$$B_2 = \{x_{v,i}x_{v,j} | v \in V, 1 \leq i < j \leq 3\};$$

$$B_3 = \{x_{u,v}x_{v,i} | uv \in E, 1 \leq i \leq 3\}.$$

Given the 3-coloring constituting the private key, we can construct a point $y$ in the zero-set of $B$ by setting the variable $x_{v,i}$ equal to 1 if the vertex $v$ has color $i$ and 0 otherwise. One can show (Exercise 4.2) that the zero-set of $B$ is nonempty if and only if $G$ has a 3-coloring. Other examples can be found in [12, p. 107].

See [1, pp. 102-105] for a method to test whether a given graph is 3-colorable, using computational algebra.

4. **A Generalization**

A generalization of the Polly Cracker system can be constructed as follows. Alice chooses as private key a Gröbner basis $G$ of an ideal $I \subset k[x]$. Let $M$ be a set of representatives for the elements of $k[x]/I$, which cannot be reduced modulo $G$, e.g., $M$ could be the set of standard monomials. Suppose that $M$ is publicly known and comprises the set of message units. That is, a message to be encrypted is an element of $M$.

As an example, let $y = (y_1, \ldots, y_n) \in k^n$ be a secretly chosen point, and let

$$I = \langle x_1 - y_1, \ldots, x_n - y_n \rangle \subset k[x_1, \ldots, x_n].$$

then the generators of $I$ form a Gröbner basis (check this) and the set $M$ is just the set of constant polynomials, that is, the elements of $k$.

Next, Alice chooses a set $P = \{q_j\}$ of polynomials in $I$, which is her public key. Let $J$ be the ideal they generate. To send a message $m \in M$ to Alice, Bob randomly chooses an element

$$p = \sum h_j q_j \in J$$

and sends Alice the cyphertext polynomial

$$c = p + m.$$
Alice deciphers the message by reducing $c$ modulo the Gröbner basis $G$. The special case of the example above, where $I = \langle x_1 - y_1, \ldots, x_n - y_n \rangle$, gives the earlier Polly Cracker system.

It can be shown (Exercise 4.1) that in order to break this system, it is sufficient to find a Gröbner basis of the ideal $J$. Thus it is important to choose the $q_j$ in such a way that computing such a Gröbner basis is prohibitively expensive. As discussed in [12], at this point it is unclear whether Polly Cracker is a secure system, and some skepticism has been expressed by members of the computational algebra community. Security considerations and possible attempts to break it are discussed in [12] and [9].

**Exercises.**

**Exercise 4.1.** Show that the generalized Polly Cracker system can be broken by computing a Gröbner basis for the ideal $J$.

**Exercise 4.2.** Show that the zero set of $B$ is nonempty if and only if $G$ has a 3-coloring.
Bibliography